

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL – UFRGS

DEPARTAMENTO DE GENÉTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR – PPGBM

**RELAÇÕES FILOGENÉTICAS NO GÊNERO *DROSOPHILA* (DIPTERA, DROSOPHILIDAE):**

**UMA ABORDAGEM MOLECULAR**

LIZANDRA JAQUELINE ROBE

Tese de Doutorado submetida ao Programa de  
Pós-Graduação em Genética e Biologia Molecular  
da UFRGS como requisito parcial  
para a obtenção do grau de Doutor em Ciências.

Orientadora: Profa. Dra. Vera Lúcia da Silva Valente Gaiesky

Co-Orientador: Prof. Dr. Elgion Lúcio da Silva Loreto

Porto Alegre, março de 2008.

O presente trabalho foi desenvolvido junto ao Laboratório 210 da Universidade Federal do Rio Grande do Sul, com o apoio do LabDros da Universidade Federal de Santa Maria e do Laboratório de Genoma Estrutural e Funcional do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul, e com o financiamento do CNPq, FAPERGS, CAPES e PROPESQ-UFRGS.

*“Nada na biologia faz sentido, exceto à luz da evolução.”*

(Theodosius Dobzhansky)

*“A classificação por descendência não pode ser inventada por biólogos,  
ela pode apenas ser descoberta.”*

(Theodosius Dobzhansky)

Dedico este trabalho ao Sandro e aos meus pais,  
pelo amor, pela confiança e pelo apoio incondicional.

## AGRADECIMENTOS

---

À minha querida orientadora Verinha, pela orientação, pelas conversas e conselhos, pelo excelente ambiente de trabalho que sempre propiciou a todos os seus alunos, e, principalmente, pelo apoio e dedicação à minha formação.

Ao Elgion, pela orientação, pela confiança e pelos conselhos e ensinamentos constantes ao longo dos últimos 10 anos. Agradeço também pelo empenho em aprimorar minha formação.

Aos Profs. Arnaldo Zaha e Henrique Ferreira, pela disponibilização do eletroporador e da mão-de-obra amiga (leia-se Fabiano Torres, Larissa Heinzemann, Newton Vidal e Paulo Marcos Pinto) junto ao Laboratório de Genoma Estrutural e Funcional.

Aos Profs. responsáveis pelo Laboratório de Imunogenética e pelo Laboratório de Genética Vegetal, pelos incontáveis auxílios.

Ao Elmo e à Ellen, pela ajuda de sempre.

Ao Fabiano, à Larissa, ao Newton e ao Paulo, que além de grandes amigos, me auxiliaram incondicionalmente durante a execução do trabalho prático.

À Paloma, por me ajudar com as inúmeras reações de seqüenciamento e, depois de tudo, ainda estar disposta a se tornar minha amiga.

Ao Hermes, ao Jonas, ao Marco e à Sabrina, pela ajuda com a identificação das moscas e com a obtenção dos estoques. Ao Hermes e ao Marco, em especial, pelos agradáveis momentos de discussão.

À Nina, por ter me suportado durante toda sua iniciação científica e ainda continuar sendo minha amiga depois de tudo... Pelas conversas amigas, pelos conselhos, pela ajuda nas correções de inglês, e por tornar a elaboração de um artigo uma etapa mais divertida.

À Adriana Ludwig pelo DNA do grupo *flavopilosa*, por ouvir atentamente meus insólitos desabafos, pela ajuda nas repicagens, pelos momentos de discussão, mas também de diversão e, principalmente, pela amizade.

À Maríndia, pelos seus conselhos, pelo seu ombro amigo, pela ajuda e pela companhia com as hibridizações de embriões.

À Ju, pelos inúmeros auxílios, pela amizade, pelas conversas e pelas festas.

Ao Ronaldo e ao Daniel, pela eterna amizade, daquelas que persistem a ventos e trovoadas. Obrigada por nossas longas conversas e pela ajuda nos momentos de aperto.

À Rosane, pelas sábias conversas e conselhos. A Mônica, minha querida amiga, sempre disposta a ajudar.

À Adri Sassi e à Marícia, por propiciarem momentos de alegria durante o trabalho, pelo coleguismo e pela companhia no laboratório.

À Ana, pelas linhagens e pela ajuda com a procura por papers nestes últimos meses de sufoco.

Aos demais colegas de laboratório: Carol, Cláudia, Dirleani, Gisele, Grazia, Juliana, Luís e Mário, por tornarem nosso laboratório um ambiente cada vez mais agradável.

Às amigas de sempre: Deise, Giovana, Giselda, Liliana, Luciane, Raquel, Vivian; sua torcida também foi fundamental.

À família Santos, obrigada pela confiança e pelo apoio.

Ao Elvis, por sua dedicação em me ajudar sempre. À minha irmã, por seu incentivo e por sua torcida. À minha Vózinha, eterna fonte de inspiração.

Aos meus pais, pelo seu amor, pelo seu exemplo, por estarem sempre dispostos a me socorrer nos momentos de aperto.

Ao Charles, pelo companheirismo e pelo carinho sempre presentes.

Ao Sandro, pelo amor, pela compreensão, pelo incentivo e pela dedicação.

*Agradeço também àqueles a quem não mencionei, mas que se fizeram presentes, de uma forma ou de outra, no decorrer dos últimos quatro anos.*

## SUMÁRIO

---

Resumo .....	12
Abstract.....	13
<b>Capítulo I – Introdução.....</b>	<b>14</b>
A Ordem Diptera .....	14
A Família Drosophilidae .....	15
1 – O gênero <i>Drosophila</i> .....	18
1.1 – O subgênero <i>Sophophora</i> .....	21
1.1.1 – Grupo <i>melanogaster</i> .....	23
1.1.2 – Grupo <i>obscura</i> .....	27
1.1.3 – Grupo <i>saltans</i> .....	29
1.1.4 – Grupo <i>willistoni</i> .....	33
1.2 – O subgênero <i>Drosophila</i> .....	39
1.2.1 – Seção <i>virilis-repleta</i> .....	41
1.2.1.1 - Grupo <i>annulimana</i> .....	43
1.2.1.2 – Grupo <i>bromeliae</i> .....	45
1.2.1.3 – Grupo <i>canalineae</i> .....	45
1.2.1.4 – Grupo <i>carsoni</i> .....	46
1.2.1.5 – Grupo <i>coffeata</i> .....	46
1.2.1.6 – Grupo <i>dreyfusi</i> .....	46
1.2.1.7 – Grupo <i>flavopilosa</i> .....	46
1.2.1.8 – Grupo <i>melanica</i> .....	48
1.2.1.9 – Grupo <i>mesophragmatica</i> .....	49
1.2.1.10 – Grupo <i>nannoptera</i> .....	50

1.2.1.11 – Grupo <i>peruviana</i> .....	51
1.2.1.12 – Grupo <i>polychaeta</i> .....	51
1.2.1.13 – Grupo <i>repleta</i> .....	52
1.2.1.14 – Grupo <i>robusta</i> .....	55
1.2.1.15 – Grupo <i>virilis</i> .....	55
1.2.1.16 – Grupo <i>tumiditarsus</i> .....	57
1.2.2 – Seção <i>quinaria-tripunctata</i> .....	57
1.2.2.1 – Grupo <i>calloptera</i> .....	58
1.2.2.2 – Grupo <i>cardini</i> .....	59
1.2.2.3 – Grupo <i>funeris</i> .....	64
1.2.2.4 – Grupo <i>guarani</i> .....	65
1.2.2.5 – Grupo <i>histrion</i> .....	67
1.2.2.6 – Grupo <i>immigrans</i> .....	68
1.2.2.7 – Grupo <i>macroptera</i> .....	69
1.2.2.8 – Grupo <i>pallidipennis</i> .....	70
1.2.2.9 – Grupo <i>quinaria</i> .....	70
1.2.2.10 – Grupo <i>rubrifrons</i> .....	71
1.2.2.11 – Grupo <i>sticta</i> .....	72
1.2.2.12 – Grupo <i>testacea</i> .....	72
1.2.2.13 – Grupo <i>tripunctata</i> .....	72
1.2.3 – Drosofilídeos hawaianos .....	77
1.3 – O subgênero <i>Dorsilopha</i> .....	81
A importância do estabelecimento das relações filogenéticas de Drosophilidae .....	82
1 – Estudos filogenéticos com base em caracteres morfológicos.....	84
1.1 – Throckmorton (1962 e 1975).....	84



1.2 – Grimaldi (1990) .....	91
2 – A sistemática molecular de Drosophilidae .....	93
3 – Os problemas na filogenia de Drosophilidae.....	95
3.1 – Grupo(s)-externo(s) ao gênero <i>Drosophila</i> .....	95
3.2 – Posicionamento de <i>Zaprionus</i> .....	96
3.3 – Posicionamento de <i>Hirtodrosophila</i> .....	97
3.4 – Posicionamento de <i>Liodrosophila</i> .....	98
3.5 – Posicionamento de <i>Samoaia</i> .....	99
3.6 – Posicionamento de <i>Dorsilopha</i> .....	100
3.7 – Monofilia e posicionamento do subgênero <i>Sophophora</i> .....	101
3.8 – Filogenia do subgênero <i>Drosophila</i> .....	103
Objetivos.....	113
<b>Capítulo II – Molecular phylogeny of the subgenus <i>Drosophila</i> (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: a nuclear versus mitochondrial gene approach .....</b>	<b>115</b>
Abstract.....	116
1. Introduction .....	116
2. Materials and methods.....	117
3. Results and discussion.....	120
4. Conclusion.....	129
Acknowledgments .....	131
References .....	131
<b>Capítulo III – On the phylogenetic relationships of the subgenus <i>Drosophila</i> (Diptera, Drosophilidae): New evidences from the <i>Amd</i>, <i>Ddc</i> and <i>hunchback</i> gene partitions .....</b>	<b>134</b>

Abstract.....	135
1. Introduction .....	136
2. Materials and methods.....	138
3. Results .....	141
4. Discussion.....	151
Acknowledgments .....	157
References .....	157
Figure Legends .....	164
Tables 1-7 .....	167
Figures 1-5.....	178
Supplementary Material – Figure 1S.....	183
<b>Capítulo IV – Phylogeny of the <i>Drosophila mesophragmatica</i> group (Diptera, Drosophilidae): an example of Andean evolution.....</b>	<b>184</b>
Abstract.....	185
Introduction .....	186
Materials and methods.....	188
Results .....	190
Discussion.....	191
Acknowledgments .....	196
References .....	197
Figure Legends .....	201
Tables 1-3 .....	202
Figures 1-2.....	205
<b>Capítulo V – Taxonomic boundaries and the phylogenetic relationships within the <i>Drosophila willistoni</i> subgroup .....</b>	<b>207</b>

Abstract.....	208
Introduction .....	209
Materials and methods.....	212
Results .....	216
Discussion.....	222
Acknowledgments .....	228
References .....	228
Figure Legends .....	234
Tables 1-5 .....	236
Supplementary Material – Table 1S .....	241
Figures 1-4.....	242
<b>Capítulo VI – Conclusões Gerais e Perspectivas .....</b>	<b>247</b>
<b>Referências Bibliográficas .....</b>	<b>260</b>
<b>Anexo I.....</b>	<b>286</b>
Os marcadores utilizados.....	286
O gene <i>Adh</i> ( <i>álcool desidrogenase</i> ) .....	287
Os genes parálogos <i>Amd</i> ( <i>alfa metildopa</i> ) e <i>Ddc</i> ( <i>dopa descarboxilase</i> ).....	289
O gene <i>Hb</i> ( <i>hunchback</i> ).....	294
O gene <i>Per</i> ( <i>period</i> ).....	296
Os genes mitocondriais <i>COI</i> ( <i>citocromo oxidase I</i> ) e <i>COII</i> ( <i>citocromo oxidase II</i> ).....	298
Referências Bibliográficas.....	300

## RESUMO

---

O gênero *Drosophila* abrange pelo menos 1.149 espécies, subdivididas em oito subgêneros, muitas das quais vêm desempenhando um papel fundamental no desenvolvimento do conhecimento biológico, principalmente nos âmbitos genético e evolutivo. Toda esta diversidade, entretanto, ao mesmo tempo em que gera inúmeras oportunidades de pesquisa, impõe uma série de desafios taxonômicos e filogenéticos. Desta forma, apesar de *Drosophila* compor um dos organismos experimentais mais bem estudados, muitas questões acerca de sua evolução permanecem irresolvidos ao longo do tempo. A presente tese foi, portanto, desenvolvida com o objetivo de auxiliar na resolução de aspectos particulares relacionados à história evolutiva deste complexo grupo de espécies. Neste sentido, os **Capítulos II e III** foram designados com o intuito de elucidar as relações filogenéticas entre espécies e grupos tradicionalmente situados no subgênero *Drosophila*, com ênfase naqueles de distribuição primariamente Neotropical. A inclusão de outros gêneros pertencentes à família Drosophilidae levou, porém, a definições parafiléticas para o gênero e para o subgênero *Drosophila*. Mesmo assim, níveis adicionais de resolução foram acrescentados à filogenia do subgênero *Drosophila*, que se apresentou subdividido em duas (**Capítulo II**) ou três (**Capítulo III**) grandes radiações. Os **Capítulos IV e V**, por outro lado, buscaram a inferência de cenários evolutivos mais restritos. No **Capítulo IV** o foco foi um pequeno grupo de espécies pertencente ao subgênero *Drosophila*, o grupo *mesophragmatica*, para o qual se obteve descrições evolutivas confiáveis. No **Capítulo V**, por outro lado, o grupo *willistoni*, pertencente ao subgênero *Sophophora* teve sua complexidade evolutiva mais uma vez reiterada. Em cada um destes casos, entretanto, novos estudos são incentivados para que um dia realmente nos aproximemos da história evolutiva verdadeira do gênero *Drosophila*.

## ABSTRACT

---

The genus *Drosophila* encompasses at least 1.149 species subdivided into eight subgenera, many of which presented a central role in the development of the biological knowledge, mainly in the genetical and evolutionary areas. However, at the same time that this diversity provides numberless opportunities of research, it imposes a set of taxonomic and phylogenetic challenges. Thus, although *Drosophila* composes one of the best studied experimental organisms, many questions concerning its evolution remain unsolved through time. This thesis was developed with the aim of helping to solve particular questions related to the evolutionary history of this complex group of species. In this sense, **Chapters II and III** were designed with the aim of elucidating the phylogenetic relationships within and between species groups traditionally placed in the *Drosophila* subgenus, with an emphasis in those mainly Neotropical in distribution. The inclusion of other Drosophilidae genera, nevertheless, led the *Drosophila* genus and subgenus entirely paraphyletic. Even so, improved levels of resolution were provided to the subgenus *Drosophila* phylogeny, which was subdivided into two (**Chapter II**) or three (**Chapter III**) major radiations. **Chapters IV and V**, on the other hand, addressed more restricted evolutionary scenarios. In **Chapter IV** the focus was a small group of species included in the *Drosophila* subgenus, the *mesophragmatica* group, for which a confident evolutionary description was attained. In **Chapter V**, otherwise, the *willistoni* group, included in the *Sophophora* subgenus, had its evolutionary complexity once again reiterated. In each of these cases, nevertheless, new studies are encouraged in order to approach the genus *Drosophila* real evolutionary history.

## CAPÍTULO I

---

### INTRODUÇÃO

#### A ORDEM DIPTERA

Dípteros são insetos holometábolos amplamente diversificados, cuja origem é postulada para o Permiano, há cerca de 250 milhões de anos atrás (Yeates & Wiegmann, 2005). Sua morfologia, entretanto, é tão especializada que retém poucas pistas de sua história filogenética, de forma que existem controvérsias acerca de seu posicionamento frente às demais dez ordens de Holometabola (Whiting, 2005). Tradicionalmente, Diptera é considerada membro da Super-Ordem Mecopterida, que também inclui as ordens Trichoptera (friganeídeos), Lepidoptera (borboletas e mariposas), Siphonaptera (pulgas e bichos-de-pé), Mecoptera (moscas-escorpião) e, talvez, Strepsiptera (estrepssípteros parasitas). Segundo Whiting (2005), Diptera juntamente a estas três últimas ordens constituem, ainda, o clado Antliosphora, e, embora questões referentes ao correto grupo-irmão de Diptera permaneçam em aberto, evidências recentes sugerem que Strepsiptera componha sua ordem mais próxima.

Os dípteros encontram-se, certamente, entre os grupos mais diversos, anatomicamente variáveis e ecologicamente inovativos do planeta (Yeates & Wiegmann, 1999). Com mais de 150.000 espécies descritas, e perfazendo 10-15% das espécies animais conhecidas, os dípteros são classificados em 150 famílias e cerca de 10.000 gêneros (Yeates & Wiegmann, 2005). Embora a taxonomia de Diptera seja alvo constante de debates, uma versão frequentemente aceita divide a ordem da seguinte forma (Wheeler, 1981; Thompson, 2005):

Ordem Diptera

Subordem Nematocera

Subordem Brachycera

Subordem Cyclorrhapha

Divisão Aschiza

Divisão Schizophora

Seção Calyptratae

Seção Acalyptratae

Embora Diptera seja uma dos grupos mais bem estudados filogeneticamente [em grande parte devido aos trabalhos de Hennig (revisão em Meier, 2005)], trabalhos recentes têm demonstrado o caráter merofilético de grande parte de seus táxons. Este é o caso, por exemplo, dos agrupamentos Nematocera, Brachycera e Aschiza, conforme esquema acima apresentado (Yeates & Wiegmann, 1999; Yeates & Wiegmann, 2005).

#### **A FAMÍLIA DROSOPHILIDAE**

A família Drosophilidae é apenas uma das mais de 150 famílias que compõem a ordem Diptera (Yeates & Wiegmann, 2005), mas encontra-se entre as mais diversas e amplamente distribuídas. Dentro dos ecossistemas, drosofilídeos desempenham um papel fundamental nas cadeias alimentares saprofíticas, uma vez que seus estágios imaturos alimentam-se de organismos fermentadores.

Throckmorton (1975), com base em dados fósseis e biogeográficos, postulou a origem da família Drosophilidae para as regiões tropicais do Velho Mundo, há cerca de 50 milhões de anos atrás (Eoceno). Segundo este autor, a evolução desta família teria se dado a partir de ancestrais que utilizavam fungos crescidos em folhíço como principal

recurso alimentar, de forma que a diversificação de substratos seria o resultado de oportunismos derivados, com uma certa ênfase na utilização de frutos.

Taxonomicamente, Drosophilidae pertence à superfamília Ephydroidea da seção Acalyptratae, juntamente com as famílias Camilidae, Curtonotidae, Diastatidae e Ephydridae (Thompson, 2005). Throckmorton (1975) sugere Diastatidae como grupo-irmão de Drosophilidae, uma vez que diastatídeos apresentam hábito saprófago de folhiço. Entretanto, McAlpine (1989) apresenta Camilidae como clado-irmão de Drosophilidae, enquanto Grimaldi (1990) favorece Curtonotidae para esta posição<sup>1</sup>, de forma que o correto posicionamento de Drosophilidae dentre os Ephydroidea permanece em aberto.

A monofilia da família Drosophilidae foi confirmada por Grimaldi (1990). Internamente, Drosophilidae apresenta mais de 3.900 espécies descritas, distribuídas em 76 gêneros (Bächli, 2008). Estes gêneros subdividem-se em duas subfamílias: Steganinae, que apresenta 28 gêneros; e Drosophilinae, que perfaz outros 47 gêneros (Tabela 1) (Toda, 2007; Bächli, 2008). Segundo Grimaldi (1988), evidências fósseis e biogeográficas sugerem que a divergência entre estas duas subfamílias ocorreu há pelo menos 70 milhões de anos atrás. A monofilia das duas subfamílias foi suportada por Grimaldi (1990) e, mais recentemente, por Remsen & O'Grady (2002), embora neste último caso apenas representantes dos gêneros *Amiota*, *Stegana*, *Rhinoleucophenga* e *Gitona* tenham sido amostrados no que diz respeito à subfamília Steganinae.

---

<sup>1</sup> Embora a filogenia consenso das suas três árvores de máxima parcimônia não auxilie na resolução desta questão.



Tabela 1. Classificação de 75 dos 76 gêneros pertencentes à família Drosophilidae em suas respectivas subfamílias [Obtido a partir de Toda (2007) e Bächli (2008)]. O número de espécies apresentado por cada táxon encontra-se indicado entre parênteses.

Gêneros da subfamília Steganinae	Gêneros da subfamília Drosophilinae	
1- <i>Acletoxenus</i> (4)	29- <i>Baeodrosophila</i> (5)	57- <i>Neotanygastrella</i> (18)
2- <i>Allopygaea</i> (3)	30- <i>Balara</i> (1)	58- <i>Palmomyia</i> (1)
3- <i>Amiota</i> (116)	31- <i>Bialba</i> (1)	59- <i>Palmophila</i> (2)
4- <i>Apenthecia</i> (12)	32- <i>Calodrosophila</i> (1)	60- <i>Paraliodrosophila</i> (4)
5- <i>Apsiphortica</i> (2)	33- <i>Celidosoma</i> (1)	61- <i>Paramycodrosophila</i> (16)
6- <i>Cacoxenus</i> (35)	34- <i>Chymomyza</i> (56)	62- <i>Phorticella</i> (11)
7- <i>Crincosia</i> (3)	35- <i>Cladochaeta</i> (124)	63- <i>Poliocephala</i> (1)
8- <i>Electrophortica</i> (1)	36- <i>Collesia</i> (5)	64- <i>Protochymomyza</i> (1)
9- <i>Eostegana</i> (13)	37- <i>Colocasiomyia</i> (26)	65- <i>Samoaia</i> (7)
10- <i>Erima</i> (1)	38- <i>Dettopsomyia</i> (13)	66- <i>Scaptodrosophila</i> (279)
11- <i>Gitona</i> (15)	39- <i>Diathoneura</i> (39)	67- <i>Scaptomyza</i> (263)
12- <i>Hyalistata</i> (6)	40- <i>Dichaetophora</i> (61)	68- <i>Sphaerogastrella</i> (10)
13- <i>Leucophenga</i> (205)	41- <i>Dicladochaeta</i> (1)	69- <i>Sphyrnoceps</i> (1)
14- <i>Luzonimyia</i> (4)	42- <i>Drosophila</i> (1.149)	70- <i>Styloptera</i> (10)
15- <i>Mayagueza</i> (1)	43- <i>Hirtodrosophila</i> (158)	71- <i>Tambourella</i> (3)
16- <i>Paraleucophenga</i> (8)	44- <i>Hypselothyrea</i> (30)	72- <i>Zaprionus</i> (56)
17- <i>Paraphortica</i> (1)	45- <i>Idiomyia</i> (379)	73- <i>Zapriothrica</i> (5)
18- <i>Pararhinoleucophenga</i> (2)	46- <i>Jeannelopsis</i> (3)	74- <i>Zaropunis</i> (1)
19- <i>Parastegana</i> (3)	47- <i>Laccodrosophila</i> (5)	75- <i>Zygothrica</i> (124)
20- <i>Phortica</i> (97)	48- <i>Liodrosophila</i> (64)	
21- <i>Pseudiasata</i> (6)	49- <i>Lissocephala</i> (32)	
22- <i>Pseudocacoxenus</i> (1)	50- <i>Lordiphosa</i> (58)	
23- <i>Pseudostegana</i> (33)	51- <i>Marquesia</i> (2)	
24- <i>Pyrgometopa</i> (1)	52- <i>Microdrosophila</i> (77)	
25- <i>Rhinoleucophenga</i> (18)	53- <i>Miomyia</i> (1)	
26- <i>Soederbomia</i> (1)	54- <i>Mulgravea</i> (14)	
27- <i>Stegana</i> (119)	55- <i>Mycodrosophila</i> (118)	
28- <i>Trachyleucophenga</i> (1)	56- <i>Neorhinoleucophenga</i> (2)	

Obs.: O gênero monoespecífico *Apacrochaeta* não foi incluído nesta tabela, uma vez que sua localização não foi determinada.

A classificação padrão de Drosophilidae foi publicada por Wheeler (1981 e 1986). Entretanto, os táxons sugeridos por este autor são inconsistentes com as relações filogenéticas entre as espécies, tanto quando estas são inferidas com base em dados morfológicos (Throckmorton, 1975; Grimaldi, 1990), como quando as inferências são baseadas em caracteres moleculares (DeSalle, 1992a e b; Pélandakis & Solignac, 1993; Thomas & Hunt, 1993; Russo *et al.*, 1995; Kwiatowski *et al.*, 1997; Remsen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001). Grimaldi (1990) realizou uma análise cladística de 217 caracteres morfológicos em 120 espécies de Drosophilidae, a partir do que propôs uma revisão na taxonomia da família. Entretanto, seus achados são também freqüentemente questionados (Thomas & Hunt, 1991 e 1993; Russo *et al.*, 1995; Kwiatowski *et al.*, 1997; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999 e 2001). Desta forma, preferiu-se seguir, nesta tese, a classificação apresentada por Bächli (2008), que embora mantenha inúmeros táxons parafiléticos, apresenta atualizações com relação ao status de alguns deles.

## **1. O GÊNERO *DROSOPHILA***

Dentre toda esta diversidade que caracteriza a família Drosophilidae, é possível dar destaque especial ao gênero *Drosophila*, que sozinho abrange cerca de um terço de suas espécies. Existem, entretanto, grandes discordâncias no que diz respeito ao posicionamento do gênero *Drosophila* dentro da família Drosophilidae. Throckmorton (1975) apresenta uma série de outros gêneros originando-se a partir do gênero *Drosophila*, de modo que este táxon ocuparia um posicionamento basal na família. Grimaldi (1990), entretanto, recuperou um gênero *Drosophila* derivado e monofilético, dadas as elevações taxonômicas sugeridas para os então subgêneros *Scaptodrosophila*, *Engiscaptomyza*, *Hirtodrosophila*, *Antopocerus* e *Lordiphosa*. Segundo Grimaldi

(1990), o clado constituído por diversos outros gêneros, tais como *Zaprionus*, *Samoaia* e *Liodrosophila*, constitui a linhagem irmã do gênero *Drosophila* assim definido. Hipóteses desenvolvidas a partir de dados moleculares, por outro lado, têm apresentado maiores semelhanças aos achados de Throckmorton, de forma que inúmeros trabalhos têm recuperado a parafilia do gênero *Drosophila* com relação a diferentes táxons: *Scaptomyza* [Thomas & Hunt, 1991 e 1993; DeSalle, 1992a; Russo *et al.*, 1995; Remsen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000; Remsen & O'Grady, 2002; Zilversmit *et al.*, 2002; DaLage *et al.*, 2007]; *Hirtodrosophila* [Kwiatowski *et al.*, 1997; Remsen & DeSalle, 1998; Tatarenkov *et al.*, 1999a e 2001; Kwiatowski & Ayala, 1999; DaLage *et al.*, 2007]; *Liodrosophila* [DeSalle, 1992b; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000; DaLage *et al.*, 2007]; *Zaprionus* [Pélandakis & Solignac, 1993; Thomas & Hunt, 1993; Russo *et al.*, 1995; Kwiatowski *et al.*, 1997; Remsen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000; Remsen & O'Grady, 2002; Zilversmit *et al.*, 2002; Yotoko *et al.*, 2003; DaLage *et al.*, 2007]; *Samoaia* [Pélandakis & Solignac, 1993; Tatarenkov *et al.*, 1999a e 2001]; *Mycodrosophila* [Katoh *et al.*, 2000; DaLage *et al.*, 2007]; *Lordiphosa* [Katoh *et al.*, 2000] e *Scaptodrosophila* e *Chymomyza* [Remsen & O'Grady, 2002]. Por este motivo, pode-se afirmar que o gênero *Drosophila* é essencialmente um agrupamento artificial.

A despeito desta composição parafilética, o gênero *Drosophila* encontra-se constituído por 1.149 espécies, subdivididas em oito subgêneros (Tabela 2) (Bächli, 2008). Toda esta diversidade, entretanto, ao mesmo tempo em que fornece diversas oportunidades de investigações acerca de padrões evolutivos, impõe uma série de desafios taxonômicos e filogenéticos.

Tabela 2. Apresentação dos oito subgêneros que compõem o gênero *Drosophila*, com seu respectivo número de espécies indicado entre parênteses. [Obtido a partir de Bächli (2008)]

---

Subgêneros do gênero <i>Drosophila</i>
1. <i>Chusqueophila</i> (1)
2. <i>Dorsilopha</i> (3)
3. <i>Drosophila</i> (721)
4. <i>Dudaica</i> (2)
5. <i>Phloridosa</i> (8)
6. <i>Psilodorha</i> (2)
7. <i>Siphlodora</i> (2)
8. <i>Sophophora</i> (332)

---

Obs.: De acordo com Bächli (2008) um total de 78 espécies não se encontram situadas em nenhum subgênero.

Como pode ser diagnosticado pela tabela 2, dentre os oito subgêneros que compõem o gênero *Drosophila*, destacam-se, pela diversidade apresentada, os subgêneros *Sophophora* e *Drosophila*. Segundo Russo *et al.* (1995), estes dois subgêneros divergiram há cerca de 40 milhões de anos. Tamura *et al.* (2004), por outro lado, ao corrigir distâncias sinônimas para os efeitos de viés na utilização de códons, obtém estimativas mais antigas, da ordem de 63 milhões de anos. Além disso, Tamura *et al.* (2004) demonstra que eventos de especiação em *Drosophila* não se distribuem regularmente no tempo, mas encontram-se inteiramente agrupados em determinados períodos do Cenozóico, especialmente aqueles em que ocorreram resfriamentos climáticos. Neste caso, a adaptação a novos ambientes áridos e a fragmentação das antigas florestas são os fatores postulados pelos autores como potenciais promotores da especiação no gênero.

## 1.1 - O subgênero *Sophophora*

O subgênero *Sophophora* foi estabelecido por Sturtevant (1939), e compreende 332 espécies subdivididas em 8 grupos (Bächli, 2008): *dentissima*, *dispar*, *fima*, *melanogaster*, *obscura*, *populi*, *saltans* e *willistoni*. Apesar de algumas relações filogenéticas dentro deste subgênero terem sido extensivamente estudadas, em especial aquelas referentes ao posicionamento de espécies e subgrupos dentro de grupos, algumas questões permanecem em aberto, entre as quais merece especial destaque aquelas referentes à monofilia e ao posicionamento do subgênero *Sophophora* dentro do gênero *Drosophila*. Esta questão será mais bem discutida a seguir, entretanto, a maior parte das evidências leva a crer que este subgênero é monofilético e basal dentro da filogenia do gênero *Drosophila* (Thomas & Hunt, 1993; Russo *et al.*, 1995; Kwiatowski *et al.*, 1997; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; DaLage *et al.*, 2007; Katoh *et al.*, 2007).

De acordo com Throckmorton (1962 e 1975), a radiação *Sophophora* subdivide-se em pelo menos três linhagens: (1) a linhagem derivada a partir de um ancestral “*proto-melanogaster*”, que dá origem aos grupos *melanogaster* e *obscura*<sup>2</sup>; (2) a linhagem Neotropical que compreende os grupos *saltans* e *willistoni*; (3) a linhagem que abrange os gêneros *Chymomyza* e *Neotanygastrella*. Além disso, embora não entre nos detalhes da questão, Throckmorton (1975) parece favorecer, em seu esquema de relações, a hipótese de que a radiação que dá origem ao subgênero *Drosophila* origina-se a partir de níveis inferiores da radiação *Sophophora*, de modo a definir um subgênero *Sophophora* parafilético também com relação ao subgênero *Drosophila*, resultado este refutado em uma série de estudos posteriores (Russo *et al.*, 1995; Remsen & DeSalle,

---

<sup>2</sup> Throckmorton (1962) considera igualmente provável a possibilidade de que os grupos *melanogaster* e *obscura* tenham se originado separadamente.

1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999 e 2001). Embora nos trabalhos de Pélandakis *et al.* (1991), Pélandakis & Solignac (1993) e Kwiatowski *et al.* (1997), bem como em algumas das topologias obtidas por Katoh *et al.* (2000) e por DaLage *et al.* (2007), o subgênero *Sophophora* tenha sido parafileticamente definido com relação ao subgênero *Drosophila*, dada a ramificação basal do grupo *willistoni*, este resultado não adquiriu maiores níveis de significância em nenhum destes trabalhos.

Ainda com relação ao cenário evolutivo proposto por Throckmorton (1975), trabalhos executados por Pélandakis *et al.* (1991), DeSalle (1992b), Kwiatowski *et al.* (1997), Katoh *et al.* (2000) e DaLage *et al.* (2007) demonstraram que o gênero *Chymomyza* não é um membro da radiação *Sophophora*. Entretanto, a parafilia do subgênero *Sophophora* foi diferentemente suportada por Katoh *et al.* (2000), que apresenta representantes do gênero *Lordiphosa* como constituintes do táxon irmão do cluster formado pelos grupos *saltans* e *willistoni*. Mais recentemente, resultados apresentados por O'Grady & Kidwell (1998), Kwiatowski & Ayala (1999), Remsen & O'Grady (2002) e DaLage *et al.* (2007), apoiaram a monofilia dos cladoss-irmãos *melanogaster/obscura* e *saltans/willistoni*. No trabalho de Pélandakis *et al.* (1991), por outro lado, o grupo *melanogaster* é parafiléticamente definido com relação aos grupos *obscura* e *fima*, enquanto que o grupo *willistoni* é parafilético com relação ao grupo *saltans*.

Throckmorton (1975) sugeriu, ainda, que a radiação *Sophophora* originou-se nos trópicos do Velho Mundo. Segundo Powell & DeSalle (1995) a linhagem que levou aos grupos *saltans* e *willistoni* divergiu da linhagem ancestral aos grupos *melanogaster* e *obscura* há de 30/40 milhões de anos. Mais exatamente, Russo *et al.* (1995), postula a datação desta divergência para cerca de 36 milhões de anos, enquanto que, segundo estes autores, os grupos *melanogaster* e *obscura* teriam divergido mais recentemente, há

aproximadamente 25 milhões de anos atrás. Tamura *et al.* (2004), por outro lado, estima que a divergência *saltans/willistoni* vs. *melanogaster/obscura* ocorreu há cerca de 62 milhões de anos, enquanto que a divergência *melanogaster* vs. *obscura* teria ocorrido há aproximadamente 55 milhões de anos.

Uma deficiência em todos estes trabalhos refere-se ao fato de que outros grupos tradicionalmente situados neste subgênero (*dentissima*, *dispar* e *populi*) ainda não foram incluídos em estudos envolvendo estas espécies, de modo a estabelecer uma análise de amplo nível.

#### 1.1.1 - Grupo *melanogaster*

O grupo *melanogaster* é formado por 185 espécies (Bächli, 2008), atualmente subdivididas em 12 subgrupos: *ananassae* (24 espécies), *montium* (91 espécies), *elegans* (5 espécies), *ficuspila* (6 espécies), *eugracilis* (1 espécie), *takahashii* (14 espécies), *suzuki* (18 espécies), *denticulata* (4 espécies), *flavohirta* (1 espécie), *rhopaloo* (5 espécies), *longissima* (2 espécies) e *melanogaster* (9 espécies). De acordo com Ashburner *et al.* (1984), Pélandakis & Solignac (1993), Goto & Kimura (2001), Kastanis *et al.* (2003), Yang *et al.* (2004) e Da Lage *et al.* (2007), estes grupos podem ser agrupados em três linhagens principais: (1) subgrupo *montium*; (2) subgrupo *ananassae*; (3) subgrupo *melanogaster* e demais subgrupos orientais analisados (*elegans*, *ficuspila*, *eugracilis*, *takahashii*, *suzuki* e *rhopaloo*). Entretanto, existem controvérsias quanto à identidade da primeira linhagem a divergir dentro do grupo. Segundo Pélandakis & Solignac (1993), Goto & Kimura (2001), Kastanis *et al.* (2003) e Da Lage *et al.* (2007), esta posição seria ocupada pelo grupo *ananassae*, enquanto Yang *et al.* (2004) defendem que o subgrupo *montium* teria sido o primeiro a divergir dentro do grupo. Uma terceira alternativa é proposta por Shawaroch (2002), em que os

subgrupos *ananassae* e *montium* são apresentados como grupos-irmãos entre si, de forma a constituir o clado-irmão do subgrupo *melanogaster* agrupado às espécies Orientais. Pélandakis *et al.* (1991) apresenta um resultado similar a este; entretanto, neste trabalho, o clado irmão do subgrupo *montium* é composto pelo subgrupo *ananassae*, bem como pelos grupos *fima* e *obscura*, de forma a constituir um grupo *melanogaster* parafilético. No estudo realizado por Young & Coleman (2004) com base em seqüências do espaçador intergênico ITS2 (*Internal Transcribed Spacer*) do rDNA nuclear, por outro lado, o grupo *melanogaster* foi parafileticamente definido com relação aos grupos *saltans* e *willistoni*, dado o agrupamento de *D. ananassae* com o clado constituído por *D. sturtevanti* e *D. willistoni*.

Sinais filogenéticos conflitantes também são encontrados nas relações entre o subgrupo *melanogaster* e os demais subgrupos Orientais. Wong *et al.* (2006) propõem que estas linhagens especiarão-se rapidamente a partir de um ancestral comum amplamente polimórfico, de forma que atualmente, fortes sinais de incongruência entre diferentes genes são encontrados. Com relação a estes subgrupos, resultados recentes apresentados por Lewis *et al.* (2005) e Da Lage *et al.* (2007) não suportam a monofilia do subgrupo *suzuki*, enquanto este último trabalho também apresenta *ficuspila* como polifilético. Ao mesmo tempo, Da Lage *et al.* (2007) propõem elevar o status taxonômico dos subgrupos *ananassae* e *montium* para o ranking de grupos de espécies, de forma que o grupo *melanogaster* agruparia apenas o subgrupo *melanogaster* e as espécies Orientais. Neste sentido, estes autores apresentam a linhagem composta pelo subgrupo *elegans* associado à *D. lucipennis* (representante do subgrupo *suzukii*) como ramificação basal do grupo *melanogaster*, sob sua nova composição. Powell & DeSalle (1995), com base em características sinapomórficas encontradas previamente para duplicações do gene *Gapdh* (gliceraldeído fosfato desidrogenase), propõem que o grupo



*melanogaster* pode ser subdividido em duas linhagens monofiléticas: a primeira composta pelos subgrupos *melanogaster* e *takahashii*, e a segunda formada pelos demais subgrupos.

Outro ponto controverso diz respeito às relações entre espécies que compõem o subgrupo *melanogaster*. De acordo com Bächli (2008), este subgrupo é formado pelas seguintes espécies: *D. melanogaster*, *D. simulans*, *D. mauritiana*, *D. sechellia*, *D. erecta*, *D. orena*, *D. yakuba*, *D. teissieri* e *D. santomea*. Segundo Caccone *et al.* (1996), Russo *et al.* (1995), Kastanis *et al.* (2003) e Da Lage *et al.* (2007), estas espécies evoluíram a partir de três linhagens principais: o clado *orena/erecta*, o clado *yakuba/teissieri/santomea* e o clado formado pelas espécies crípticas *melanogaster/simulans/sechellia/mauritiana*. Controvérsias existem quanto às relações entre estas linhagens, de forma que Lachaise *et al.* (1988) e Russo *et al.* (1995) apresentam o clado *orena/erecta* como basal dentro do subgrupo, enquanto Kastanis *et al.* (2003), Lewis *et al.* (2005), Wong *et al.* (2006) e Da Lage *et al.* (2007) defendem que os clados *orena/erecta* e *yakuba/teissieri/santomea* representam clados-irmãos. Segundo Young & Coleman (2004), por outro lado, o clado composto por *D. yakuba* constitui a ramificação basal do subgrupo *melanogaster*. O trabalho recente de Pollard *et al.* (2006), entretanto, parece ter contribuído definitivamente para a resolução desta questão. Eles demonstraram, em uma análise de escala genômica, que a filogenia agrupando as linhagens *erecta* e *yakuba* como irmãs fornece a hipótese melhor suportada, sendo que a ampla incongruência quanto a esta questão é atribuída pelos autores ao fenômeno de “incomplete lineage sorting”, associado ao curto espaço de tempo transcorrido entre os eventos de especiação que deram origem ao subgrupo *melanogaster*.

Outras questões que permanecem em aberto relacionam-se ao correto agrupamento da tríade *simulans/sechellia/mauritiana*. Caccone *et al.* (1996), com base em um conjunto de 13 loci nucleares, favorecem a hipótese de que *D. sechellia* e *D. mauritiana* são espécies-irmãs, visão esta defendida por Powell & DeSalle (1995) e Young & Coleman (2004). Por outro lado, Kastanis *et al.* (2003) apresentaram *D. mauritiana* e *D. simulans* como espécies-irmãs, enquanto que segundo Lewis *et al.* (2005), *D. sechellia* e *D. simulans* seriam mais proximamente relacionadas. Wong *et al.* (2006) argumentam que a especiação entre estas três espécies também parece ter ocorrido apenas recentemente, a partir de uma população grande e polimórfica, de forma que amplos níveis de polimorfismo ancestral são ainda compartilhados entre estas espécies.

Embora as espécies pertencentes ao grupo *melanogaster* distribuam-se principalmente pelos trópicos do Velho Mundo, de modo que uma origem Paleotropical é postulada (Pringent *et al.*, 2003), alguns de seus representantes, tais como *D. ananassae* (grupo *ananassae*), *D. kikkawai* (grupo *montium*), *D. simulans* e *D. melanogaster* (grupo *melanogaster*), apresentam atualmente uma distribuição cosmopolita. Segundo Throckmorton (1975) este grupo quase certamente originou-se no sul da Ásia, enquanto o subgrupo *melanogaster* parece ter-se originado na África sub-Sahariana e ilhas da costa leste da África (Caccone *et al.*, 1996). Algumas estimativas de divergência dentro do grupo *melanogaster* são fornecidas por Tamura *et al.* (2004): subgrupos *melanogaster* vs. *ananassae*: 44 milhões de anos; subgrupos *melanogaster* vs. *montium*: 41 milhões de anos; subgrupos *melanogaster* vs. *takahashii*: 36 milhões de anos; clados *melanogaster/simulans* vs. *yakuba/teissieri* ou *orena/erecta*: 13 milhões de anos; *D. melanogaster* vs. *D. simulans*: 5 milhões de anos; *D. simulans* vs. *D. mauritiana*: 1 milhão de anos.

### 1.1.2 – Grupo *obscura*

O grupo *obscura* é constituído por 42 espécies descritas (Bächli, 2008), grande parte das quais habita florestas temperadas Holárticas, com algumas adaptações secundárias a ambientes de altas altitudes, com clima similar ao temperado, em regiões Afrotropicais (subgrupo *microlabis*), Neotropicais (*D. azteca* e *D. tolteca*, membros do subgrupo *affinis*) e Orientais (Gao *et al.*, 2007). De modo geral, entretanto, uma origem Paleártica é postulada (Pringent *et al.*, 2003). Atualmente, as espécies do grupo *obscura* são subdivididas em 6 subgrupos: *affinis* (10 espécies), *microlabis* (2 espécies), *obscura* (16 espécies), *pseudoobscura* (6 espécies), *sinobscura* (3 espécies) e *subobscura* (3 espécies), além de duas espécies, *D. alpina* e *D. helvetica*, não classificadas em nenhum subgrupo (Bächli, 2008). Originalmente, entretanto, Sturtevant (1942) dividiu o grupo *obscura* em apenas duas linhagens principais: o subgrupo *affinis*, exclusivamente Neártico, e o subgrupo *obscura*, com representantes Neárticos e Paleárticos. Trabalhos posteriores (Lakovaara & Saura, 1982; Latorre *et al.*, 1988) demonstraram que os membros Neárticos do subgrupo *obscura* eram mais proximamente relacionados ao subgrupo *affinis* do que aos demais membros do subgrupo *obscura*, de modo que foi proposta sua elevação taxonômica para o novo subgrupo *pseudoobscura*, enquanto as espécies Paleárticas continuaram agrupadas no subgrupo *obscura*. Barrío *et al.* (1994) propuseram que os representantes do “heterogêneo subgrupo *obscura*” fossem redistribuídos entre dois novos subgrupos: o subgrupo *obscura*, perfazendo a maior parte das espécies, e o subgrupo *subobscura*, constituído por *D. subobscura* e duas espécies endêmicas de ilhas, *D. guanche* e *D. madeirensis*. Além disso, espécies Afrotropicais recentemente descritas para a Etiópia levaram à proposição do subgrupo *microlabis* (Tsacas *et al.*, 1985).

Segundo Barrío & Ayala (1997), o grupo *obscura* constitui um táxon monofilético derivado de diversas linhagens: (1) o subgrupo *subobscura*, que, segundo estes autores talvez represente o clado ancestral de todo grupo, hipótese esta recentemente suportada por DaLage *et al.* (2007); (2) uma ou mais linhagens contendo representantes do subgrupo *obscura*, que apresenta uma composição poli ou parafilética na maior parte dos cladogramas apresentados naquele trabalho; (3) um cluster Neártico constituído por dois cladogramas-irmãos, os subgrupos *affinis* e *pseudoobscura*. Uma quarta linhagem poderia ser ainda adicionada para inclusão do subgrupo *microlabis*, provável grupo-irmão da linhagem Neártica, segundo DaLage *et al.* (2007). Entretanto, segundo Gao *et al.* (2007), *microlabis* constitui o subgrupo-irmão de *subobscura*, de forma que o clado assim estabelecido corresponde a linhagem basal do grupo *obscura*.

Por outro lado, Ramos-Onsins *et al.* (1997) e O'Grady (1999) defendem que dois grandes cladogramas compõem o grupo *obscura*: o clado do Velho Mundo, formado pelos subgrupos-irmãos *obscura* e *subobscura*, e o clado do Novo Mundo, constituído pelos subgrupos-irmãos *affinis* e *pseudoobscura*. O trabalho recente de Gao *et al.* (2007), entretanto, ao apresentar uma amostragem mais extensa de representantes do Velho Mundo, recupera apenas a monofilia do clado do Novo Mundo, a despeito da parafilia dos representantes Paleárticos. Neste estudo, os subgrupos *subobscura*, *sinobscura*, *affinis* e *pseudoobscura* foram monofileticamente recuperados, enquanto o subgrupo *obscura*, mais uma vez, apresentou uma constituição polifilética, subdividindo-se em três cladogramas diferentes: o cluster *D. ambigua/D. obscura/D. tristis/D. subsilvestris/D. dianensis*, irmão do subgrupo *sinobscura*; o cluster formado pelo par *D. bifasciata/D. imaii*; e o cluster formado pelo par *D. limingi/D. tsukubaensis*. Baseados neste resultado, os autores propõem uma nova revisão taxonômica com referência ao subgrupo *obscura*, de modo que o mesmo passe a referir-se apenas às cinco espécies

acima explicitadas, bem como OTU's relacionadas futuramente definidas. Gleason *et al.* (1997) e O'Grady (1999), por outro lado, suportam a monofilia do subgrupo *obscura*.

Throckmorton (1962 e 1975) defende que a divergência *melanogaster-obscura* teria ocorrido antes do Oligoceno Médio, de modo o grupo *obscura* teria se originado a partir de uma linhagem *protomelanogaster*, nos trópicos do Velho Mundo. Segundo ele, a adaptação a ambientes temperados teria surgido posteriormente, acompanhada de “homo-seleção” e efeito-fundador, e com a expansão das florestas temperadas Paleárticas durante o Oligoceno Médio. Ainda segundo este autor, a expansão de parte do grupo *obscura* para o Novo Mundo teria ocorrido a partir do Estreito de Bering, no período Mioceno, há cerca de 20 milhões de anos. O trabalho recente de Gao *et al.* (2007) demonstrou que pelo menos metade das grandes linhagens constituintes do grupo *obscura* parecem ter se originado a partir de radiações ocorridas nas imediações do Mioceno Médio, há cerca de 15 milhões de anos, época em que houve o desenvolvimento final e a fragmentação das florestas temperadas no Hemisfério Norte.

### 1.1.3 – Grupo *saltans*

Os grupos *saltans* e *willistoni* são os representantes da radiação *Sophophora* ocorrida no Novo Mundo (Throckmorton, 1975). O grupo *saltans* é formado por um total de 21 espécies descritas, subdivididas em cinco subgrupos (Bächli, 2008): subgrupo *cordata* (2 espécies), subgrupo *elliptica* (4 espécies), subgrupo *parasaltans* (2 espécies), subgrupo *saltans* (7 espécies) e subgrupo *sturtevantii* (6 espécies).

Segundo Throckmorton & Magalhães (1962), a ordem ascendente de divergência dos subgrupos, a partir do mais primitivo, seria: *cordata*, *elliptica*, *sturtevantii*, *parasaltans* e *saltans*. Baseado nesta filogenia e nos padrões contemporâneos de distribuição destes subgrupos, Throckmorton (1975) postulou que o

ancestral do grupo originou-se nos trópicos da América do Norte, local de distribuição atual dos subgrupos “mais primitivos”, *cordata* e *elliptica*. A partir daí teria havido a colonização da América do Sul, de forma que Throckmorton defende que os subgrupos “derivados” *sturtevantii*, *saltans* e *parasaltans* teriam se diversificado no continente sul-americano antes da formação do Istmo do Panamá. Concordando com estes resultados, o trabalho de O’Grady *et al.* (1998) recuperou em sua filogenia de evidências totais a ramificação basal dos subgrupos *cordata* e *elliptica*, sendo os subgrupos-irmãos *saltans* e *parasaltans* apresentaram um posicionamento derivado também em relação ao subgrupo *sturtevantii*, com ramificação intermediária.

Entretanto, a filogenia apresentada posteriormente por Rodríguez-Trelles *et al.* (1999a) difere significativamente daquelas obtidas por Throckmorton & Magalhães (1962) e por O’Grady *et al.* (1998). De acordo com a hipótese sugerida por Rodríguez-Trelles *et al.* (1999a) e Tarrío *et al.* (2000) a partir da análise de 2.085 pb do gene *Xdh* (xantina-desidrogenase), o subgrupo *parasaltans* é a linhagem mais antiga do grupo *saltans*, seguida, em progressão sucessiva, pelos subgrupos *sturtevantii*, *elliptica*, *cordata* e *saltans*. O ponto mais inconsistente nas topologias apresentadas por estes autores diz respeito à relação entre os subgrupos *elliptica* e *cordata*, ora apresentados como táxons-irmãos, ora como ramos sucessivos. Análises executadas por Rodríguez-Trelles *et al.* (1999a) a partir dos dados gerados por O’Grady *et al.* (1998), demonstram que inclusive estes datasets, quando analisados sob um modelo evolutivo mais coerente, favorecem sua nova hipótese evolutiva. Neste caso, como a distribuição geográfica do subgrupo *parasaltans*, agora mais antigo na filogenia do grupo *saltans*, é postulada para o Brasil e para as Ilhas do Caribe, o novo cenário biogeográfico emergente favorece a hipótese de que *saltans* teria se originado nos trópicos da América do Sul, e não na América do Norte.

Trabalhos recentes têm, ainda, revelado, que o grupo *saltans*, juntamente com o grupo *willistoni*, exibe padrões de substituição diferentes daqueles observados em outras linhagens do subgênero *Sophophora*, apresentando um menor conteúdo GC, principalmente na terceira posição dos seus códons, e em menor proporção na primeira posição dos seus códons e em seus íntrons (Rodríguez-Trelles *et al.*, 1999b e 2000a; Tarrío *et al.*, 2000). Segundo Rodríguez-Trelles *et al.* (1999b, 2000a e 2000b), as diferenças composicionais são mais acentuadas em regiões e genes menos restritos funcionalmente. Além disso, as espécies dos grupos *saltans* e *willistoni* também apresentam padrões de viés na utilização de códons menores e diferentes daqueles previamente conhecidos para o subgênero *Sophophora* (Rodríguez-Trelles *et al.*, 1999b e 2000a; Powell *et al.*, 2003), com um aumento significativo no uso de códons terminados em U, em detrimento dos terminados em C, para cerca de um terço dos aminoácidos (Powell *et al.*, 2003). Paralelamente, na linhagem *saltans/willistoni*, o gene *Xdh*, assim como diversos outros genes menos restritos (*period*, *Amyrel*, *dopa-decarboxilase*) apresentam um déficit de aminoácidos com códons ricos em GC, contrastado por um excesso de aminoácidos codificados por códons ricos em AT (Rodríguez-Trellez *et al.*, 1999b, 2000a e 2000b). Rodríguez-Trellez *et al.* (1999b e 2000b) apresentaram, também, evidências de que os grupos *saltans* e *willistoni* estão evoluindo a taxas maiores que os grupos *obscura* e *melanogaster*, ao menos no que se refere a substituições não-sinônimas no gene *Xdh*. O trabalho de Rodríguez-Trelles *et al.* (2000b) demonstra, ainda, a partir da utilização de modelos realísticos não-homogêneos e não-estacionários aplicados sobre uma filogenia conhecida, que o desvio nos padrões mutacionais ocorreu na linhagem ancestral aos grupos *saltans* e *willistoni*, sendo que, aparentemente, o grupo *saltans* continua a aumentar seu conteúdo AT.

Todos estes resultados são congruentes com um deslocamento de pressão mutacional ocorrido nestas linhagens (Rodríguez-Trellez *et al.*, 1999b, 2000a e 2000b), uma vez que não foram encontradas evidências de seleção positiva ou de relaxamento de restrições seletivas, fenômenos que poderiam ser diagnosticados por um aumento significativo na razão entre a taxa de substituição não-sinônima ( $dN$ ) e a taxa de substituição sinônima ( $dS$ ) (Rodríguez-Trelles *et al.*, 2000a). Além disso, Rodríguez-Trelles *et al.* (1999b) encontraram uma correlação positiva significativa entre o conteúdo GC apresentado por íntrons e éxons do gene *Xdh*, sendo que as modificações são mais conspícuas em regiões e genes menos restritas funcionalmente, o que é compatível com o cenário de deslocamento de viés mutacional. Neste caso, o aumento nas taxas evolutivas do gene *Xdh* também é condizente com o fenômeno de explosão mutacional, previsto em casos de mudanças no viés mutacional como uma tentativa de atingir um novo equilíbrio composicional, e cujo efeito é mais pronunciado para regiões e genes menos restritos. Por outro lado, Powell *et al.* (2003), ao demonstrar que apenas cerca de um terço de todos os aminoácidos apresentam desvios com relação ao uso preferencial de códons apresentado por *D. melanogaster*, favorecem a hipótese de que mudanças na abundância relativa de alguns tRNA's isoceptores é a melhor explicação disponível para o atual conjunto de observações. Outra evidência, apresentada por estes autores, em favor desta hipótese refere-se ao fato de que genes portadores de um maior viés na utilização de códons apresentam o maior desvio com relação ao uso preferencial de códons apresentado por *D. melanogaster*. Este cenário, entretanto, não explica as diferenças de composição nucleotídica encontradas em regiões não-codificadoras. Mais recentemente, Singh *et al.* (2006) forneceram novos *insights* a essa questão, ao demonstrar que o deslocamento dos padrões background de substituição na linhagem *saltans/willistoni* deve-se primariamente a mudanças nos padrões mutacionais.



Entretanto, a diferença estável de conteúdo GC predita por estes resultados explica apenas as mudanças de composição intrônicas, sendo insuficientes para explicar a maior magnitude de mudanças encontradas nas regiões codificadoras. Neste caso, os autores sugerem que as modificações no uso preferencial de códons ocorridas na linhagem *saltans/willistoni* refletem mudanças nas definições do pool preferido de códons, ou uma menor pressão seletiva sobre viés na utilização de códons.

#### 1.1.4 – Grupo *willistoni*

O grupo *willistoni* é constituído por 23 espécies, subdivididas atualmente em três subgrupos (Bächli, 2008): *alagitans* (5 espécies), *bocainensis* (12 espécies), *willistoni* (6 espécies). O subgrupo *willistoni* agrupa *D. willistoni* e suas cinco espécies crípticas: *D. equinoxialis*, *D. paulistorum*, *D. pavlovsiana*, *D. tropicalis* e *D. insularis*. Estas espécies exibem graus variados de isolamento pré-zigótico e geralmente falham em hibridizar (Ehrman & Powell, 1982). A mais notável exceção refere-se ao cruzamento entre *D. paulistorum* e *D. pavlovskiana*, capaz de produzir prole constituída por machos estéreis e fêmeas férteis (Spassky *et al.*, 1971), embora Winge (1965) e Winge & Cordeiro (1995) também tenham relatado a obtenção de híbridos nos cruzamentos envolvendo outras espécies crípticas. As espécies do subgrupo *willistoni* são, em geral, muito similares em sua morfologia externa, sendo que os métodos utilizados em sua identificação restringem-se à inspeção da genitália dos machos (Spassky, 1957), ao estudo detalhado dos cromossomos politênicos (Rohde *et al.*, 2006) e a testes de cruzamento (Spassky *et al.*, 1971), embora recentemente, Garcia *et al.* (2006) tenham descrito a utilização da variação da mobilidade eletroforética da enzima fosfatase ácida-1 (*Acph1*) na diagnose rápida e eficiente de cinco das seis espécies crípticas pertencentes ao subgrupo.

O subgrupo *willistoni* contém, ainda, as seis semi-espécies que compõem a superespécie *D. paulistorum*, que são: Andino-Brasileira, Amazônica, Centro-Americana, Interior, Transicional e Orinocana. Estas semi-espécies exibem entre si um isolamento sexual etológico pronunciado, embora não absoluto, sendo que quando o cruzamento entre elas é efetivado, o que acontece essencialmente em condições laboratoriais, híbridos inviáveis ou fêmeas férteis e machos estéreis são produzidos (Ehrman & Powell, 1982). A semi-espécie Transicional, entretanto, parece ocupar um status intermediário, sendo capaz de produzir híbridos férteis e viáveis, ao menos quando inter-cruzada com representantes das semi-espécies Andino-Brasileira e Centro-Americana (Dobzhansky *et al.*, 1964; Ehrman & Powell, 1982).

Além disso, o subgrupo críptico também apresenta diferenciação em nível de subespécies, em que graus menores de isolamento etológico são encontrados, mas apenas machos híbridos estéreis são produzidos: *D. willistoni* diferencia-se nas subespécies *willistoni* e *quechua* (restrita a áreas da Colômbia, Peru, Equador e Ilhas Galápagos); *D. equinoxialis* apresenta as subespécies *equinoxialis* e *caribbensis* (encontrada em áreas da Costa Rica, do Porto Rico, da Jamaica e de Hispaniola); e *D. tropicalis* subdivide-se nas subespécies *tropicalis* e *cubana* (distribuída em determinadas regiões do Brasil, Cuba, Porto Rico, Jamaica e Hispaniola)<sup>3</sup>. Também alocada ao grupo *willistoni* está uma linhagem de status taxonômico indefinido, chamada de linhagem Carmody (Carmody, 1965), capaz de intercruzar com algumas linhagens de *D. equinoxialis*, bem como com uma das semi-espécies de *D. paulistorum*. Assim sendo, o grupo *willistoni* é certamente um complexo de vários níveis taxonômicos, contendo inclusive diversas espécies incipientes em *statu nascendi*

---

<sup>3</sup> Winge (1965) argumenta que estas duas subespécies deveriam ser tratadas como espécies separadas, embora Rohde *et al.* (2006) não tenham detectado diferenças nos arranjos cromossômicos por elas apresentados.

(Dobzhansky & Spassky, 1959). Ele fornece, portanto, um material muito conveniente para diversos estudos biológicos, com um arranjo seqüencial de estágios evolutivos, ideal para análises envolvendo mecanismos evolutivos.

Quanto à sua distribuição, o grupo *willistoni* é, essencialmente, Neotropical, sendo que o subgrupo *willistoni* chega a constituir um dos mais notáveis e abundantes grupos de drosofilídeos nesta região (Garcia *et al.*, 2006). Throckmorton (1975) argumenta que a origem deste grupo se deu na América do Sul, após uma migração ancestral a partir da América do Norte. Através de estudos biogeográficos, Spassky *et al.* (1971) e Ehrman & Powell (1982) apontam *D. willistoni* como a espécie com mais ampla distribuição dentro do grupo, existindo em altas densidades desde a Flórida até o sul do Brasil e Argentina, e desde o Oceano Pacífico até o Atlântico. A distribuição das demais espécies do subgrupo críptico encontra-se, em geral, dentro da extensão ocupada por *D. willistoni*, sendo que esta espécie, juntamente com *D. paulistorum*, *D. equinoxialis* e *D. tropicalis* são simpátricas desde a Guatemala, através da América Central e do Norte da América do Sul, até a região central do Peru e norte do Brasil (Ayala *et al.*, 1970; Spassky *et al.*, 1971). Segundo Spassky *et al.* (1971), ao sul do estado de São Paulo, no Brasil, apenas *D. paulistorum* e *D. willistoni* são encontradas. As outras duas espécies do subgrupo *willistoni* apresentam uma ocorrência restrita e endêmica (Ehrman & Powell, 1982): *D. insularis*, habitante de algumas ilhas das Pequenas Antilhas; e *D. pavlovskiana*, espécie rara que ocorre nas Guianas. Quanto à super-espécie *D. paulistorum*, a região da América Central ao norte do Panamá tem ocupação restrita pela semi-espécie Centro-Americana; e o nordeste, centro e sul do Brasil, juntamente com o Peru e a Bolívia, apresentam apenas a semi-espécie Andino-Brasileira de *D. paulistorum* (Spassky *et al.*, 1971). Na região mais ao sul do Panamá e no norte da América do Sul registra-se a ocorrência, isolada ou em simpatria, das outras

quatro semi-espécies, juntamente com a Andino-Brasileira, que ocupa a maior área de distribuição: a semi-espécie Transicional ocorre apenas no norte da Colômbia e Venezuela; a semi-espécie Amazônica estende-se desde o sul do Panamá e Trinidad, até os estuários dos rios Amazonas e Tocantins; Orinocana parece confinada à região sul do Panamá até o centro da Colômbia e da Venezuela; e Interior habita regiões mais ao sul da Colômbia e da Venezuela, até as bacias dos rios Negro, Solimões e Amazonas (Spassky *et al.*, 1971). Ainda segundo Spassky *et al.* (1971) e Ehrman & Powell (1982), esta distribuição das semi-espécies de *D. paulistorum* parece estar relacionada à ocorrência de refúgios na região da Floresta Amazônica durante o Pleistoceno.

Até o momento, poucos estudos foram desenvolvidos com o objetivo de auxiliar na resolução das relações filogenéticas dentro do grupo *willistoni*. O trabalho de Spassky *et al.* (1971), desenvolvido com base em evidências genéticas, citogenéticas, bioquímicas e biogeográficas então disponíveis, apresenta um diagrama esquemático das relações dentro do subgrupo *willistoni*. Segundo este diagrama, *D. equinoxialis*, *D. pavlovskiana* e *D. paulistorum* são espécies-irmãs, da mesma forma que *D. insularis* e *D. tropicalis*, enquanto *D. willistoni* não aparece agrupada a nenhuma outra espécie. Ainda segundo Spassky *et al.* (1971), dentre as semi-espécies de *D. paulistorum*, três linhagens principais podem ser distinguidas: (1) a que agrupa as semi-espécies Transicional, Centro-Americana e Andino-Brasileira; (2) a que une Interior e Orinocana; (3) a mono-específica, constituída pela semi-espécie Amazônica. O trabalho de Ayala *et al.* (1974), por outro lado, já apresenta um dendrograma de relações para o subgrupo, construído a partir de dados de diferenciação genética obtidos mediante análise de 36 loci de alozimas. Neste trabalho o clado formado pelo agrupamento de *D. pavlovskiana/D. equinoxialis/D. paulistorum* junto a *D. willistoni/D. tropicalis* apresenta *D. insularis* como ramificação basal. Ainda segundo Ayala *et al.* (1974), as

semi-espécies Orinocana e Interior de *D. paulistorum* compõem a linhagem irmã do clado que agrupa a semi-espécie Amazônica ao cluster formado pelas semi-espécies Centro-Americana/Transicional/Andino-Brazileira. Mais recentemente, Gleason *et al.* (1998) publicaram o primeiro trabalho de filogenia do subgrupo *willistoni* realizado a partir da análise de seqüências codificadoras de DNA – a partir dos genes *Period (per)*, *Álcool-desidrogenase (Adh)* e *Citocromo Oxidase I (COI)*. A hipótese melhor suportada por estes dados apresenta *D. insularis* como ramificação basal do subgrupo *willistoni*, seguida por *D. tropicalis*, de forma que *D. willistoni* aparece como espécie-irmã do clado que agrupa *D. equinoxialis* e *D. paulistorum*. Resultados semelhantes a estes foram recuperados por Tarrío *et al.* (2000), a partir da análise de seqüências do gene *xantina-desidrogenase (Xdh)*. Entretanto, a configuração das espécies dentro do subgrupo *willistoni* apresentada por O’Grady & Kidwell (2002) é incongruente com aquela apresentada por Gleason *et al.* (1998) e por Tarrío *et al.* (2000), uma vez que naquele trabalho *D. equinoxialis* aparece como ramificação basal, e *D. insularis* agrupa-se com o clado formado por *D. willistoni* e *D. tropicalis*, de forma a constituir o grupo-irmão do clado que une *D. paulistorum* e *D. pavlovskiana*. A filogenia recentemente publicada por Rohde *et al.* (2006), com base na análise evolutiva dos padrões cromossômicos encontrados no cromossomo IIR, vem, por fim, confirmar a proximidade evolutiva entre *D. equinoxialis* e *D. paulistorum*, ao mesmo tempo em que suporta a idéia de que *D. insularis* seja a espécie mais divergente dentro do subgrupo *willistoni*. Este trabalho também apóia a efetividade do isolamento existente entre as cinco espécies crípticas (*D. pavlovskiana* não foi incluída), ao demonstrar que todas as inversões segregantes apresentam uma distribuição restrita a apenas uma das espécies. Por outro lado, o trabalho de Gleason *et al.* (1998) apresenta a linhagem *Carmody* agrupada dentro do clado composto por diferentes populações de *D. equinoxialis*, de

modo que, filogeneticamente, esta linhagem deveria ser classificada como *D. equinoxialis*. Outro ponto de incongruência com a taxonomia tradicional diz respeito ao posicionamento de *D. pavlovskiana*, que segundo o conceito biológico de espécie seria uma espécie diferente, mas dentro de uma perspectiva filogenética, agrupa-se dentro do cluster formado pelas diferentes semi-espécies de *D. paulistorum* (Gleason *et al.*, 1997 e 1998).

Quanto às relações entre as semi-espécies de *D. paulistorum*, Gleason *et al.* (1998), com base na análise do gene *per*, favorece a hipótese de que Amazônica foi a primeira semi-espécie a divergir, no que é seguida pelo clado que agrupa Andino-Brasileira e Orinocana. Neste caso os clusters que incluem as semi-espécies Transicional/Interior e Centro-Americana/*D. pavlovskiana* aparecem como derivados e irmãos entre si. O trabalho de Rohde *et al.* (2006), embora inclua apenas quatro das seis semi-espécies, é congruente com estes resultados ao apresentar Andino-Brasileira e Orinocana como homossequenciais entre si, ao mesmo tempo em que relatam o compartilhamento de uma inversão encontrada em estado polimórfico na semi-espécie Centro-Americana e fixa para a semi-espécie Interior.

Poucos trabalhos têm sido executados, ainda, no sentido de auxiliar no estabelecimento das relações apresentadas pelas demais espécies do grupo *willistoni*. Entre as exceções, pode-se citar os estudos de Tarrío *et al.* (2000), O'Grady & Kidwell (2002) e Powell *et al.* (2003), que, de modo geral, estão em concordância ao sugerir existência de três grandes linhagens dentro do grupo *willistoni*: a primeira unindo *D. capricorni* e *D. sucinea*; a segunda, mono-específica, formada por *D. nebulosa*; e a terceira constituída pelas espécies crípticas. Entretanto, segundo Tarrío *et al.* (2000) o subgrupo *bocainensis* apresenta uma composição parafilética com relação ao subgrupo *willistoni*, sendo *D. nebulosa* mais proximamente relacionada às espécies crípticas do

que aos demais membros de seu subgrupo. Por outro lado, no trabalho de O'Grady & Kidwell (2002), o subgrupo *bocainensis* é polifileticamente definido, uma vez que *D. nebulosa* agrupa isoladamente com o grupo *saltans*, *D. fumipennis* aparece como espécie-irmã do clado *saltans/willistoni*, e o clado *capricorni/sucinea* constitui o grupo-irmão do subgrupo *willistoni*. Já na análise de Powell *et al.* (2003), o posicionamento de *D. nebulosa* apresentou-se inconstante com relação aos diferentes métodos de reconstrução filogenética, sendo que o maior corpo de evidências leva à proposição de que esta é a espécie-irmã do subgrupo *willistoni*.

Outro ponto relevante e controverso diz respeito à monofilia do grupo *willistoni*. No estudo de Pélandakis *et al.* (1991) e em algumas filogenias apresentadas por O'Grady & Kidwell (2002), este grupo é parafileticamente definido com relação ao grupo *saltans*. O trabalho de Tarrío *et al.* (2000), entretanto, auxilia na resolução deste debate ao apoiar confidentemente a monofilia do grupo *willistoni* em sua filogenia. Estes autores afirmam, ainda, ter encontrado uma deleção de um íntron de *Adh* sinapomórfico para o grupo. Além disso, os padrões de substituição peculiares demonstrados por diversos autores para as espécies do grupo *willistoni* (Powell *et al.*, 2003; Rodríguez-Trelles *et al.*, 2000a e 2000b; Tarrío *et al.*, 2000), com um elevado conteúdo A+T, juntamente com o fenômeno de atração de ramos longos, parecem ainda explicar parte das dificuldades previamente encontradas em recuperar sua monofilia.

## 1.2 – O subgênero *Drosophila*

O subgênero *Drosophila* é o maior representante do gênero *Drosophila*, com pelo menos 721 espécies, subdivididas em 43 grupos (Bächli, 2008): *angor*, *annulimana*, *antioquia*, *aureata*, *bizonata*, *bromeliae*, *calloptera*, *canalinae*, *carbonaria*, *cardini*, *carsoni*, *coffeata*, *dreyfusi*, *flavopilosa*, *funnebris*, *guarani*, *guttifera*,

*histro*, *immigrans*, *macroptera*, *melanica*, *mesophragmatica*, *morelia*, *nannoptera*, *nigrosarsa*, *onychophora*, *pallidipennis*, *peruviana*, *picta*, *pinicola*, *polychaeta*, *quadrissetata*, *quinaria*, *repleta*, *robusta*, *rubrifrons*, *simulivora*, *sticta*, *testacea*, *tripunctata*, *tumiditarsus*, *virilis* e *xanthopallescens*. Segundo Powell & DeSalle (1995), esta radiação ocorreu entre 25 e 36 milhões de anos atrás.

O subgênero *Drosophila* constitui o maior representante da família Drosophilidae em regiões Neotropicais, onde apresenta, pelo menos, 25 grupos de espécies (Val *et al.*, 1981). Malogolowkin (1953) e Throckmorton (1975) dividiram as espécies do subgênero *Drosophila* em duas linhagens principais: a seção ou radiação *virilis-repleta* e a seção ou radiação *quinaria-tripunctata*, além do grupo *funeris*, basal em relação a ambos segundo Throckmorton (1975). Segundo Russo *et al.* (1995) membros das duas seções ou radiações divergiram entre si há aproximadamente 33 milhões de anos. Suporte para a existência das duas radiações é, entretanto, raramente encontrado, sendo que nos poucos casos em que isto acontece, os clados apresentados têm sua composição alterada com relação ao que foi proposto por Throckmorton em 1975 (Remsen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Remsen & O'Grady, 2002; Yotoko *et al.*, 2003; DaLage *et al.*, 2007; Katoh *et al.*, 2007). Em DeSalle (1992b), membros de ambas as seções encontram-se irregularmente distribuídos ao longo da filogenia do subgênero *Drosophila*.

Diversos autores têm defendido que o subgênero *Drosophila* é, na verdade, parafilético com relação a diversos outros subgêneros de *Drosophila* e até mesmo com relação a outros gêneros de Drosophilidae. Katoh *et al.* (2000), por exemplo, apresenta o subgênero *Drosophila* subdividido em diversas linhagens intimamente associadas a membros dos gêneros *Liodrosophila*, *Mycodrosophila*, *Hirtodrosophila*, *Lordiphosa* e *Scaptomyza*. A parafilia do subgênero *Drosophila* com relação ao gênero *Scaptomyza*



foi ainda suportada por Russo *et al.* (1995), Remsen & DeSalle (1998), Kwiatowski & Ayala (1999), Tatarenkov *et al.* (1999a e 2001), Remsen & O'Grady (2002) e DaLage *et al.* (2007). Kwiatowski & Ayala (1999) defendem ainda a parafilia do subgênero *Drosophila* com relação a *Hirtodrosophila*. Segundo Russo *et al.* (1995), o gênero *Zaprionus* constitui a linhagem irmã do subgênero *Drosophila* inclusivamente constituído, local este ocupado pelo gênero *Liodrosophila* no trabalho de DeSalle (1992b), que apresenta o gênero *Zaprionus* como grupo irmão do gênero *Drosophila*. Em Remsen & DeSalle (1998), por outro lado, *Hirtodrosophila* constitui o grupo irmão do subgênero *Drosophila*, sendo que o clado assim constituído apresenta *Zaprionus* como linhagem irmã. Por outro lado, DaLage *et al.* (2007) apontam diversos outros gêneros apresentados a seguir em progressão sucessiva, como basais em relação ao subgênero *Drosophila*: *Zaprionus*, *Hirtodrosophila*, *Mycodrosophila* e *Liodrosophila*.

### 1.2.1 – Seção *virilis-repleta*

A seção *virilis-repleta* é formada pelos grupos *angor*, *annulimana*, *bromeliae*, *canalinae*, *carbonaria*, *carsoni*, *coffeata*, *dreyfusi*, *flavopilosa*, *melanica*, *mesophragmatica*, *polychaeta*, *quadrisetata*, *repleta*, *robusta*, *virilis*, *xanthopallescens*, bem como os grupos *nannoptera* e *peruviana*, que habitam regiões desérticas. Segundo Wang *et al.* (2006), espécies pertencentes a esta seção distribuem-se por diferentes zonas climáticas, sendo encontradas desde a zona circum-tropical e subtropical (grupo *polychaeta*, por exemplo), em zonas temperadas quentes (grupos *melanica* e *robusta*, por exemplo), até zonas temperadas frias e regiões circum-polares (grupo *virilis*, por exemplo). Além disso, representantes do grupo *repleta* são encontrados em regiões desérticas, áridas e semi-áridas, do Novo Mundo, onde se encontram associados a diversas espécies de cactus. Throckmorton (1975) sugeriu o sul da China como

potencial local de origem e expansão da radiação *virilis-repleta*, hipótese esta parcialmente suportada por Wang *et al.* (2006), que propuseram, de forma mais geral, os trópicos do Velho Mundo como ponto de origem desta seção. A partir daí, Wang *et al.* (2006) postulam a ocorrência de cinco migrações independentes para o Novo Mundo, via Estreito de Bering: (1) e (2) das phylads *virilis*; e *montana* do grupo *virilis*; (3) de *D. robusta*; (4) do grupo *repleta*; (5) do grupo *melanica*.

Segundo Pélandakis & Solignac (1993), Tatarenkov & Ayala (2001) e DaLage *et al.* (2007), a seção ou radiação *virilis-repleta* constitui um agrupamento monofilético de táxons, dada a exclusão do grupo *polychaeta* no primeiro, e do grupo *tumiditarsus* no segundo e no terceiro destes estudos. No trabalho de Remsen & O'Grady (2002), por outro lado, a seção *virilis-repleta* foi parafileticamente definida com relação ao subgênero *Siphodora*, dada a exclusão do grupo *polychaeta*. Os drosofilídeos havaianos, pertencentes tanto ao gênero *Drosophila*, quanto ao gênero *Scaptomyza*<sup>4</sup>, parecem ainda constituir a linhagem-irmã da seção *virilis-repleta* (Stalker, 1972; Russo *et al.*, 1995; Remsen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000; Remsen & O'Grady, 2002), embora trabalhos anteriores, realizados a partir de uma amostragem bastante reduzida de OTU's, tenham apresentado o subgênero *Drosophila* como grupo irmão dos drosofilídeos havaianos (Thomas & Hunt, 1991 e 1993; DeSalle, 1992a e b). Segundo Russo *et al.* (1995) a divergência entre os clados que originam a seção *virilis-repleta* e os drosofilídeos havaianos ocorreu há aproximadamente 32 milhões de anos.

A parte mais derivada da árvore de Tatarenkov & Ayala (2001), construída a partir de seqüências dos genes *Amd* e *Ddc* para membros da seção *virilis-repleta*,

---

<sup>4</sup> Previamente tido como um subgênero do gênero *Drosophila* (Wheeler, 1981), *Engiscaptomyza*, é considerado, atualmente, um gênero à parte (Grimaldi, 1990), ou até mesmo um subgênero do gênero *Scaptomyza* (Bächli, 2008).

apresenta o grupo *repleta*, monofileticamente definido, juntamente com os grupos *mesophragmatica*, *dreyfusi* e *canalineae*, de modo a recuperar a radiação *repleta* proposta por Throckmorton (1975). A proximidade filogenética entre estes grupos foi ainda confirmada por Durando *et al.* (2000) e DaLage *et al.* (2007). Segundo Tatarenkov & Ayala (2001), o clado constituído pelos grupos *bromeliae* e *nannopectera* constitui o grupo-irmão da radiação *repleta*, sendo o grupo *annulimana* a linhagem mais próxima do cluster assim definido. Dentre as ramificações mais basais apresentadas neste trabalho encontra-se o grupo *virilis*, o cluster constituído pelos grupos-irmãos *robusta* e *melanica*, e o grupo *polychaeta*, que se destaca como o primeiro a divergir dentro da seção *virilis-repleta*. A ramificação basal do grupo *polychaeta* foi recentemente recuperada por Wang *et al.* (2006) e DaLage *et al.* (2007). Entretanto, enquanto a base da radiação *virilis-repleta* apresentada por DaLage *et al.* (2007) assemelha-se ao cenário apresentado por Tatarenkov & Ayala (2001), no trabalho de Wang *et al.* (2006), o grupo *repleta*, junto aos grupos *annulimana* e *nannopectera*, constituem a segunda linhagem a divergir dentro da seção, de modo a constituir o cluster-irmão do clado que reúne os grupos *angor*, *virilis*, *robusta*, *melanica*, *quadrisetata* e *clefta* na árvore de evidências totais.

#### 1.2.1.1 – Grupo *annulimana*

O grupo *annulimana* é formado por 16 espécies descritas (Bächli, 2008). Este grupo foi subdividido em três subgrupos, diferenciáveis mediante inspeção de caracteres morfológicos externos: subgrupo com cerdas pré-escutelares, que inclui *D. annulimana*, *D. araicus*, *D. aragua* e *D. arauana*; subgrupo sem cerdas pré-escutelares, e com escutelares anteriores convergentes, formado por *D. gibberosa*, *D. ararama*, *D. talamancana* e *D. gigas*; subgrupo sem cerdas pré-escutelares, e com cerdas escutelares

anteriores divergentes, que abrange *D. arassari* e *D. arapuan* (Breuer & Pavan, 1950; Vilela & Pereira, 1982). Segundo Breuer & Pavan (1950) é extremamente difícil separar as espécies pertencentes a um mesmo subgrupo com base apenas em sua morfologia externa, sendo necessário, nesses casos, fazer uso da análise da morfologia da genitália dos machos. Val *et al.* (1981) destacam, entretanto, que esta subdivisão é extremamente artificial, dada a existência de uma espécie, *D. breuerae*, com características intermediárias entre o segundo e o terceiro subgrupos. Neste sentido, Breuer & Pavan (1950) ressaltam que, enquanto as espécies pertencentes ao subgrupo 1 e ao subgrupo 3 são extremamente semelhantes também em suas genitálias, *D. gibberosa* e *D. ararama*, do subgrupo 2, tem suas genitálias bastante modificadas. Mais recentemente, Tosi & Pereira (1993), com base em informações cariotípicas, confirmaram alguns dos achados de Breuer & Pavan (1950) ao propor três subgrupos para o grupo: o primeiro formado por *D. shineri*, *D. ararama*, *D. pseudotalamanca* e *D. gibberosa*; o segundo constituído por *D. arapuan* e *D. arassari*; e o terceiro composto por *D. annulimana*, *D. aragua*, *D. aracataca* e *D. arana*. *D. gigas* foi deslocada deste grupo e não está alocada a nenhum outro até o momento (Bächli, 2008).

Segundo Val *et al.* (1981), o primeiro e o terceiro subgrupos [se definidos conforme descrição original de Breuer & Pavan (1950)] apresentam uma distribuição exclusivamente Sul-Americana, enquanto o segundo subgrupo ocorre desde o México, até o Brasil e a Argentina.

Embora Malogolowkin (1953) tenha situado o grupo *annulimana* junto aos grupos *funebri*, *guttifera*, *immigrans* e *dreyfusi*, Throckmorton (1975) o transferiu para a base da radiação *virilis-repleta*, junto aos grupos *tumiditarsus*, *polychaeta* e *carbonaria*. Mais recentemente, entretanto, Tatarenkov & Ayala (2001) confirmaram a presença do grupo *annulimana* dentro da seção *virilis-repleta*, como ramificação basal

do clado constituído pela radiação *repleta* agrupada aos grupos-irmão *bromeliae* e *nannoptera*.

#### 1.2.1.2 – Grupo *bromeliae*

O grupo *bromeliae* apresenta apenas cinco espécies descritas, encontradas em regiões Neotropicais em associação com flores de determinadas espécies de plantas. Tatarenkov & Ayala (2001) apresentam o grupo *bromeliae* como irmão do grupo *nannoptera*, enquanto Remsen & O’Grady (2002) favorecem o grupo *mesophragmatica* como táxon irmão do grupo *bromeliae*. Segundo Throckmorton (1975), os grupos *nannoptera*, *bromeliae* e *peruviana*, juntamente aos membros da radiação *repleta* parecem ter se originado próximo à regiões do México.

#### 1.2.1.3 – Grupo *canalineae*

Um total de 11 espécies descritas corresponde ao grupo *canalineae* em sua composição atual (Bächli, 2008). Segundo Throckmorton (1975), estas espécies são essencialmente habitantes de florestas, cuja distribuição centraliza-se na América do Sul.

Este grupo de espécies foi localizado junto à radiação *repleta* de Throckmorton (1975), que favoreceu o grupo *dreyfusi* como seu táxon irmão. A afinidade filogenética entre os grupos *canalineae*, *dreyfusi*, *mesophragmatica* e *repleta* foi recentemente confirmada por outros autores (Durando *et al.*, 2000; O’Grady *et al.*, 2001; Tatarenkov & Ayala, 2001).

#### 1.2.1.4 – Grupo *carsoni*

Trata-se de um grupo monotípico, cuja única espécie *D. carsoni*, apresenta uma distribuição restrita aos Estados Unidos (Bächli, 2008). Este grupo foi situado junto aos grupos *robusta*, *melanica* e *virilis*, na radiação *virilis-repleta* (Throckmorton, 1975).

#### 1.2.1.5 – Grupo *coffeata*

Sinonimizado ao grupo *castanea* (Markow & O’Grady, 2006) e posicionado por Throckmorton (1975) junto à radiação *repleta*, este grupo perfaz atualmente um total de quatro espécies de distribuição exclusivamente Neotropical (Bächli, 2008).

#### 1.2.1.6 – Grupo *dreyfusi*

Outro membro da radiação *repleta* de Throckmorton (1975), o grupo *dreyfusi* é composto atualmente por nove espécies (Bächli, 2008), cuja distribuição centra-se em regiões sul-americanas (Throckmorton, 1975). Segundo Durando *et al.* (2000) e O’Grady *et al.* (2001), espécies deste grupo agrupam-se internamente ao grupo *repleta*, de modo a torná-lo um táxon parafilético. Embora em Tataronov & Ayala (2001) o grupo *repleta* tenha sido monofileticamente recuperado, o posicionamento do grupo *dreyfusi* foi inconstante mediante as análises com diferentes partições (*Amd* ou *Ddc* ou *Amd+Ddc*), ora situando-se como grupo externo ao clado composto pelos grupos *repleta* e *mesophragmatica* (*Amd*), ora agrupando-se como linhagem-irmã do grupo *mesophragmatica* (*Ddc* e *Amd+Ddc*).

#### 1.2.1.7 – Grupo *flavopilosa*

O grupo *flavopilosa* é constituído por um total de 17 espécies de distribuição exclusivamente Neotropical (Bächli, 2008), subdivididas em 2 subgrupos: *flavopilosa*,

que apresenta *D. acroria*, *D. crossoptera*, *D. lauta*, *D. cordeiroi*, *D. cestri*, *D. flavopilosa* e *D. hollisae*; e *nesiota*, que inclui *D. incompa*, *D. melina*, *D. nesiota*, *D. gentica*, *D. mariahelena*.

Todas as espécies do grupo *flavopilosa* apresentam uma ecologia restrita, na medida em que utilizam flores da família *Solanaceae* como único sítio de oviposição, desenvolvimento e alimentação (Brncic, 1978; Hofmann, 1985). São, portanto, espécies monófagas que utilizam como suas hospedeiras flores, abertas ou fechadas, de 11 diferentes espécies de *Cestrum* e *Sessea brasiliensis* (Santos & Vilela, 2005), sendo que diferentes espécies do grupo apresentam preferências por espécies particulares de planta, havendo ampla sobreposição, mas diferentes níveis de restrição. Neste sentido, a distribuição das espécies do grupo *flavopilosa* de *Drosophila* depende amplamente da distribuição de suas espécies hospedeiras, que, em sua maior parte, ocorrem em relativa abundância na região Neotropical (Hofmann, 1985). No Brasil, seis espécies de *Drosophila* do grupo *flavopilosa* foram encontradas (*D. cestri*, *D. cordeiroi*, *D. flavopilosa*, *D. hollisae*, *D. incompta* e *D. mariahelena*) em um total de sete diferentes espécies de *Cestrum* (*C. amictum*, *C. calycinum*, *C. corymboum*, *C. intermedium*, *C. parqui*, *C. schlenchtendalii* e *C. sendtnerianum*) e *Sessea brasiliensis* (Santos & Vilela, 2005).

Embora o período de florescimento destas diferentes espécies de plantas varie ao longo do ano, podendo haver sobreposições, uma questão que permanece em aberto acerca da ecologia do grupo *flavopilosa* refere-se a sua sobrevivência nos meses de inverno, ou privados de recursos. Neste sentido, em alguns locais, espécies de *Cestrum* com um florescimento quase que contínuo foram encontradas (Santos & Vilela, 2005), sendo que alguns autores defendem ainda que alguns adultos tenham a capacidade de

sobreviver ao longo do inverno, de modo a fundar novas populações quando o período de escassez terminar (Hofmann, 1985).

Throckmorton (1975) situa as espécies do grupo *flavopilosa* em sua radiação *virilis-repleta*, deixando seu relacionamento com os demais grupos como uma questão em aberto. Neste sentido, este autor sugere que o grupo *flavopilosa* possa ser derivado, tanto da radiação Neotropical que dá origem aos grupos *peruviana*, *bromeliae* e *nannoptera*, quanto da radiação *repleta*. Grimaldi (1990), por outro lado, apresenta o grupo *flavopilosa* junto ao clado formado essencialmente por scaptomizóides hawaianos, levantando a possibilidade de que o mesmo sequer faça parte do gênero *Drosophila*.

#### 1.2.1.8 – Grupo *melanica*

O grupo *melanica* encontra-se constituído por 13 espécies (Bächli, 2008), embora Wang *et al.* (2006) tenham demonstrado que a inclusão de *D. moriwakii* torna este grupo parafilético com relação a representantes do grupo *robusta*. As demais espécies do grupo *melanica* parecem, entretanto, formar um agrupamento monofilético, subdividido em dois clados maiores: o primeiro constituído por membros do Velho Mundo (*D. tsigana* e *D. longiserrata*) e outro por representantes do Novo Mundo (*D. micromelanica*, *D. melanica*, *D. paramelanica* e *D. euronotus*) (Wang *et al.*, 2006). Dentre os representantes do Novo Mundo, Stalker (1966), em sua filogenia estabelecida com base em dados citogenéticos, posiciona *D. micromelanica* e *D. nigromelanica* como basais em relação a *D. paramelanica*, *D. melanica* e *D. euronotus*.



### 1.2.1.9 – Grupo *mesophragmatica*

O grupo *mesophragmatica*, descrito originalmente por Brncic & Koref-Santibañez (1957), compreende cerca de 13 espécies de distribuição exclusivamente Neotropical (Bächli, 2008). Quanto à sua distribuição, com exceção de *D. gaucha*, que se distribui até o sul e sudeste do Brasil, todas as demais espécies do grupo são essencialmente Andinas (Brncic *et al.*, 1971). As espécies pertencentes ao grupo *mesophragmatica* são bastante similares morfológicamente, sendo que algumas delas, como *D. gaucha* e *D. pavani*, são espécies crípticas, indistinguíveis mediante inspeção de traços morfológicos, inclusive genitália externa (Nacrur, 1958).

Segundo Brncic *et al.* (1958 e 1971), variados graus de isolamento reprodutivo podem ser encontrados entre as espécies do grupo, de forma que *D. gaucha* e *D. pavani*, bem como *D. altiplanica* e *D. orkui*, quando cruzadas em condições laboratoriais, produzem híbridos estéreis, enquanto o cruzamento entre *D. mesophragmatica* e *D. pavani* ou *D. gaucha* resulta, normalmente, na produção de híbridos inviáveis. Por outro lado, *D. gasici*, *D. brncici* e *D. viracochi* parecem ser totalmente isoladas reprodutivamente (Brncic *et al.*, 1971).

Nacrur (1958) propõe a subdivisão do grupo *mesophragmatica* em dois subgrupos, com base em caracteres morfológicos externos, dado que este autor não encontrou maiores diferenças mediante inspeção de suas genitálias. Esta proposição foi recentemente validada e renomeada por Vela e Rafael (2004), segundo os quais os subgrupos apresentam a seguinte composição: subgrupo *mesophragmatica*, com cerdas escutelares anteriores divergentes, composto por *D. gasici*, *D. brncici*, *D. mesophragmatica*, *D. gaucha*, *D. pavani*, *D. altiplanica*, *D. orkui*, *D. amaguana*, *D. canescens*, *D. camaronensis* e *D. shyri*; subgrupo *viracochi*, com cerdas escutelares convergentes, composto apenas por *D. viracochi* e *D. ruminahuii*. Entretanto, Brncic *et*

*al.* (1971) e Nair *et al.* (1971), com base em dados citogenéticos, e de isozimas, demonstraram a subdivisão do grupo em três linhagens principais: a primeira formada por *D. gasici*, *D. brncici* e *D. mesophragmatica*, a segunda contendo *D. gaucha* e *D. pavani*, e a terceira, monoespecífica, contendo *D. viracochi*. Estas linhagens coincidem com a proposição inicial de Brncic & Koref-Santibañez (1957), segundo os quais o grupo apresenta três claras subdivisões [apresentadas em ordem de correspondência com a ordem acima apresentada para os dados de Brncic *et al.* (1971) e Nair *et al.* (1971)]: a primeira, com espécies escuras e com cerdas escutelares divergentes, a segunda com espécies mais claras e com cerdas escutelares divergentes e a terceira, cujas cerdas escutelares são convergentes. Ainda segundo Brncic *et al.* (1971) e Nair *et al.* (1971), a linhagem *gaucha/pavani* é mais diferenciada, geneticamente, das demais.

Segundo Throckmorton (1975), este grupo é proximamente relacionado aos grupos *repleta*, *coffeata* (= *castanea*), *canalinea* e *dreyfusi*, de modo a compor um dos membros da radiação *repleta*, resultados estes confirmados recentemente na análise executada por Tatarenkov & Ayala (2001). Neste trabalho, o grupo *mesophragmatica* foi posicionado como irmão do grupo *repleta* (nas análises realizadas com base em seqüências do gene *Amd*), ou como irmão do grupo *dreyfusi* (nas análises realizadas a partir de *Ddc* e de *Amd+Ddc*). Remsen & O'Grady (2002), entretanto, situam o grupo *mesophragmatica* como irmão do grupo *bromeliae*.

#### 1.2.1.10 – Grupo *nannoptera*

Trata-se de um grupo pequeno, com quatro espécies descritas (Bächli, 2008), que ocorrem em regiões desérticas ao longo do continente americano. Suas relações filogenéticas foram estabelecidas por Ward & Heed (1970), que, a partir da análise de

cromossomos politênicos, apresentam *D. wassermani* e *D. nannoptera* como espécies irmãs derivadas com relação a *D. acanthoptera* e *D. pachea*.

Throckmorton (1975), bem como achados recentes de Tatarenkov & Ayala (2001) favorecem a visão de que os grupos *nannoptera* e *bromeliae* são irmãos entre si.

#### 1.2.1.11 – Grupo *peruviana*

O grupo Neotropical monoespecífico *peruviana* também tem sua distribuição associada a flores em regiões desérticas. Este grupo foi posicionado por Throckmorton (1975) na radiação *virilis-repleta*, como intimamente relacionado aos grupos *bromeliae* e *nannoptera*.

#### 1.2.1.12 – Grupo *polychaeta*

Composto atualmente por sete espécies (Bächli, 2008), o grupo *polychaeta* teve sua monofilia recentemente validada (Wang *et al.*, 2006; DaLage *et al.*, 2007). O trabalho de Wang *et al.* (2006) também forneceu suporte para a proximidade filogenética entre o grupo *polychaeta* e *D. fluvialis*, uma espécie tropical até o momento não designada a nenhum grupo, sendo que o clado assim constituído estabelece uma das ramificações basais da radiação *virilis-repleta*. Resultados semelhantes foram também sugeridos por Throckmorton (1975), Tatarenkov & Ayala (2001) e DaLage *et al.* (2007) com relação ao grupo *polychaeta*. As espécies deste grupo apresentam, em geral, uma distribuição tropical disjunta, de forma que, embora a maioria das espécies seja Asiática, *D. hirtipes* é encontrada na África e *D. illota* nos Neotrópicos.

### 1.2.1.13 – Grupo *repleta*

O grupo *repleta* compreende, atualmente, pelo menos 100 espécies, subdivididas em seis subgrupos (Bächli, 2008): *fasciola* (21 espécies), *hydei* (7 espécies), *inca* (3 espécies), *mercatorum* (4 espécies), *mulleri* (50 espécies) e *repleta* (9 espécies). A taxonomia clássica divide, ainda, estes subgrupos em complexos, clusters e subclusters. Espécies pertencentes ao subgrupo *hydei*, por exemplo, encontram-se agrupadas nos complexos *bifurca* e *hydei*, enquanto que o subgrupo *repleta* tem suas espécies agrupadas nos complexos *fulvimacula* e *repleta* (Bächli, 2008). Já o subgrupo *mulleri* encontra-se subdividido em dez complexos de espécies: *anceps*, *buzzatii*, *eremophila*, *longicornis*, *martensis*, *meridiana*, *mojavensis*, *mulleri*, *ritae* e *stalker* (Bächli, 2008). De acordo Wang *et al.* (2006), ancestrais do grupo *repleta* devem ter evoluído nos trópicos do Velho Mundo, a partir de onde se dispersaram e, atingindo o Novo Mundo, finalmente se diversificaram, sendo que os representantes que permaneceram no Velho Mundo extinguíram-se. Com isso, o grupo *repleta* é, atualmente, um grupo tipicamente Neotropical (Val *et al.*, 1981), onde distribui-se por zonas áridas e semi-áridas, em associação a diversas espécies de cactus, embora representantes do subgrupo *fasciola* sejam encontrados em áreas mais úmidas da América Central e da América do Sul (Throckmorton, 1975). Throckmorton (1975) associou a diversificação do grupo com o desenvolvimento dos desertos americanos. Datações de Russo *et al.* (1995), postulam que o grupo *repleta* originou-se há cerca de 18 milhões de anos.

Muitas espécies deste grupo apresentam, entre si, variados graus de isolamento reprodutivo, tornando-as organismos-modelo adequados para estudos de especiação. Este é o caso, por exemplo, das três espécies pertencentes ao cluster *mojavensis* do subgrupo *mulleri*: enquanto que o cruzamento de *D. arizonae* com *D. navojoa*, resulta em híbridos estéreis (Ruiz *et al.*, 1990), *D. mojavensis* e *D. arizonae*, duas espécies

críticas entre si, produzem híbridos férteis em condições laboratoriais (embora híbridos não tenham sido coletados em ambiente natural). Entretanto, recentemente, Oliveira *et al.* (2003) demonstraram que eventos de reticulação recente podem estar desempenhando um papel fundamental na evolução destas espécies. Morán & Fontdevila (2005) também recorrem a fenômenos antigos de introgressão, bem como à possível retenção de polimorfismo ancestral, na tentativa de explicar as incongruências encontradas em sua análise filogenética do subgrupo *hydei*, já que eventos internos de hibridação foram previamente notificados (Wasserman, 1992).

Durando *et al.* (2000) e O'Grady *et al.* (2001) sugerem que possivelmente o grupo *repleta* é parafilético com relação aos grupos *mesophragmatica*, *dreyfusi* e *canalinae*, embora Durando *et al.* (2000) admitam que esta inferência pode ser resultante da resolução deficiente apresentada na base de sua filogenia. Neste último trabalho, este resultado é, ainda, pobremente suportado, dado que o acréscimo de apenas dois passos ao comprimento da árvore de máxima parcimônia torna o grupo *repleta* monofilético. Dadas estas considerações, Katoh *et al.* (2000), Tatarenkov & Ayala (2001), Wang *et al.* (2006) e DaLage *et al.* (2007), recuperam a monofilia do grupo *repleta*, embora este resultado deva ser interpretado com cautela, dado o baixo número de espécies incluídas nestes trabalhos (oito, oito, doze e seis, respectivamente).

Segundo Tatarenkov & Ayala (2001), o subgrupo *fasciola* é o primeiro a divergir dentre os membros do grupo *repleta*, enquanto o clado constituído pelos subgrupos-irmãos *hydei* e *mulleri* apresenta-se como derivado em relação aos demais. Já os subgrupos *repleta* e *mercatorum* apresentam um posicionamento intermediário nas filogenias obtidas por Tatarenkov & Ayala (2001), aparecendo, ora como táxons-irmãos, ora como ramificações independentes, embora Throckmorton (1975) e Wasserman (1992) defendam o primeiro destes cenários. A afinidade filogenética entre

os subgrupos *repleta* e *mercatorum* foi, ainda, suportada por Durando *et al.* (2000). Neste trabalho, entretanto, apenas os subgrupos *mercatorum* e *hydei* são monofileticamente recuperados, de forma que o subgrupo *repleta* é parafileticamente definido com relação ao subgrupo *mercatorum*, enquanto o subgrupo *mulleri* representa um agrupamento polifilético de espécies.

De acordo com Durando *et al.* (2000), internamente ao subgrupo *mulleri*, os complexos *mulleri*, *buzzatii*, *eremophila* e *meridiana* são monofiléticos, enquanto que *anceps* apresenta seu status taxonômico não-resolvido. Ainda segundo este autor os complexos *mulleri* e *buzzatii* constituem táxons-irmãos entre si, enquanto os demais complexos apresentam ramificações independentes. Neste caso, o subgrupo *eremophila* parece constituir o clado irmão do subgrupo *hydei*, resultado este também suportado por Russo *et al.* (1995). Dentro do complexo *mulleri* recuperado por Durando *et al.* (2000), membros dos clusters *mojavensis* misturam-se ao cluster *mulleri*, da mesma forma que os clusters *ritae* e *longicornis*. No complexo *buzzatii*, por outro lado, os três clusters são monofileticamente definidos, sendo que *stalker* representa a ramificação basal, e *martensis* e *buzzatii* são clusters-irmãos entre si. Rodríguez-Trelles *et al.* (2000c) apresenta um resultado similar a este, exceto que neste caso *stalker* apresenta uma composição parafiletica bem suportada.

A filogenia do subgrupo *hydei* foi recentemente estudada por Morán & Fontdevila (2005 e 2006). No primeiro destes estudos, tanto o complexo *bifurca*, quanto o complexo *hydei* foram parafileticamente definidos com relação um ao outro. Entretanto, no trabalho de 2006, estes autores consistentemente validam a monofilia do complexo *hydei*, contrastada pela reiteração da parafilia do complexo *bifurca*.

#### 1.2.1.14 – Grupo *robusta*

O grupo *robusta* apresenta, atualmente, 16 espécies descritas, subdivididas em três subgrupos (Bächli, 2008): *lacertosa* (sete espécies), *okadai* (três espécies) e *robusta* (quatro espécies). Entretanto, Wang *et al.* (2006) não recuperaram a monofilia de nenhum destes subgrupos, sendo que o grupo *robusta* apresentou, neste trabalho, uma constituição parafilética com relação aos grupos *quadrisetata*, *clefta* e *melanica*. De modo geral, estes autores apresentam o subgrupo *robusta* como táxon-irmão do grupo *melanica*, enquanto os subgrupos *lacertosa* e *okadai* aparecem agrupados aos grupos *quadrisetata* e *clefta*. A proximidade filogenética entre os grupos *robusta* e *melanica* foi, ainda, sugerida por Tatarenkov & Ayala (2001), enquanto Stalker (1972) consideram o grupo *robusta* como táxon irmão dos drosofilídeos havaianos.

Segundo Markow & O’Grady (2006), espécies do grupo *robusta* são encontradas em regiões temperadas da América do Norte, da Europa e da Ásia, apresentando pelo menos uma espécie Afrotropical. Para Wang *et al.* (2006) o padrão de especiação evidenciado em sua filogenia pode refletir especializações a habitats diferentes, uma vez que o subgrupo *robusta* e o grupo *melanica* parecem preferir regiões mais distantes de linhas d’água, dentro das florestas temperadas, enquanto que os subgrupos *okadai* e *lacertosa*, bem como os grupos *quadrisetata* e *clefta* apresentam uma distribuição mais ribeirinha.

#### 1.2.1.15 – Grupo *virilis*

O grupo *virilis* é formado por 12 espécies descritas (Bächli, 2008), que alimentam-se, principalmente de seiva de plantas (Pringent *et al.*, 2003). Dentre as espécies que compõem o grupo *virilis*, apenas *D. virilis* apresenta uma distribuição que se estende até a região Neotropical (Val *et al.*, 1981). As demais apresentam uma

distribuição tipicamente boreal (Markow & O'Grady, 2006). O grupo *virilis* encontra-se atualmente subdividido em quatro subgrupos (Spicer, 1991): *virilis*, que inclui *D. virilis*, *D. lummei*, *D. novamexicana*, *D. americana* (subespécies *texana* e *americana*); *montana*, formado por *D. montana*, *D. lacicola*, *D. flavomontana* e *D. borealis*; *litoralis*, formado por *D. littoralis*, *D. ezoana* e *D. canadiana*; e a linhagem monotípica *kanekoi*, composta por *D. kanekoi*. Cruzamentos entre espécies pertencentes a este grupo ilustram gradações completas de isolamento reprodutivo pós-copulatório, desde a inviabilidade dos híbridos até sua completa fertilidade (Throckmorton, 1982). Apoiados por estes resultados, Nurminsky *et al.* (1996) defendem que o grupo *virilis* diversificou-se recentemente, há apenas 9 milhões de anos atrás. Este trabalho suporta ainda a monofilia dos clados *virilis* e *montana*, sugerindo sua equivalência a subgrupos de espécies.

Considerado como um todo, o grupo *virilis* teve sua monofilia recentemente suportada por Wang *et al.* (2006) e DaLage *et al.* (2007). No primeiro destes trabalhos os quatro agrupamentos de espécies previamente propostos por Spicer (1991) foram devidamente recuperados, sendo que o clado *montana* destacou-se como a ramificação basal dentro do grupo. As relações entre os demais clados não foram, entretanto, definidas, dado o posicionamento inconstante de *D. kanekoi*.

Caletka & McAllister (2004), por outro lado, examinaram as relações filogenéticas entre espécies cromossomicamente variáveis do subgrupo *virilis*, a partir de marcadores mitocondriais. Seus resultados suportam a ramificação basal das espécies Paleárticas *D. virilis* e *D. lummei*, sendo que esta última constitui a linhagem irmã do clado Neártico formado por representantes das espécies *D. novamexicana* e *D. americana*. Relacionamentos semelhantes foram também recuperados por O'Grady *et al.* (2001), com base na análise combinada de seqüências de DNA e marcadores



citogenéticos, bem como por DaLage *et al.* (2007). Com base nestas relações, Caletka & McAllister (2004) estipulam que a radiação Neártica do subgrupo se deu a partir de eventos de dispersão de ancestrais compartilhados com *D. lummei*, a partir do Estreito de Bering, há cerca de três milhões de anos atrás. Este estudo também demonstrou a ausência de divergência entre as subespécies cromossomicamente definidas de *D. americana*, de forma que esta subdivisão não é substanciada. Neste caso, o aparecimento das duas formas cromossômicas deve ter ocorrido recentemente, bem depois do último evento de coalescência de seus alelos mitocondriais.

Com relação aos demais subgrupos, O'Grady *et al.* (2001) foram capazes de recuperar a monofilia do subgrupo *montana*, enquanto o subgrupo *litoralis* apresentou uma composição polifilética.

#### 1.2.1.16 – Grupo *tumiditarsus*

O grupo *tumiditarsus* perfaz outro grupo monoespecífico, cuja única espécie *D. repletoides*, é encontrada em regiões da China e Japão. Embora Throckmorton (1975) tenha situado o grupo *tumiditarsus* junto à radiação *virilis-repleta*, trabalhos recentes executados por Tatarenkov & Ayala (2001) e DaLage *et al.* (2007) não corroboram este posicionamento.

#### 1.2.2 – Seção *quinaria-tripunctata*

Já a seção *quinaria-tripunctata* é constituída pelos grupos *calloptera*, *cardini*, *guarani*, *immigrans*, *macroptera*, *pallidipennis*, *pinicola*, *quinaria*, *rubrifrons sticta*, *testacea* e *tripunctata*. Segundo Throckmorton (1975), representantes desta seção compõem a chamada radiação *immigrans-Hirtodrosophila*, que serve como base a partir da qual, em diferentes momentos e níveis, diversos outros táxons se originam. Entre os

gêneros e subgêneros destacados por Throckmorton (1975) como participantes desta radiação, pode-se citar: *Zaprionus*, *Samoaia*, *Hirtodrosophila*, *Mycodrosophila*, *Zygothrica*, *Scaptomyza*, *Dorsilopha*, etc.

#### 1.2.2.1 - Grupo *calloptera*

O grupo *calloptera* apresenta atualmente um total de oito espécies essencialmente latino-americanas (Bächli, 2008), sendo *D. calloptera* seu representante mais amplamente distribuído, encontrado desde o México até o sul do Brasil (Val *et al.*, 1981). *D. ornatipennis*, por outro lado, tem sua distribuição restrita às Ilhas do Caribe (Bächli, 2008).

Bastante peculiares morfologicamente, as espécies do grupo *calloptera* são caracterizadas pela presença de amplas ornamentações nas asas e no mesonoto (Burla & Pavan, 1953). Segundo Val *et al.* (1981), as espécies do grupo *calloptera* podem ser subdivididas em dois clusters: o primeiro formado por *D. kallima*, *D. lindae*, *D. quadrum* e *D. schildi*; o segundo constituído por *D. atrata*, *D. calloptera*, *D. maracaya* e *D. ornatipennis*. Mesmo sendo um grupo bastante reduzido em termos de número de espécies, *calloptera* teve sua monofilia questionada por Yotoko *et al.* (2003), sendo que os dois clusters recuperados neste trabalho também não suportam a subdivisão acima apresentada.

Throckmorton (1975) posicionou o grupo *calloptera* dentro da radiação *tripunctata*, resultado este recentemente confirmado por Remsen & O'Grady (2002) e por Yotoko *et al.* (2003). Entretanto, enquanto Remsen & O'Grady (2002) sugerem o grupo *guarani* como táxon irmão do grupo *calloptera*, Yotoko *et al.* (2003) apresenta dois membros do grupo *calloptera* (*D. calloptera* e *D. schildi*) agrupados com membros do grupo *guarani* (*D. maculifrons* e *D. griseolineata*), enquanto a terceira espécie

analisada (*D. atrata*) agrupou-se com *D. cuaso* em um clado independente. Estes resultados são parcialmente concordantes com as proposições de Burla & Pavan (1953), que acreditavam que o grupo *calloptera* era proximoamente relacionado aos grupos *tripunctata*, *quinaria*, *guarani* e *cardini*, mas apresentavam o grupo *guarani* como táxon-irmão.

#### 1.2.2.2 - Grupo *cardini*

O grupo *cardini* apresenta um total de 16 espécies descritas, embora uma delas, *D. bedicheki*, tenha sido coletada uma única vez (em 1964, em Trinidad, por S. Bedichek Pipkin de acordo com Heed & Russel, 1971). Este grupo encontra-se distribuído ao longo da América Neotropical, inclusive Ilhas do Caribe, sendo caracterizado por apresentar um amplo espectro de variação na pigmentação abdominal, no que diz respeito tanto aos padrões quanto à intensidade de pigmentação. Heed (1962) subdividiu as espécies deste grupo em dois subgrupos: subgrupos *dunni*, que inclui sete espécies (*D. belladunni*, *D. dunni*, *D. similis*, *D. caribiana*, *D. antillea*, *D. arawakana*, e *D. nigrodunni*), e *cardini*, que apresenta atualmente outras nove espécies (*D. bedicheki*, *D. cardini*, *D. cardinoides*, *D. parthenogenetica*, *D. neomorpha*, *D. polymorpha*, *D. neocardini*, *D. procardinoides* e *D. acutilabella*). Entre estas espécies, os pares *D. bedicheki* e *D. cardinoides*, *D. parthenogenetica* e *D. procardinoides*, e *D. neomorpha* e *D. polymorpha*, são compostos por espécies crípticas (Markow & O'Grady, 2006).

As espécies do subgrupo *dunni* apresentam uma distribuição endêmica para as Ilhas do Caribe, sendo que cada forma (espécie ou subespécie) distribui-se, exclusivamente, em uma única Ilha (Heed, 1962) (*D. belladunni* é endêmica da Jamaica; *D. dunni* apresenta uma subespécie em Porto Rico e outra em St. Thomas; *D. similis* tem uma subespécie em St. Vincent e outra em Grenada; *D. caribiana* é

endêmica de Martinique; *D. antillea* é endêmica de St. Lucia; *D. arawakana* tem uma subespécie em St. Kitts e outra em Guadeloupe; e *D. nigrodunni* é endêmica de Barbados). Neste caso, o estabelecimento de subespécies se deu com base em diferenças intra-específicas, com fertilidade em alopatria, quanto aos resultados dos cruzamentos com outras espécies (*D. similis*), ou com base em diferenças morfológicas e citogenéticas (*D. arawakana* e *D. dunni*) (Heed, 1962). Segundo Brisson *et al.* (2006) este subgrupo é caracterizado pela presença de um padrão monomórfico de pigmentação que, analisado interespecificamente, acaba por formar uma clina norte-sul, com espécies escuras ocupando regiões mais ao sul e espécies mais claras distribuídas nas Ilhas mais ao norte. Heed & Krishnamurthy (1959) e Heed (1962) atestam que diversos graus de isolamento reprodutivo são exibidos nos cruzamentos envolvendo os membros do subgrupo *dunni*, indo desde a fertilidade completa (como entre as duas subespécies de *D. dunni*, *D. similis* ou *D. arawakana*) ou parcial (como entre *D. arawakana* e *D. similis grenadensis*), até certo isolamento sexual (como entre representantes de *D. dunni* e *D. similis*) e esterilidade completa / inviabilidade dos híbridos (como em qualquer cruzamento envolvendo *D. caribiana*). Segundo Heed (1962), o mecanismo de isolamento mais comum para o grupo é a esterilidade dos machos, de forma que praticamente todos os cruzamentos interespecíficos produzem machos híbridos estéreis. Outra característica freqüentemente encontrada nestes cruzamentos interespecíficos é a presença de uma clina de razão sexual desigual nos híbridos de diferentes cruzamentos, cuja severidade aumenta com o aumento da distância entre as ilhas (Heed & Krishnamurthy, 1959; Heed, 1962). Segundo Heed (1962), alguns destes casos parecem resultantes do efeito Haldane, enquanto outros parecem estar relacionados a efeitos maternos. Além disso, larvas do subgrupo *dunni* exibem uma certa taxa de canibalismo (Heed & Krishnamurthy, 1959).

O subgrupo *cardini*, por outro lado, foi estabelecido para abranger todas aquelas espécies do grupo *cardini* não incluídas no subgrupo *dunni* (Heed, 1962). Ele apresenta espécies amplamente distribuídas em áreas americanas continentais, principalmente na América do Sul e Central, bem como *D. acutilabella*, encontrada desde os Estados Unidos, México e Cuba até o Panamá, bem como nas Ilhas de Hispaniola e Jamaica (Bächli, 2008). Todas as espécies do subgrupo *cardini*, com exceção de *D. procardinoides*<sup>5</sup>, apresentam um amplo grau de polimorfismo intra-específico para os padrões de pigmentação abdominal (Brisson *et al.*, 2006). Além disso, o subgrupo *cardini* apresenta uma maior frequência de inversões, quando comparado ao subgrupo *dunni* (Heed & Russell, 1971). Dentro deste subgrupo, mecanismos de isolamento reprodutivo parecem ser mais rígidos e eficientes, sendo que, como suas espécies são freqüentemente encontradas em simpatria, o isolamento sexual parece ser mais freqüente (Heed, 1962). A produção de híbridos, neste caso, restringe-se aos cruzamentos envolvendo *D. acutilabella* e outras espécies (*D. cardinoides*, *D. parthenogenetica*, *D. polymorpha* e *D. neocardini*), bem como aqueles envolvendo a tríade *D. parthenogenetica*, *D. cardinoides* e *D. procardinoides*, ou o par *D. polymorpha*, *D. neomorpha* (Heed, 1962; Heed & Russell, 1971). Em cruzamentos envolvendo espécies dos subgrupos *cardini* e *dunni*, híbridos foram produzidos apenas quando o subgrupo continental era representado por *D. acutilabella* (Heed, 1962). *D. cardini*, *D. cardinoides* e *D. acutilabella* são caracterizadas, ainda, como as espécies mais amplamente polimórficas, possivelmente ancestrais em relação a subconjuntos formados pelas demais espécies (Heed & Russell, 1971).

---

<sup>5</sup> Curiosamente, a única espécie do subgrupo a apresentar uma distribuição restrita [às altas elevações da Bolívia e do Peru, segundo Brisson *et al.* (2006)]. Isto está em inteira concordância à hipótese de que adaptações locais expliquem a configuração em que espécies de distribuição restrita apresentam padrões monomórficos de pigmentação em contraste às espécies de distribuição ampla, mais polimórficas para estes padrões.

Diversas hipóteses filogenéticas foram desenvolvidas para o grupo *cardini* ao longo dos últimos 50 anos, tanto com base na análise dos padrões de isolamento reprodutivo (Heed & Krishnamurthy, 1959) e de morfologia de genitália masculina (Heed, 1962), quanto com base em padrões de inversão cromossômica (Heed & Russel, 1971) e de similaridade genética em nível de isozimas (Napp & Cordeiro, 1981). Estas hipóteses são particularmente conflitantes no que diz respeito ao posicionamento de *D. neocardini*: esta espécie é apresentada como membro basal do clado contendo *D. acutilabella* bem como todas as espécies do subgrupo *dunni* no trabalho de Heed & Russel (1971), ou como membro do clado que agrupa ainda *D. neomorpha* e *D. polymorpha* segundo a análise de morfologia de genitália, realizada por Heed (1962). Os dados combinados de Brisson *et al.* (2006), entretanto, constituídos de seqüências de três genes nucleares (*pgd* – fosfogluconato desidrogenase, *sod* – Cu, Zn superóxido dismutase, *mlc* – cadeia leve da miosina) e três mitocondriais (*COII* – citocromo oxidase II, *cytB* – citocromo B, *16S* – RNA ribossomal 16S), rejeitaram todas as hipóteses anteriores. A filogenia de evidências totais obtida neste trabalho para o grupo *cardini* apresenta *D. cardini* como ramificação basal, no que se segue a divergência do clado contendo *D. polymorpha/D. neomorpha* e, posteriormente, do clado contendo *D. neocardini/D. cardinoides/D. parthenogenetica/D. procardinoides*. O último membro do subgrupo *cardini*, *D. acutilabella*, agrupa-se com *D. belladunni*, do subgrupo *dunni*, de modo a constituir a linhagem irmã do clado que contém os demais representantes do subgrupo *dunni*. A afinidade filogenética entre *D. acutilabella* e *D. belladunni* com os membros do subgrupo *dunni*, bem como a proximidade evolutiva entre as linhagens de *D. arawakana*, *D. caribiana*, *D. antillea* e *D. nigrodunni* entre si, foram ainda previamente sugeridos por Heed & Krishnamurthy (1959). Heed (1962), por outro lado, afirma que *acutilabella* assemelha-se tanto à tríade *neocardini/polymorpha/neomorpha*

quanto ao subgrupo *dunni*, sendo que esta tríade, ou até mesmo *D. neocardini* isoladamente, é proposta como elo de ligação entre o subgrupo continental e insular, enquanto o complexo *nigrodunni*, que envolve *D. arawakana*, *D. caribiana*, *D. antillea* e *D. nigrodunni*, é caracterizado como críptico inclusive no que diz respeito a padrões de genitália. Resultados apresentados por Heed & Russel (1971) demonstram ainda que as espécies deste complexo são homossequenciais no que diz respeito aos arranjos cromossômicos fixados. Este trabalho apresenta ainda padrões de evolução comum para o par *D. polymorpha* e *D. neomorpha*, que aparece como irmão do agrupamento formado por *D. cardinoides*, *D. bedichecki*, *D. parthenogenetica* e *D. procardinoides*. Neste sentido, *D. neocardini* é posicionada como basal em relação ao clado que apresenta *D. acutilabella* como linhagem irmã do subgrupo *dunni*. Napp & Cordeiro (1981), por outro lado, ao mesmo tempo em que concordam com o agrupamento do complexo *nigrodunni*, posicionam *D. dunni* junto a *D. neomorpha*, de forma a não apresentar suporte à monofilia do subgrupo *dunni*. Ainda neste caso, *D. cardini* é agrupada com *D. cardinoides*, enquanto *D. parthenogenetica* e *D. procardinoides* são agrupadas com *D. polymorpha*.

Desta forma, a monofilia do subgrupo *cardini* foi rejeitada por Brisson *et al.* (2006), segundo os quais espécies deste subgrupo divergiram umas das outras há mais de 6,6 milhões de anos atrás. Estes autores defendem, ainda, a transferência de *D. belladunni* para o subgrupo *cardini*, de modo a tornar o subgrupo *dunni* monofilético. Como *D. belladunni* e *D. acutilabella*, constituintes da linhagem irmã deste subgrupo, encontram-se distribuídas na América do Norte e Grandes Antilhas, uma origem setentrional é postulada para os membros do subgrupo *dunni* (Heed & Krishnamurthy, 1959). A ramificação basal de *D. cardini* é, ainda, fortemente suportada por testes de cruzamento, que a destacam como única espécie totalmente isolada reprodutivamente

(Heed, 1962); bem como por dados citológicos, que demonstram que esta é a única espécie do grupo a apresentar seis pares de cromossomos metafásicos (uma vez que as demais espécies do grupo apresentam quatro pares de cromossomos), bem como arranjos cromossômicos semelhantes aos inferidos para a seqüência ancestral ao grupo (Heed & Russel, 1971).

Quanto ao relacionamento com os demais grupos do subgênero *Drosophila*, o grupo *cardini* tem-se revelado um agrupamento monofilético (Yotoko *et al.*, 2003; Brisson *et al.*, 2006; DaLage *et al.*, 2007), situado junto à radiação *immigrans-tripunctata* (Remsen & O'Grady, 2002; Yotoko *et al.*, 2003; DaLage *et al.*, 2007) proximamente relacionado aos grupos *calloptera* (Remsen & O'Grady, 2002), *guarani* (Remsen & O'Grady, 2002; DaLage *et al.*, 2007), *tripunctata* e *pallidipennis* (Yotoko *et al.*, 2003; DaLage *et al.*, 2007).

#### 1.2.2.3 – Grupo *funnebris*

Embora Throckmorton (1975) tenha apresentado este grupo como basal em relação às radiações *virilis-repleta* e *immigrans-tripunctata*, a maior parte dos trabalhos posteriores (Kato *et al.*, 2000; Tatarenkov *et al.*, 2001; Remsen & O'Grady, 2002; Yotoko *et al.*, 2003; DaLage *et al.*, 2007) o tem situado junto aos representantes da radiação *immigrans-tripunctata*. Atualmente composto por sete espécies (Bächli, 2008), entre as quais *D. funnebris* apresenta uma distribuição cosmopolita, e *D. macrospina* e *D. subfunnebris* parecem ser principalmente Neárticas. *D. pentaspina* apresenta registros para a Índia e *D. multispina* distribui-se ao longo da Índia, da China, do Japão e da Rússia. *D. altukhovi* também ocorre na Rússia. São, em sua maioria, espécies micófagas (Pringent *et al.*, 2003).



#### 1.2.2.4 - Grupo *guarani*

O grupo *guarani*, originalmente estabelecido por Dobzhansky & Pavan (1943), abrange, atualmente, um total de 16 espécies (Bächli, 2008), subdivididas em dois subgrupos (King, 1947): *guaramunu* (3 espécies), constituído por *D. griseolineata*, *D. guaraja* e *D. maculifrons* (= *D. guaramunu*, segundo Vilela & Bächli, 1990); e *guarani* (5 espécies), formado por *D. guaru*, *D. ornatifrons* (= *D. guarani*, segundo Vilela & Bächli, 1990), *D. subbadia*, *D. tucumana* e *D. urubamba*. Outras oito espécies (*D. alexandrei*, *D. araucana*, *D. cuscungu*, *D. ecuatoriana*, *D. huilliche*, *D. limbinervis*, *D. pichinchana* e *D. quitensis*) não foram, até o momento, alocadas a nenhum subgrupo.

Todas as espécies do grupo *guarani* são Neotropicais em sua distribuição, restringindo-se a áreas da América do Sul e Central. *D. subbadia* parece ser a espécie do grupo com distribuição mais restrita, sendo encontrada apenas em áreas do México Central. *D. ornatifrons* e *D. guaru*, por outro lado, ocorrem principalmente em áreas brasileiras, embora haja registros de *D. ornatifrons* no Uruguai, Bolívia e Colômbia (Kastritsis, 1969; Bächli, 2008). Dentro do subgrupo *guaramunu*, *D. griseolineata* e *D. maculifrons* são espécies de distribuição mais ampla, em detrimento de *D. guaraja*, raramente encontrada até mesmo em território brasileiro: *D. griseolineata* é encontrada em regiões do Brasil, Bolívia, Colômbia, Equador e Venezuela; *D. maculifrons* apresenta registros no Brasil, Argentina, Bolívia e Peru (Salzano, 1955; Kastritsis, 1969; Bächli, 2008). Segundo Salzano (1955), na região Sul da América Latina, *D. maculifrons* é muito mais comum e amplamente distribuída que *D. griseolineata*, sendo que este autor encontrou uma correlação positiva entre amplitude de ocorrência e níveis de polimorfismo cromossômico na comparação entre estas duas espécies.

A partir da realização de testes de cruzamento, King (1947) estabeleceu que, dentro do subgrupo *guarani*, as espécies simpátricas *D. ornatifrons* e *D. guaru* estão

efetivamente isoladas sexualmente, enquanto *D. ornatifrons* e *D. subbadia* copulam, mas a fertilização é abortada ou os híbridos são inviabilizados em estágios precoces de desenvolvimento<sup>6</sup>. Por outro lado, segundo este autor, cruzamentos envolvendo as espécies alopátricas *D. guaru* e *D. subbadia* resultam em híbridos macho-estéreis. Nesse último caso, alguns distúrbios de desenvolvimento foram também evidenciados nos casos em que fêmeas de *D. subbadia* eram cruzadas com machos de *D. guaru*, mas não em seu cruzamento recíproco, levando à inferência de um forte efeito materno. Ainda segundo King (1947), nenhum caso de inseminação interespecífica foi detectado nos testes envolvendo combinações de espécies pertencentes aos dois subgrupos, nem mesmo naqueles envolvendo apenas espécies do subgrupo *guaramunu*.

Kastritsis (1969), com base em comparações citogenéticas, demonstrou que espécies pertencentes aos dois subgrupos acima descritos são inteiramente diferentes em sua organização cromossômica, de modo a propor que os mesmos fossem elevados taxonomicamente à categoria de grupos. Segundo este autor, a comparação a nível cromossômico torna evidente o fato de que *D. griseolineata* e *D. guaramunu* são filogeneticamente mais próximos a membros do grupo *tripunctata*, do que aos demais membros do grupo *guarani*. Neste sentido, o próprio King (1947), bem como Throckmorton (1962), destacaram o fato de que os dois subgrupos são bastante diferentes.

Mais recentemente, Robe *et al.* (2002) atestaram, molecular e morfológicamente, a monofilia do subgrupo *guarani*, mas não conseguiram eficientemente comprová-la para o subgrupo *guaramunu*, embora seus representantes tenham apresentado altos índices de similaridade entre si. A inclusão de membros de ambos subgrupos na análise mais inclusiva de Yotoko *et al.* (2003) suportou, entretanto,

---

<sup>6</sup> Kastritsis (1969) relata a obtenção de híbridos nos cruzamentos entre *D. ornatifrons* e *D. subbadia*, mas não entra em maiores detalhes quanto a esta questão.

inferências quanto à sua diferenciação. Neste caso, representantes do subgrupo *guaramunu* agruparam-se com *D. calloptera* e *D. shildi* (grupo *calloptera*), de modo a compor a linhagem irmã de *D. paramediostriata* (grupo *tripunctata*); enquanto *D. ornatifrons* compôs, juntamente com *D. mediopunctata* (grupo *tripunctata*), um clado independente do anterior. Ainda existem, entretanto, controvérsias quanto à efetivação ou não desta subdivisão. Embora não haja dúvidas referentes à diferenciação entre os subgrupos, o fato de que muitas das espécies do grupo ainda não foram devidamente alocadas a nenhum subgrupo certamente dificulta o processo (Val *et al.*, 1981).

O grupo *guarani* encontra-se situado junto à radiação *immigrans-tripunctata* de Throckmorton (1975), resultado este recentemente suportado pelos achados de Remsen & O'Grady (2002), Yotoko *et al.* (2003) e DaLage *et al.* (2007). Neste último caso, em particular, o grupo *guarani* posicionou-se dentro da radiação *tripunctata* monofileticamente definida conforme Throckmorton (1975), dado que nos dois primeiros trabalhos esta havia sido apenas inclusivamente recuperada.

#### 1.2.2.5 – Grupo *histrío*

Trata-se de um grupo composto por 16 espécies (Bächli, 2008), que distribui-se ao longo da Eurásia. Da Lage *et al.* (2007) incluem *D. sternopleuralis* em sua amostragem, porém recuperam um agrupamento inesperado para esta espécie, que constitui a ramificação basal do clado formado pelos Drosofilídeos hawaianos na árvore de máxima parcimônia, mas o grupo irmão da linhagem estabelecida por *Hirtodrosophila* mais *Mycodrosophila* na análise bayesiana. Em Perlman *et al.* (2003), *D. histrío* aparece como espécie irmã de *D. macroptera*.

#### 1.2.2.6 – Grupo *immigrans*

O grupo *immigrans* encontra-se constituído, atualmente, por 101 espécies típicas do Velho Mundo (Bächli, 2008), subdivididas em 5 subgrupos: *curviceps* (10 espécies), *hypocausta* (9 espécies), *immigrans* (34 espécies), *nasuta* (12 espécies) e *quadrilineata* (22 espécies). Segundo DaLage *et al.* (2007) este grupo constitui um agrupamento monofilético de espécies, dada a exclusão de *D. trilingata*, não alocada a nenhum dos subgrupos. Katoh *et al.* (2007), por outro lado, recuperam a monofilia de um grupo *immigrans* restritamente definido, mediante a exclusão de duas espécies do subgrupo *quadrilineata*, *D. annulipes* e *D. quadrilineata*. Quanto à filogenia dos subgrupos, DaLage *et al.* (2007) recuperam a monofilia do subgrupo *nasuta*, enquanto que os subgrupos *immigrans* e *hypocausta* misturam-se em sua composição. Katoh *et al.* (2007) também apresentam clados bem definidos para as quatro espécies do subgrupo *nasuta*, bem como para *D. siamana* + *D. hypocausta* (subgrupo *hypocausta*) e para *D. immigrans* + *D. formosana* (subgrupo *immigrans*). Entretanto, as posições de *D. ruberrima* (subgrupo *immigrans*) e *D. neohypocausta* (subgrupo *hypocausta*) oscilaram mediante utilização de diferentes métodos de análise, de forma que a monofilia dos subgrupos *immigrans* e *hypocausta* não foi resolvida. Espécies do subgrupo *quadrilineata*, por outro lado, ao mesmo tempo em que foram reveladas como remotamente relacionadas às demais espécies do grupo *immigrans*, apresentaram uma composição inteiramente polifilética na análise de Katoh *et al.* (2007). Segundo estes autores, diferenças morfológicas e citogenéticas observadas anteriormente para as espécies do subgrupo *quadrilineata* em relação às demais espécies do grupo *immigrans* apóiam seus resultados.

O grupo *immigrans* é um dos grupos mais representativos de *Drosophila* no Leste Asiático, apresentando uma distribuição confinada à região que se estende da

Costa da África, ao longo da Ásia, até as Ilhas do Pacífico (Katoh *et al.*, 2007). *D. immigrans* é a única espécie do grupo que desvia deste padrão ao atingir uma distribuição cosmopolita. Markow & O’Grady (2006) enfatizam que este grupo de espécies diversificou-se no Velho Mundo em paralelo à diversificação do grupo *tripunctata* no Novo Mundo.

Quanto ao relacionamento do grupo *immigrans* frente aos demais grupos de espécies, alguns autores o situam como linhagem basal da radiação *immigrans-tripunctata* (Throckmorton, 1975; Tatarenkov *et al.*, 2001; Remsen & O’Grady, 2002; Yotoko *et al.*, 2003), enquanto outros o localizam junto à base do subgênero *Drosophila*, de forma que sua ramificação precede a divergência dos clados *immigrans-tripunctata* e *virilis-repleta* (Katoh *et al.*, 2000; DaLage *et al.*, 2007). Katoh *et al.* (2007) apresenta os grupos *funebria* e *guttifera* como constituintes da linhagem irmã do grupo *immigrans* restritamente definido, de modo a formar um clado bem suportado dentro da radiação *immigrans-Hirtodrosophila*.

#### 1.2.2.7 – Grupo *macroptera*

Cinco diferentes espécies compõem este grupo que distribuiu-se em regiões do México e da América Central: *D. alafumosa*, *D. aurea*, *D. macroptera*, *D. magnabadia* e *D. submacroptera*. De acordo com Throckmorton (1975), este grupo participa da radiação *tripunctata*. Embora Remsen & O’Grady (2002) não tenham recuperado uma radiação *tripunctata* monofilética, eles confirmaram o posicionamento do grupo *macroptera* junto à radiação *immigrans-tripunctata*. Perlman *et al.* (2003), por outro lado, recuperaram com um suporte razoável o agrupamento entre *D. macroptera* e *D. histrio*, embora tenham apresentado uma radiação *tripunctata* polifilética. Eles destacam, entretanto, que nenhuma combinação envolvendo os grupos *tripunctata*,

*cardini*, *quinaria*, *testacea*, *histrion* e *macroptera* é significativamente diferente das demais.

#### 1.2.2.8 - Grupo *pallidipennis*

Ainda mais reduzido, o grupo *pallidipennis* é composto por uma única espécie, *D. pallidipennis* (Bächli, 2008). Segundo Val *et al.* (1981), esta espécie encontra-se subdividida em duas subespécies: *D. pallidipennis centralis*, encontrada no México e El Salvador, e *D. pallidipennis pallidipennis*, do Brasil e Peru.

Este é outro dos grupos situados dentro da radiação *tripunctata* por Throckmorton (1975), hipótese esta confirmada nos trabalhos de Yotoko *et al.* (2003) e DaLage *et al.* (2007). No primeiro destes trabalhos *D. pallidipennis* foi agrupada com *D. bandeirantorum*, membro do grupo *tripunctata*. No segundo, entretanto, seu posicionamento revelou-se instável, dependendo do método de análise filogenética: na análise de máxima parcimônia *D. pallidipennis* aparece como espécie irmã do clado que agrupa representantes dos grupos *tripunctata* e *cardini*, enquanto na análise bayesiana *D. pallidipennis* ocupa um posicionamento mais derivado, situando-se junto a dois membros do subgrupo IV do grupo *tripunctata* (*D. metzii* e *D. albirostris*).

#### 1.2.2.9 – Grupo *quinaria*

O grupo *quinaria* compreende cerca de 33 espécies Holárticas, que constituem um agrupamento monofilético segundo Yotoko *et al.* (2003) e DaLage *et al.* (2007). Segundo Spicer & Jaenike (1996), *D. phalerata* e *D. falleni* são espécies irmãs entre si, constituintes da linhagem irmã do clado que apresenta *D. guttifera* como ramificação basal em relação ao cluster que contém *D. palustris*, *D. subpalustris*, *D. recens* e *D.*

*quinaria*. Relacionamentos semelhantes foram ainda recuperados por Perlman *et al.* (2003).

Segundo Pringent *et al.* (2003), o grupo *quinaria* inclui espécies com diferentes hábitos alimentares, embora a maioria das espécies sejam micófagas. Espécies filogeneticamente mais derivadas, tais como *D. palustris* e *D. subpalustris*, entretanto modificaram seus hábitos em direção à fitofagia, alimentando-se principalmente de folhas (Perlman *et al.*, 2003).

Embora amplas amostragens envolvendo o grupo *quinaria* não sejam freqüentes, o mesmo é agrupado com os grupos *funebri* e *testaceae* no trabalho de DaLage *et al.* (2007), enquanto Yotoko *et al.* (2003) o situa como uma ramificação basal da radiação *immigrans-tripunctata*, de modo que o mesmo segue o grupo *immigrans* na ordem de divergências. Remsen & O'Grady (2002), por outro lado, situam o grupo *quinaria* junto aos grupos *bizonata*, *macroptera* e *rubrifrons*.

#### 1.2.2.10 – Grupo *rubrifrons*

O grupo *rubrifrons* contém um total de nove espécies, subdivididas em três subgrupos (Bächli, 2008): subgrupo *macroptera*, que contém *D. rubidifrons*; subgrupo *rubrifrons*, que contém *D. nubiluna*, *D. rubrifrons* e *D. spadicifrons*; e subgrupo *uninubes*, composto por *D. batmani*, *D. parachrogaster*, *D. popayan*, *D. uninubes* e *D. xalapa*. Estas espécies são encontradas, principalmente, em regiões da América do Norte e Central.

Posicionado junto a radiação *tripunctata* por Throckmorton (1975), este grupo foi também incluído na amostragem de Remsen & O'Grady (2002), onde agrupou-se com os grupos *macroptera*, *bizonata* e *quinaria* de modo a compor parte da radiação *immigrans-tripunctata*.

#### 1.2.2.11 – Grupo *sticta*

O grupo *sticta* é composto única e exclusivamente por *D. sticta*, espécie encontrada principalmente na América Central e em algumas em áreas da América do Sul (Colômbia e Brasil) (Bächli, 2008). Throckmorton (1975) posicionou esta espécie junto a radiação *tripunctata*.

#### 1.2.2.12 – Grupo *testacea*

Este pequeno grupo de quatro espécies micófagas (Bächli, 2008) é encontrado, principalmente, em áreas da América do Norte, mas também na Europa e Japão, apresentando uma distribuição Holártica. De acordo com Grimaldi *et al.* (1992), este grupo pode ser subdividido em dois subgrupos: o subgrupo *testacea*, que contém *D. testacea*, *D. orientalis* e *D. neotestacea*; e o subgrupo monoespecífico *putrida*, composto por *D. putrida*.

O grupo *testacea* foi situado por Throckmorton (1975), junto aos grupos *immigrans* e *quinaria*, na base da radiação *immigrans-tripunctata*. Nos trabalhos de Tatarenkov *et al.* (2001), Remsen & O'Grady (2002) e Yotoko *et al.* (2003), este grupo foi confirmado como representante da radiação *immigrans-tripunctata*. Entretanto, estes trabalhos sugerem a ocupação de um posicionamento mais derivado pelos membros do grupo. DaLage *et al.* (2007), entretanto, confirmam sua caracterização como grupo externo à radiação *tripunctata*, definindo-o como táxon irmão do grupo *funnebris*.

#### 1.2.2.13 - Grupo *tripunctata*

O grupo *tripunctata* foi proposto, originalmente, por Sturtevant (1942), como composto por *D. tripunctata* e, duvidosamente, por *D. histrio*, que foi, posteriormente, removida deste grupo para constituir o grupo *histrio* (atualmente com 16 integrantes).



Desde então inúmeras outras espécies foram adicionadas a este grupo, grande parte das quais descritas por Frota-Pessoa (1954). Atualmente o grupo *tripunctata* abrange um total de 79 espécies (Bächli, 2008), subdivididas em quatro subgrupos propostos com base em caracteres morfológicos (Frota-Pessoa, 1954): (1) subgrupo I (8 espécies): composto por *D. angustibucca*, *D. mediocris*, *D. medioobscurata*, *D. nappae*, *D. neoguaramunu*, *D. platitarsus*, *D. rostrata* e *D. setula*; (2) subgrupo II (14 espécies): composto por *D. bodemanna*, *D. cuaso*, *D. facialba*, *D. machachensis*, *D. medidelta*, *D. medioimpressa*, *D. mediopunctata*, *D. mediosignata* (revalidada por Bächli *et al.*, 2000), *D. morena*, *D. paraguayensis* (= *D. medionotata*, segundo Bächli *et al.*, 2000), *D. prosimilis*, *D. roehrae*, *D. tomasi*, *D. unipunctata*; (3) subgrupo III (25 espécies): abrange *D. addisoni*, *D. bandeirantium*, *D. bifilum*, *D. bipunctata*, *D. blumelae*, *D. carlosvilelai*, *D. converga*, *D. curvapex*, *D. divisa*, *D. fairchildi*, *D. fragilis*, *D. frotapessoai*, *D. johnstonae*, *D. mediopicta*, *D. mediopictoides*, *D. mediostriata* (= *D. crocina*, segundo Vilela & Bächli, 1990), *D. mesostigma*, *D. nigricincta*, *D. paramediostriata*, *D. paschoensis*, *D. trapeza* (= *D. mirassolensis*, segundo Vilela & Pereira, 1985), *D. triangula*, *D. triangulina*, *D. trifiloides* e *D. trifilum*; (4) subgrupo IV (20 espécies): *D. albescens*, *D. albicans*, *D. albirostris*, *D. arcosae*, *D. fontdevilai*, *D. greerae*, *D. leticiae*, *D. loewi*, *D. mediodiffusa*, *D. medioparva*, *D. mediovittata*, *D. metzii*, *D. pellewae*, *D. pruinifacies*, *D. semialba*, *D. spinatermina*, *D. tripunctata*, *D. tristriata*, *D. valenciai* e *D. whartoniae*. Doze outras espécies (*D. argenteifrons*, *D. cundinamarca*, *D. hansonii*, *D. hansonioides*, *D. ichubamba*, *D. mcclintockae*, *D. ninarumi*, *D. patacorona*, *D. pilaresae*, *D. quillu*, *D. surucucho* e *D. urco*) não foram, até o momento, alocadas a nenhum subgrupo (Bächli, 2008).

Exceto por *D. tripunctata*, que ocorre em regiões Neárticas, todas as demais espécies do grupo *tripunctata* são endêmicas da região Neotropical, onde o grupo é

superado, com relação à diversidade de espécies, apenas pelo grupo *repleta* (Vilela, 1992). Este grupo contém, entretanto, as espécies dominantes nas florestas Neotropicais chuvosas (Vilela & Val, 1985). Em outras localidades, especialmente no inverno, estas espécies costumam ser superadas apenas pelo grupo *willistoni* em termos de densidade populacional (Frota-Pessoa, 1954), mas isto nem sempre se aplica. As espécies do grupo *tripunctata* caracterizam-se, ainda, pela enorme versatilidade ecológica, ocupando uma grande diversidade de nichos, com algumas espécies fungívoras, outras encontradas em associação a flores ou frutos em processo de fermentação, bem como outros órgãos vegetais (Val *et al.*, 1981).

De acordo com Frota-Pessoa (1954), o subgrupo II constitui o agrupamento típico do grupo *tripunctata*, uma vez que os subgrupos I e III parecem ser transicionais em relações aos grupos *guarani* e *cardini*, respectivamente. Segundo este autor, também o subgrupo IV parece apresentar grandes afinidades ao grupo *cardini*, de modo que, em geral, o grupo *tripunctata* não apresenta caracteres que sejam comuns e ao mesmo tempo exclusivos a todas as suas espécies, o que torna a sua composição extremamente subjetiva. Este fato, aliado à grande variação intra-específica apresentada pelas espécies do grupo *tripunctata*, torna a identificação de suas espécies bastante problemática. Na maior parte dos casos torna-se necessário fazer uso da análise de genitália masculina (Vilela & Pereira, 1985; Vilela & Val, 1985; Vilela & Pereira, 1986; Vilela & Bächli, 1990; Vilela, 1992; Bächli *et al.*, 2000; Vilela & Bächli, 2000).

Patterson (1957) realizaram testes de cruzamento envolvendo onze espécies do grupo, e demonstraram que estas são, em sua maior parte, isoladas reprodutivamente umas das outras, embora algumas exceções tenham sido encontradas, como nos cruzamentos envolvendo *D. mediopunctata* e *D. unipunctata*, bem como *D. mediodstriata* e *D. paramediodstriata*. Com base na análise de padrões morfológicos

presentes na genitália masculina de algumas espécies do grupo *tripunctata*, Vilela & Pereira (1985), Vilela & Val (1985), Vilela & Pereira (1986), Vilela (1992), Bächli *et al.* (2000), Vilela & Bächli (2000) também apresentam, para muitas das espécies descritas ou redescritas, hipóteses de relacionamentos evolutivos para com outras espécies de seu subgrupo, entre as quais pode-se destacar as sugestões de proximidade entre *D. mediopictoides* e *D. mediopicta*, entre *D. bandeirantum* e *D. triangula*, entre *D. paramediostriata* e *D. mediostriata*, bem como entre a tríade de espécies crípticas *D. paraguayensis*, *D. mediosignata* e *D. cuaso*, por exemplo. Outro trabalho que buscou avaliar algumas relações interespecíficas no grupo *tripunctata* foi executado por Franc & Napp (1992), que, a partir da análise eletroforética de nove sistemas isoenzimáticos, apresentaram um dendrograma de UPGMA bastante discordante da classificação de Frota-Pessoa (1954). As espécies do grupo também apresentaram, neste trabalho, um baixo coeficiente médio de similaridade (15%), remetendo à idéia de que, possivelmente, este táxon constitua um agrupamento heterogêneo de espécies.

Em termos morfológicos, espécies do grupo *tripunctata* assemelham-se enormemente aos grupos *cardini*, *guarani* e *calloptera*, sendo que algumas de suas espécies são apontadas como transicionais em relação aos dois primeiros destes grupos (Frota-Pessoa, 1954). Throckmorton (1975) incluiu o grupo *tripunctata* em sua radiação *tripunctata*, juntamente aos grupos *rubrifrons*, *pallidipennis*, *sticta*, *macroptera*, *cardini*, *guarani* e *calloptera*. Segundo este autor, a radiação *tripunctata* é originária dos trópicos do Novo Mundo e apresenta um posicionamento derivado com relação aos demais membros da radiação *immigrans-Hirtodrosophila*, composta, entre outros, pelos grupos *testacea*, *quinaria* e *immigrans*. Ainda segundo este autor, o grupo *tripunctata* constitui a base a partir da qual, em diferentes períodos e níveis, outros grupos surgiram,

sendo que em seu trabalho de 1962, o mesmo afirma que este grupo possivelmente forneça uma forte exceção a hipóteses de monofilia.

Análises moleculares posteriores têm posicionado o grupo *tripunctata* junto a outros representantes da radiação *immigrans-tripunctata*, monofileticamente definida (Pélandakis & Solignac, 1993; Tatarenkov *et al.*, 2001; Remsen & O'Grady, 2002; Yotoko *et al.*, 2003). Tatarenkov *et al.* (2001) apresentam *D. tripunctata* como espécie-irmã de *D. putrida* (grupo *testacea*), de modo a constituir o clado-irmão do agrupamento que inclui *D. phalerata* (grupo *quinaria*) e *D. funebris* (grupo *funebris*). Também no trabalho de Remsen & O'Grady (2002) o grupo *tripunctata* agrupou-se ao grupo *testacea*, mas neste caso para constituir a ramificação basal do clado que inclui representantes dos grupos *funebris*, *pinicola*, *guarani*, *calloptera*, *cardini*, *bizonata*, *macroptera*, *rubrifrons* e *quinaria*, levando a parafilia da radiação *tripunctata*.

A inclusão de mais de uma espécie do grupo *tripunctata*, nos trabalhos de Yotoko *et al.* (2003) e DaLage *et al.* (2007) tem, por outro lado, levantado novos questionamentos. No primeiro destes trabalhos o grupo *tripunctata* foi parafileticamente constituído com relação aos grupos *cardini*, *guarani*, *pallidipennis*, *calloptera*, *funebris* e *testacea*, novamente recuperando uma radiação *tripunctata* parafilética. DaLage *et al.* (2007), por outro lado, recuperam a monofilia da radiação *tripunctata*, embora o grupo *tripunctata per se* novamente apresente uma composição parafilética com relação apenas ao grupo *cardini*, na análise de máxima parcimônia, mas com relação aos grupos *cardini*, *guarani* e *pallidipennis* na análise bayesiana. No trabalho de Yotoko *et al.* (2003), particularmente, nenhum dos agrupamentos propostos por Frota-Pessoa (1954) foram recuperados.

### 1.2.3 – Drosofilídeos hawaianos

O grupo dos drosofilídeos hawaianos fornece um dos exemplos mais notáveis de radiação explosiva, sendo composto por aproximadamente 700 espécies, agrupadas em duas linhagens principais: drosofilóides e scaptomizóides (Throckmorton, 1975). Estimativas atuais do número de espécies do grupo, entretanto, apontam a presença de cerca de 900 espécies (Markow & O’Grady, 2006). Segundo Throckmorton (1975) as linhagens drosofilóide e scaptomizóide parecem ter evoluído a partir de uma única população fundadora ancestral residente no sudeste asiático. Grimaldi (1990), entretanto, defende a hipótese de que invasões independentes originaram a atual diversidade de drosofilídeos hawaianos, de forma que drosofilóides não constituem o grupo-irmão dos scaptomizóides.

A distribuição atual destas espécies restringe-se a áreas de florestas chuvosas encontradas em seis ilhas hawaianas de formação recente: Kauai (5,1 - 4,7 milhões de anos), Oahu (3,5 - 2,6 milhões de anos), Molokai (2 - 1,8 milhões de anos), Maui (1,2 - 1,5 milhões de anos), Lanai (1,5 milhões de anos) e Hawai (0,6 - 0,5 milhões de anos)<sup>7</sup>. Segundo Russo *et al.* (1995), 98% dos drosofilídeos hawaianos apresentam uma distribuição endêmica a uma única ilha, de forma que, eventos biogeográficos de origem das ilhas podem ser tomados como estimativas aproximadas do tempo de origem de suas espécies habitantes. Como estas ilhas apresentam uma origem relativamente recente, se comparada às estimativas de diversificação do grupo [entre 24 e 30 milhões de anos segundo Thomas & Hunt (1991); cerca de 26 milhões de anos, segundo Russo *et al.* (1995); cerca de 30 milhões de anos segundo Tamura *et al.* (2004)], especula-se que ilhas antigas, não mais existentes, ou não mais ocupadas por florestas úmidas,

---

<sup>7</sup> Estimativas do período de formação geológica destas ilhas foram realizadas pelo uso de técnicas de datação radiométrica potássio-argônio (K-Ar) (Bonacum *et al.*, 2005).

tenham fornecido as condições ideais para a proliferação inicial dos drosofilídeos havaianos. Aparentemente, enquanto algumas das linhagens basais deste distinto grupo já estavam presentes há cerca de 10 milhões de anos, a maior parte da sua diversidade originou-se nos últimos 5 milhões de anos (Bonacum *et al.*, 2005).

Estimativas do tempo de divergência entre espécies havaianas são freqüentemente utilizadas na calibração do relógio molecular para Drosophilidae (Thomas & Hunt, 1991 e 1993; Russo *et al.*, 1995; Tamura *et al.*, 2004). Neste caso, diferentes autores apóiam a utilização da datação da formação geológica de Kauai (5,1-4,7 milhões de anos), a ilha mais antiga do Arquipélago Hawaiano, como tempo de divergência entre *D. picticornis* (encontrada apenas em Kauai e representante da ramificação basal do grupo *planitibia* – Bonacum *et al.*, 2005) e outros membros do subgrupo *planitibia*.

Throckmorton (1975) agrupou os drosofilídeos havaianos, tanto drosofilídeos como scaptomizóides, como uma única linhagem interna à sua radiação *immigrans-Hirtodrosophila*. Grimaldi (1990), por outro lado, propôs que as espécies havaianas de *Drosophila* compõem um gênero à parte, *Idiomyia*, que, uma vez agrupado ao seu clado-irmão composto por gêneros micófagos como *Hirtodrosophila*, *Zygothrica* e *Paramycodrosophila*, constitui uma ramificação basal de Drosophilinae. Ainda segundo este autor, *Scaptomyza* e *Engiscaptomyza* constituem um agrupamento distinto e não proximamente relacionado a *Idiomyia*, mas sim irmão do clado composto pelos gêneros *Drosophila*, *Zaprionus*, *Liodrosophila*, *Samoaia*, entre outros. Evidências moleculares têm, entretanto, conflitado com os achados de Grimaldi, de forma que trabalhos mais recentes têm confirmado a origem monofilética dos drosofilídeos havaianos, bem como sua afinidade filogenética com o subgênero *Drosophila* (Thomas & Hunt, 1991 e 1993; DeSalle, 1992a; Powell & DeSalle, 1995). Mais precisamente, e ao contrário do que foi

defendido por Throckmorton (1975), a seção *virilis-repleta* parece constituir a linhagem irmã dos drosofilídeos hawaianos (Stalker, 1972; Russo *et al.*, 1995; Rensen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000; Rensen & O'Grady, 2002; Zilversmit *et al.*, 2002). Mais recentemente, DaLage *et al.* (2007) recuperaram *D. adamsi*<sup>8</sup> como grupo irmão dos drosofilídeos hawaianos, de forma que o clado assim constituído estabeleça a linhagem-irmã da seção *virilis-repleta*. Por outro lado, Katoh *et al.* (2007) apresenta as espécies *D. annulipes* (grupo *immigrans*) e *D. maculinotata* (grupo *funnebris*) como constituintes do clado irmão das espécies hawaianas de *Drosophila*, de modo que *Scaptomyza* e a seção *virilis-repleta* ramifiquem externamente ao cluster assim definido.

Ainda, embora Throckmorton (1975) considere *Engiscaptomyza* como intermediária entre *Scaptomyza* e *Drosophila*, uma série de outros trabalhos têm destacado a proximidade filogenética entre *Engiscaptomyza* e *Scaptomyza* (Grimaldi, 1990; DeSalle, 1992a; Thomas & Hunt, 1993; Russo *et al.*, 1995; Rensen & DeSalle, 1998). Segundo Russo *et al.* (1995) estes dois táxons divergiram entre si há cerca de 11 milhões de anos. Neste sentido, embora Grimaldi (1990) reserve um status genérico para *Engiscaptomyza*, outros o têm considerado um subgênero de *Scaptomyza* (Bächli, 2008). Também os antigos gêneros *Titanochaeta* e *Grimshawomyia* foram recentemente incorporados ao gênero *Scaptomyza* (Markow & O'Grady, 2006).

Os grupos de hawaianos *Drosophila* compõem um total de 379 espécies, ainda classificadas como *Idiomya* por Bächli (2008). Sua subdivisão se deu com base em caracteres sexuais secundários encontrados nos machos, ou em caracteres dimórficos entre machos e fêmeas, tais como padrões de pigmentação das asas (*picture wing group* – 143 espécies), alterações nos apêndices bucais (*modified mouth parts group* – 47

---

<sup>8</sup> Uma espécie até o momento não alocada a nenhum grupo (Bächli, 2007).

espécies) ou tarsais (*modified tarsus* group – 42 espécies; *spoon tarsi* group – 6 espécies; *ciliated tarsus*), etc. Estes grupos são, ainda, adicionalmente subdivididos. Espécies do grupo *picture wing*, por exemplo, são distribuídas em 13 subgrupos: *adiastola*, *anomalipes*, *distinguenda*, *grimshawi*, *hawaiiensis*, *limitada*, *odontophalus*, *orphnopeza*, *pilimana*, *planitibia*, *primaeva*, *punalua* e *vesciseta*. O grupo *modified mouth parts*, por outro lado, subdivide-se em 5 subgrupos: *eurypeza*, *fuscoamoeba*, *mimica*, *mitchelli* e *semifuscata*. Além disso, os próprios subgrupos são divididos em sub-complexos: *planitibia* subdivide-se nas seções *picticornis*, *planitibia*, *neopicta* e *cyrtoloma* (Bächli, 2008). De acordo com Throckmorton (1966), dentre os grupos hawaianos de *Drosophila*, *haleakalae* e *ciliated tarsus* compõem a ramificação basal. Neste caso, ocupando um posicionamento derivado em sua árvore, aparecem os grupos irmãos *modified tarsus* e *antopocerus*, que estabelecem a linhagem irmã do clado que apresenta os grupos *modified mouthparts*, *picture wing*, *Nudidrosophila* e *Ateledrosophila*. Mais recentemente, estes dois últimos gêneros foram incorporados ao gênero *Drosophila* (Markow & O’Grady, 2006).

O gênero *Scaptomyza*, por outro lado, é composto por 263 espécies (Bächli, 2008), algumas das quais apresentam uma distribuição restrita às Ilhas Hawaianas, enquanto outras ocorrem em áreas continentais. Segundo Markow & O’Grady (2006), cerca de 150 espécies perfazem a fração do gênero *Scaptomyza* endêmica ao Hawaii, enquanto as demais encontram-se distribuídas em nove grupos continentais. Tamura *et al.* (1995), a partir de análises do gene *Adh* em diversas espécies do gênero, concluíram que ele constitui um agrupamento monofilético de espécies. O fato de que o cluster composto por representantes continentais e insulares de *Scaptomyza* agrupe fortemente com os demais drosofilídeos hawaianos (Tatarenkov *et al.*, 1999a), é consistente com a teoria de origem hawaiana.



### 1.3 – O Subgênero *Dorsilopha*

O subgênero *Dorsilopha* apresenta um total de três espécies descritas (Bächli, 2008): *D. busckii*, *D. linearidentata* e *D. neobusckii*. Destas, *D. busckii* apresenta uma distribuição cosmopolita, enquanto as outras duas espécies encontram-se restritas a determinadas regiões da Ásia.

Throckmorton (1975) posicionou *Dorsilopha* junto a outros gêneros, tais como *Hirtodrosopila*, como pertencente à radiação *immigrans-Hirtodrosophila*. Kwiatowski *et al.* (1997) também favoreceram a visão de *Dorsilopha* e *Hirtodrosophila* como táxons irmãos, constituintes da linhagem irmã do subgênero *Drosophila*. Trabalhos posteriores, como Kwiatowski & Ayala (1999) e Tataronkov & Ayala (1999a e 2001) concordam quanto ao posicionamento de *Dorsilopha* como um dos primeiros grupos a divergir após *Sophophora*. Remsen & O’Grady (2002), apresentam *Dorsilopha* como linhagem irmã de *Zaprionus*.

## **A importância do estabelecimento das relações filogenéticas dentro de *Drosophilidae***

Ecologicamente, espécies pertencentes à família *Drosophilidae* apresentam uma importância fundamental nas cadeias saprofíticas dos mais diversos ecossistemas, uma vez que seus estágios imaturos são, na maior parte dos casos, dependentes de organismos fermentadores (Throckmorton, 1975). Alguns representantes desta família são ainda parasitas (gênero *Cladochaeta* e alguns membros do gênero *Drosophila*) e até mesmo predadores de insetos em seu estágio larval (Remsen & O'Grady, 2002), embora a grande maioria se desenvolva a partir de fungos, flores, cactos e, principalmente, frutos (Noor & Akashi, 2001).

Por outro lado, a amplitude de espécies do gênero *Drosophila*, associada à sua enorme diversidade morfológica, e sua distribuição em uma ampla variedade de nichos, o torna uma mina-de-ouro para estudos evolutivos (DaLage *et al.*, 2007). A existência de espécies crípticas, por exemplo, fornece um material valioso para estudos de especiação.

Desta forma, não apenas a sua ampla variabilidade ecológica, mas também sua diferenciação morfológica tem atraído a atenção de inúmeros pesquisadores ao longo dos anos. Tudo isto, aliado ainda às suas facilidades de manutenção e manipulação, ao seu ciclo de vida relativamente curto, e à riqueza de dados biológicos já acumulados, tem transformado estas espécies em organismos modelo ideais para uma série de estudos biológicos, principalmente no âmbito genético e evolutivo. Pode-se dizer, inclusive, que dos muitos milhões de espécies de animais e plantas existentes atualmente, nenhum organismo tem recebido maior atenção como organismo experimental quanto *Drosophila*, que tem desempenhado um papel central na

elucidação de inúmeros processos biológicos, entre os quais destacam-se aspectos genéticos, evolutivos, moleculares e ontogenéticos. Sem exagero, pode-se afirmar com convicção que o gênero *Drosophila* tem sido fundamental na origem e no subsequente desenvolvimento da genética evolutiva, inaugurando estudos de biologia evolutiva com interface em ecologia e genética e estreitando o abismo antes existente entre os estudos fenotípicos e genotípicos, como é claramente mostrado por Noor & Akashi (2001). Segundo Grimaldi (1990), ao longo dos últimos 100 anos, *Drosophila* tem sido o organismo experimental *premier* para a biologia de eucariotos.

Concomitante ao contínuo interesse em *Drosophila* como um organismo modelo, tem havido um crescente interesse na utilização da biologia comparativa para obtenção de um melhor entendimento de uma série de padrões e processos biológicos. Assim, além do interesse primário nas relações filogenéticas *per se* do gênero, pode-se dizer que a compreensão das mesmas é cientificamente crucial, uma vez que propicia: 1) a análise da história evolutiva de uma grande diversidade de características já conhecidas e estudadas dentro do grupo, ou ainda futuramente descobertas; 2) a extrapolação ou não das eventuais conclusões obtidas a outros organismos, a partir da análise do contexto filogenético em questão; 3) a definição prévia das estratégias amostrais necessárias em qualquer tipo de projeto (Powel & DeSalle, 1995). Estudos acerca das relações filogenéticas dentro do gênero *Drosophila* tornam-se ainda mais imprescindíveis com a disponibilidade das seqüências genômicas de 12 de suas espécies (Clark *et al.*, 2007), uma vez que a maior parte das análises de genômica comparativa envolve inferências sobre história evolutiva, que requerem uma descrição acurada e confiável das relações entre as espécies em consideração.

A sistemática de Drosophilidae tem sido objeto de interesse de morfologistas desde Sturtevant (1921, 1939 e 1942), pesquisador que, segundo Grimaldi (1990),

merece o título de fundador da “drosofilologia”. Foi Sturtevant (1939 e 1942) que desenvolveu, por exemplo, a classificação subgenérica do gênero *Drosophila*, ainda usada atualmente. Neste período, entretanto, os interesses eram, acima de tudo, taxonômicos, sendo que foi apenas mais recentemente que um enfoque filogenético foi adicionado a este cenário. Com a mudança de perspectiva em curso desde então, inúmeros trabalhos buscando auxiliar na resolução das relações filogenéticas entre membros da família Drosophilidae tem sido desenvolvidos, conforme pôde ser visto anteriormente. Entretanto, os mesmos ainda não tiveram um profundo impacto na taxonomia do grupo, que permanece amplamente controversa. Entre as maiores contribuições à filogenia de Drosophilidae, pode-se, seguramente, apontar os trabalhos de Throckmorton (1962 e 1975) e de Grimaldi (1990).

## **1- Estudos filogenéticos com base em caracteres morfológicos:**

### **1.1 - Throckmorton (1962 e 1975)**

Throckmorton (1962 e 1975), com base na análise de caracteres morfológicos, principalmente caracteres reprodutivos internos, descreveu parte da filogenia, ecologia e geografia do gênero *Drosophila* dentro da família Drosophilidae. Com base em uma metodologia que assumia a possibilidade de uma evolução mais dinâmica, em que polimorfismo ancestral e integração independente eram admitidos como explicações para a distribuição irregular de características, Throckmorton propôs um cenário evolutivo complexo para a família Drosophilidae. Segundo ele, a regra-geral na evolução desta família é a noção de que grupos altamente especiosos resultam da fragmentação de populações amplamente polimórficas, ou seja, o potencial adaptativo de uma população estando diretamente relacionado à sua heterozigosidade. Neste caso,

segundo ele, grupos especializados e monomórficos, embora sejam produzidos ao longo da evolução de Drosophilidae, tendem a não deixar representantes.

Em seu trabalho de 1975, Throckmorton defende que a evolução da família Drosophilidae teria se dado através de sucessivas radiações, onde se verifica o surgimento de muitas linhagens a partir de um mesmo nível filogenético ancestral. No total, teriam sido cinco radiações, todas iniciadas antes do final do Eoceno e altamente influenciadas pela disjunção tropical, ocorrida no início do Oligoceno, e pela disjunção das florestas temperadas, ocorrida no Mioceno Médio. Os detalhes finais dos padrões atuais teriam sido gerados principalmente durante o Plioceno (com a formação do Istmo do Panamá, com o desenvolvimento de uma nova conexão África-Ásia, e com a formação de ilhas havaianas, por exemplo), e Pleistoceno (com suas glaciações e refúgios), embora efeitos recentes também possam ser, por vezes, visualizados.

Throckmorton (1962 e 1975) apresenta uma série de outros gêneros surgindo a partir do gênero *Drosophila*, que ocupa, portanto, uma posição basal dentro da filogenia da subfamília Drosophilinae, e um caráter parafilético na taxonomia do grupo. Ele ainda demonstra que a parafilia não ocorre somente com o gênero *Drosophila*, sendo amplamente difundida dentro da taxonomia de Drosophilidae. Para melhor visualizar seus resultados, é importante inspecionar as principais radiações que Throckmorton propõe para a família Drosophilidae, que são (Fig. 1):

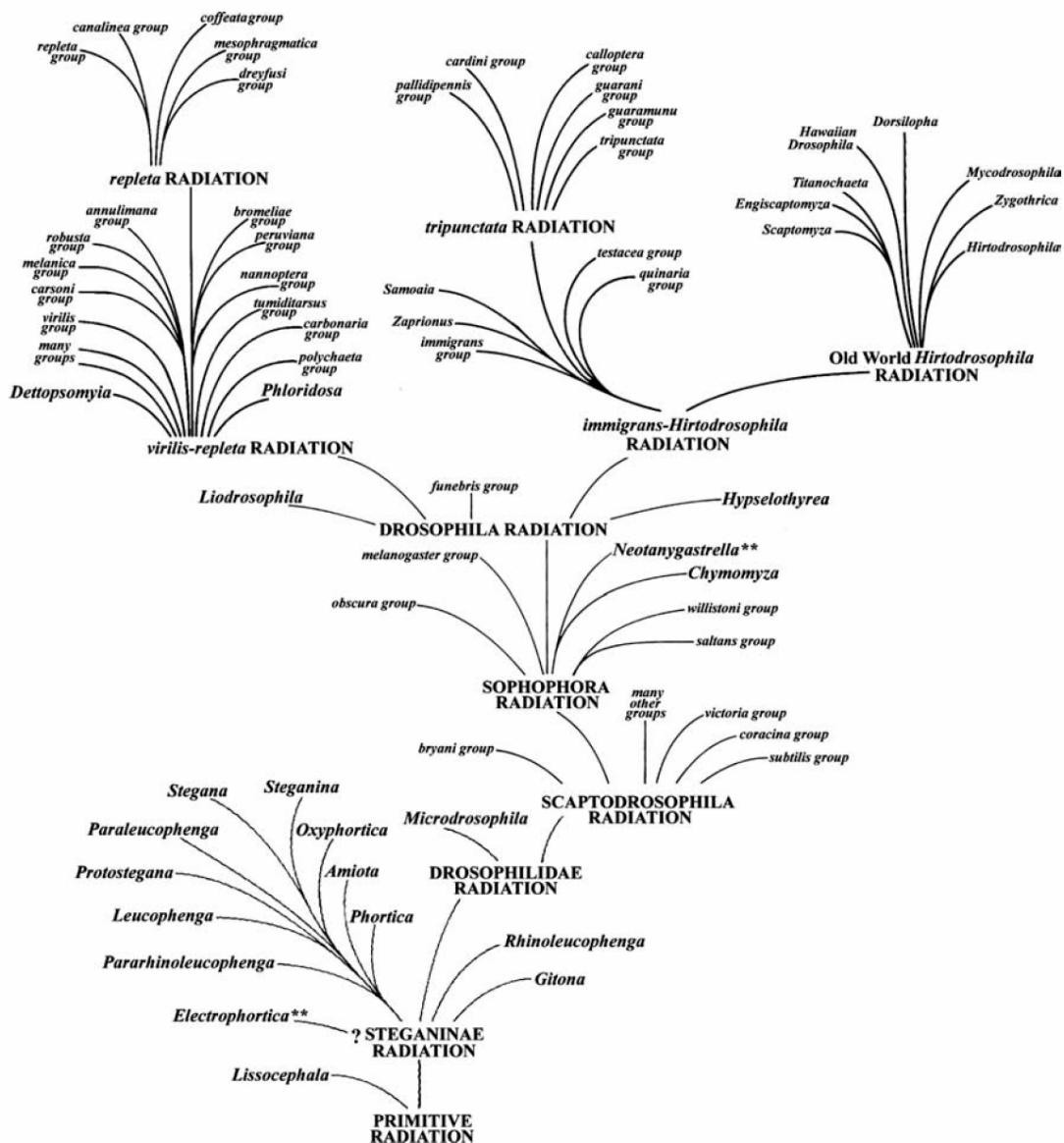


Figura 1 – Esquema das relações filogenéticas propostas por Throckmorton (1975) para a família Drosophilidae. (Extraído de Markow & O’Grady, 2006)

### 1 – Radiação Steganinae

Esta radiação abrange, entre outros, o gênero pantropical *Lissocephala*, considerado por Throckmorton (1975) como o gênero mais antigo de Drosophilidae, e o gênero fóssil *Electrophortica*, descrito por Hennig (1965) e datado para o período do Eoceno. Além destes, fazem ainda parte da radiação Steganinae os gêneros *Leucophenga*, *Rhinoleucophenga*, *Pararhinoleucophenga*, *Phortica*, *Oxyphortica*,

*Amiota*, *Gitona*, *Stegana*, entre outros. Grande parte destes gêneros tem sua distribuição principal nos trópicos do Velho e/ou do Novo mundo, embora alguns de seus representantes sejam cosmopolitas (*Amiota* e *Stegana*).

## 2 – Radiação *Drosophilinae*

Dentre seus representantes, espécies do gênero *Microdrosophila* apresentam uma distribuição pantropical, enquanto que os membros da radiação *Scaptodrosophila* são primariamente tropicais.

## 3 – Radiação *Scaptodrosophila*

Provavelmente originária da Ásia tropical, esta radiação é constituída pelos grupos *latifasciaeformis*, *bryani*, *coracina*, *victoria* e *subtilis* do então subgênero *Scaptodrosophila*, além de abranger outros gêneros, como por exemplo o gênero *Zaprionotricha*, representante da radiação no Novo Mundo. Em contraste aos outros representantes desta radiação, os grupos *victoria* e *subtilis* apresentam uma distribuição primária temperada disjunta, considerada derivada em relação às demais, e resultante de um pronunciado efeito fundador. Segundo Throckmorton (1975) este é o primeiro grupo capaz de explorar consideravelmente frutas em fermentação.

## 4 – Radiação *Sophophora*

Esta radiação compreende a linhagem *saltans-willistoni*, do Novo Mundo, o grupo *melanogaster*, presente principalmente nos trópicos do Velho Mundo, e o grupo *obscura*, mais homogêneo e distribuído ao longo da zona temperada Holártica. Além disso, outros representantes do subgênero *Drosophila*, como os grupos *populi* e *fima*, assim como outros dois gêneros proximamente relacionados, *Chymomyza*, e *Neotanygastrella* (cosmopolita e pantropical, respectivamente), fazem parte desta radiação.

Como na radiação anterior, as linhagens do Novo Mundo são consideradas derivadas com relação às formas do Velho Mundo, de forma que a radiação *Sophophora* é referida como originária dos trópicos do Velho Mundo, a partir de ancestrais primitivos pertencentes à radiação *Scaptodrosophila*, sendo *melanogaster* e *obscura* as linhagens-irmãs que primeiro se diversificaram.

## 5 – Radiação *Drosophila*

Esta radiação é composta de outras duas sub-radiações, além do grupo *funbris*, de distribuição primariamente temperada, e de outros dois gêneros distintos, restritos ao Velho Mundo, *Liodrosophila* e *Hypselothyrea*.

### 5.1 – Radiação *virilis-repleta*

Segundo Throckmorton (1975), esta radiação iniciou-se nos trópicos do Velho Mundo, expandindo-se, eventualmente, para o Novo Mundo, de forma que o grupo *repleta* acaba por tornar-se endêmico das Américas. Ainda a partir de migrações desta linhagem inicial, e como resultado de efeito fundador, parecem ter se originado novos táxons de distribuição temperada.

A radiação tropical inicial é representada por um conjunto de espécies bastante heterogêneas entre si, entre as quais se destacam representantes dos grupos *tumiditarsus*, *polychaeta*, *annulimana* e *carbonaria*, do gênero *Dettopsomyia* e do subgênero *Phloridosa*. As linhagens temperadas derivadas são constituídas pelos grupos *virilis*, *melanica*, *carsoni* e *robusta*. Outro conjunto de espécies proximamente relacionadas, com origem similar, características intermediárias, e distribuição Neotropical, é formado pelos grupos *nannoptera*, *peruviana*, *bromeliae* e talvez pelo grupo *flavopilosa*, bem como pelo subgênero *Sordophila*. O restante da radiação *virilis-repleta* é composta pelo que se convencionou chamar de radiação *repleta*, centrada nos



trópicos do Novo Mundo, e composta pelo grupo fundamentalmente andino *mesophragmatica*, pelo grupo *castanea* (atual *coffeata*), pela linhagem constituída pelos grupos *dreyfusi* e *canalinae*, e pelo grande grupo *repleta*. Todas as linhagens originadas no Novo Mundo experimentaram ampla diversificação de formas e habitats, com algumas invadindo regiões de altas altitudes (grupo *mesophragmatica*), outras se expandindo até regiões áridas e desertas (grupos *nannoptera* e *repleta*), enquanto grande parte especializa-se na exploração de florestas (grupos *canalinae*, *dreyfusi*, *peruviana*, *bromeliae* e *castanea*).

## 5.2 – Radiação *immigrans-Hirtodrosophila*

Esta radiação, assim como a radiação *virilis-repleta*, originou-se e diversificou-se inicialmente nos trópicos do Velho Mundo, para então expandir-se para a região Neotropical na forma de duas diferentes sub-radiações: a radiação *tripunctata* e a radiação *Hirtodrosophila*. Então, principalmente a partir destes dois centros, fundadores isolados levaram à invasão da zona temperada.

Representantes dos grupos *immigrans* (originário do Velho Mundo), *testaceae* e *quinaria* (predominantemente de regiões temperadas), do subgênero *Siphlodora* (Neotropical), bem como dos gêneros *Zaprionus* (tipicamente africano) e *Samoaia* (endêmicos de Samoa) ocupam a base da radiação *immigrans-Hirtodrosophila* de Throckmorton (1975). Segundo ele, a radiação *tripunctata*, centrada em regiões Neotropicais, teria se originado principalmente a partir da “subdivisão explosiva” de uma população ancestral grande e polimórfica, constituída pelos ancestrais do próprio grupo *tripunctata*, que em diferentes épocas e níveis, teria derivado os outros grupos: *macroptera*, *sticta*, *rubrifrons*, *calloptera*, *pallidipennis*, *cardini*, *guarani* e *guaramunu*. Já a radiação *Hirtodrosophila* é composta por membros do subgênero heterogêneo e disperso *Hirtodrosophila*, como, por exemplo, pelo grupo temperado *melanderi*; por

membros do subgênero cosmopolita *Dorsilopha*; pelo gênero Neotropical *Zygothrica*; além dos gêneros *Nesiodrosophila*, *Paramycodrosophila*, *Mycodrosophila*, *Scaptomyza* e dos drosofilídeos hawaianos.

Segundo Throckmorton (1962), as radiações *Sophophora* e *Drosophila* apresentaram padrões evolutivos contrastantes ao longo de sua história. Neste caso, embora ambas tenham evoluído a partir de populações ancestrais altamente polimórficas, a evolução do subgênero *Sophophora* teria se dado, principalmente, a partir da restrição do polimorfismo ancestral, seguido da integração de novas formas surgidas por mutação e recombinação. A evolução do subgênero *Drosophila*, por outro lado, teria ocorrido com a manutenção, e até mesmo a ampliação dos níveis de polimorfismo, o que possibilitou sua expansão e a exploração de novos e diferentes nichos, até mesmo em regiões inexploradas. Este padrão está em inteira concordância com os níveis de versatilidade ecológica de cada um dos subgêneros, bem como com as dificuldades no estabelecimento das relações filogenéticas entre suas espécies e grupos formadores, com o subgênero *Drosophila* apresentando filogenias com altos níveis de politomia e baixos índices de suporte.

Entretanto, os achados de Throckmorton são, muitas vezes, questionados, devido às inconsistências e contradições entre alguns de seus resultados prévios e aqueles apresentados no seu paper de 1975 (Grimaldi, 1990), bem como devido à ausência de métodos filogenéticos explícitos e medidas de suporte para cada relação (Remsen & O'Grady, 2002). A partir da execução de uma análise cladística dos caracteres recodificados de Throckmorton, Grimaldi (1990) também demonstrou que a árvore obtida é bastante diferente da apresentada pelo autor, demonstrando que, de modo geral,

seus resultados são pouco reproduzíveis<sup>9</sup>. Outro ponto relevante é que Throckmorton mantém, convenientemente, uma série de táxons parafiléticos e, apesar de suas análises filogenéticas, não propõe nenhuma nova classificação para Drosophilidae, de forma que, ainda hoje, a taxonomia do grupo permanece inconsistente com suas relações filogenéticas. Além disso, Powell e DeSalle (1995) argumentam que as árvores de Throckmorton são extremamente difíceis de interpretar, uma vez que não explicitam os padrões dicotômicos apresentados pelos cladogramas modernos. DeSalle (1992b), por outro lado, reflete acerca da veracidade do cenário proposto, amplamente constituído por politomias.

## 1.2 - Grimaldi (1990)

A partir destas críticas, Grimaldi (1990) realizou a análise cladística de 217 caracteres morfológicos adultos em mais de 120 espécies pertencentes a 50 gêneros da família Drosophilidae, com o objetivo de propor um novo sistema de classificação para o grupo, dividindo-o em uma hierarquia de táxons monofiléticos, tais como tribos, subtribos, infratribos e grupos de gêneros, importantes para a manipulação taxonômica de diversos aspectos de sua diversidade. Neste trabalho, Grimaldi diagnostica a monofilia da família Drosophilidae, e mantém sua tradicional subdivisão em duas subfamílias.

Entretanto, com base nos resultados obtidos, este autor defende que o gênero *Drosophila* é, na verdade, derivado e monofilético, uma vez que seus subgêneros *Hirtodrosophila*, *Scaptodrosophila*, *Antopocerus*, *Engiscaptomyza* e *Lordiphosa* sejam considerados como gêneros à parte. Neste caso, o clado constituído pela união do grupo genérico *Zaprionus* (constituído por *Zaprionus*, *Samoia* e *Phorticella*), com o grupo

---

<sup>9</sup> Há que se levar em conta, neste ponto, o fato de que os métodos de análise foram essencialmente diferentes.

genérico *Styloptera* (composto por *Liodrosophila*, *Sphaerogastrella*, *Dettopsomyia*, além de outros gêneros), representaria o grupo-irmão do gênero *Drosophila*, que apresenta *Dorsilopha* como ramificação basal e *Sophophora* e *Drosophila* como clados-irmãos.

Ainda segundo Grimaldi (1990), os drosofilídeos havaianos, previamente agrupados junto ao subgênero *Drosophila*, não pertencem ao gênero *Drosophila*. Mais do que isso, Grimaldi defende que espécies havaianas de *Drosophila* e *Scaptomyza* não constituem um grupo monofilético. Desta forma, ele translocou as espécies havaianas de *Drosophila*, juntamente com os gêneros *Ateledrosophila* e *Nudidrosophila*, para um gênero à parte, *Idiomyia*, que constitui o clado-irmão do grupo genérico *Zygothrica*, formado pelos gêneros micófagos *Hirtodrosophila*, *Zygothrica*, *Mycodrosophila*, *Paramycodrosophila* e *Paraliodrosophila*. Segundo este autor, *Scaptomyza* constitui um agrupamento monofilético, irmão dos gêneros *Lordiphosa* e *Engiscaptomyza*, sendo que o clado assim estabelecido acaba por formar o grupo-irmão do clado constituído pelos gêneros *Drosophila*, junto aos grupos genéricos *Zaprionus* e *Styloptera*.

Embora Grimaldi (1990) tenha realizado um amplo estudo, notável pela sua amostragem e grande número de caracteres analisados, seus achados são também constantemente criticados, principalmente, devido às inúmeras incongruências apresentadas com relação, tanto a estudos anteriores, quanto a análises posteriores. Além disso, Remsem & O'Grady (2002) afirmam que a metodologia empregada por Grimaldi foi altamente tendenciosa, uma vez que, das cerca de 30.000 árvores igualmente parcimoniosas geradas por seus dados, ele considerou apenas 100 na execução das pesagens sucessivas, isto devido às restrições de memória e velocidade computacional então existentes. Além disso, Grimaldi também não emprega medidas de

suporte estatístico aos clados. Porém, há que se ressaltar que o estudo de Grimaldi alcançou uma amostragem taxonômica ainda não superada por qualquer outra tentativa.

## **2 - A sistemática molecular de Drosophilidae**

Pelo fato da família Drosophilidae ser altamente especiosa, em um grande número de casos existe uma certa insuficiência de caracteres morfológicos, etológicos e até mesmo citológicos com os quais trabalhar. Neste caso, dados moleculares têm-se mostrado extremamente valiosos, uma vez que adicionam inúmeros novos caracteres filogeneticamente informativos, capazes de revelar relações previamente obscuras. Além disso, caracteres moleculares apresentam um padrão evolutivo mais simples, bem como uma enorme diversidade de marcadores aplicáveis ao estudo dos mais diferentes graus de divergência taxonômica. DeSalle & Grimaldi (1993) demonstram que amplos níveis de homoplasia são revelados tanto por caracteres morfológicos quanto por caracteres moleculares, principalmente quando níveis taxonômicos menos restritos são analisados. Entretanto, enquanto caracteres moleculares permitem a utilização de marcadores conservados, a escolha de caracteres morfológicos adequados para diferentes níveis de divergência é muito mais complicada.

Embora o acréscimo fornecido pela sistemática molecular tenha auxiliado na resolução de alguns problemas filogenéticos, em geral seus marcadores foram incapazes de decifrar algumas das mais interessantes politomias/incongruências inferidas com base em dados morfológicos e etológicos. Ao contrário do que se pode imaginar, a sistemática molecular, ao invés de resolver padrões previamente confusos, trouxe novas incertezas ao apresentar resultados incongruentes com análises prévias, e até mesmo incongruentes entre si.

Além disso, a maior parte do esforço de trabalho concentrou-se na inferência das relações entre espécies proximamente relacionadas dentro do gênero *Drosophila*,

especialmente aquelas pertencentes ao subgênero *Sophophora*. Quando são consideradas relações em níveis mais restritos, como àquelas referentes a espécies de um mesmo grupo, alguns cenários são amplamente trabalhados, à despeito de outros. Um exame detalhado deste quadro permite dizer que, enquanto algumas relações filogenéticas dentro de Drosophilidae foram fortemente enfocadas por um grande conjunto de dados, outras permanecem controversas em todos os seus níveis, ou até mesmo ignoradas. Por outro lado, freqüentemente, estudos delineados para a análise das relações filogenéticas em níveis mais inclusivos, como aquelas referentes a família Drosophilidae como um todo, acabam por esbarrar em amostragens extremamente limitadas.

Desta forma, uma série de questões permanecem em aberto devido à não-abordagem ou a devido às incongruências apresentadas pelos diferentes estudos que visaram a sua resolução. Isto se deve, principalmente, às diferenças entre a filogenia do gene X filogenia das espécies associadas a eventos de transferência horizontal ou “incomplete lineage sorting”, às variações entre padrões e taxas evolutivas apresentadas por diferentes OTU’s, à escolha inadequada de marcadores saturados ou não-informativos, e até mesmo às amostragens divergentes, que impedem a construção de uma estrutura filogenética mais confiável. Na verdade, não é esperado de forma alguma que um único gene seja capaz de produzir uma filogenia inequívoca para a família, de forma que somente com o acúmulo de dados, para diferentes genes e espécies, poderemos nos aproximar do “verdadeiro” cenário evolutivo que deu origem à família Drosophilidae. Talvez este cenário seja alcançado apenas com o uso de um enfoque “filogenômico”, em que diferentes loci, evoluindo a diferentes taxas, são amostrados em conjunto, de forma a aumentar o suporte nos diferentes níveis da topologia. Zilversmit *et al.* (2002) defendem que a utilização deste tipo de enfoque é fundamental para que os

estudos de sistemática filogenética, bem como de evolução genômica, atinjam todo o seu potencial.

### **3 - Os problemas na filogenia de Drosophilidae**

#### 3.1 – Grupo(s)-externo(s) ao gênero *Drosophila*

Segundo Pélandakis *et al.* (1991), Kwiatowski *et al.* (1997), Kwiatowski & Ayala (1999), Tatarenkov *et al.* (1999a), Katoh *et al.* (2000) e DaLage *et al.* (2007), os gêneros *Scaptodrosophila* e *Chymomyza* divergiram antes dos demais Drosophilinae. Embora nestes quatro primeiros trabalhos, bem como em DaLage *et al.* (2007), a ordem correta de ramificação não pôde ser determinada, Katoh *et al.* (2000) apresenta *Scaptodrosophila* como táxon mais proximamente relacionado ao clado maior composto por representantes dos subgêneros *Drosophila* e *Sophophora*, bem como dos gêneros *Liodrosophila*, *Lordiphosa*, *Mycodrosophila*, *Hirtodrosophila* e *Scaptomyza*. Por outro lado, Throckmorton (1975) posiciona *Scaptodrosophila* como basal em relação aos demais representantes da subfamília Drosophilinae, enquanto, segundo este autor, *Chymomyza* faria parte da radiação *Sophophora*. Já Grimaldi (1990) posiciona ambos os gêneros dentro de um clado maior que constitui a base da subfamília Drosophilinae.

Os trabalhos de DeSalle (1992b), Powell & DeSalle (1995) e Russo *et al.* (1995) também apontam *Scaptodrosophila* como a linhagem mais basal. Neste caso, embora Russo *et al.* (1995) não tenha incluído representantes de *Chymomyza*, DeSalle (1992b) e Powell & DeSalle (1995) defendem que este gênero ocupa um posicionamento intermediário, divergindo apenas depois de *Hirtodrosophila*, um resultado inconsistente com uma série de outros estudos, segundo os quais *Hirtodrosophila* é proximamente relacionado a membros do subgênero *Drosophila* (Throckmorton, 1975; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000; DaLage *et al.*, 2007).

Entretanto, Yotoko *et al.* (2003), ao utilizar *Ceratitis capitata* como grupo-externo, recuperam, com baixo suporte, a ramificação basal de *Hirtodrosophila*.

Trabalhos mais recentes, por outro lado, têm ampliado a representatividade amostral da família, de modo a apresentar novas alternativas a esta questão. Neste sentido, Remsen & O'Grady (2002) apontam representantes do clado *Zygothrica* (composto por *Hirtodrosophila*, *Mycodrosophila* e *Paramycodrosophila*) juntamente com *Samoia* como o primeiro grupo de Drosophilinae a divergir, seguidos pelo clado *Formosa* (constituído por *Liodrosophila*, *Sphaerogastrella* e *Dettopsomyia*). *Scaptodrosophila* e *Chymomyza* acabam por ocupar um posicionamento mais derivado, constituindo o clado irmão do subgênero *Sophophora*. Entretanto, este resultado é incongruente com os achados de Kwiatowski *et al.* (1994), que evidenciaram a ausência de um íntron do gene *Sod* (superóxido-dismutase) em todos os drosofilídeos com exceção de *Scaptodrosophila* e *Chymomyza*.

A questão acerca de qual destes dois gêneros divergiu primeiro, entretanto, parece ter sido corretamente endereçada por Tarrío *et al.* (2001) que, ao utilizar modelos evolutivos não uniformes e não estacionários, recuperaram *Scaptodrosophila* como ramificação basal, ao que se segue a diversificação de *Chymomyza*.

### 3.2 – Posicionamento de *Zaprionus*

O gênero *Zaprionus* foi apresentado como grupo irmão do clado constituído pelos subgêneros irmãos *Sophophora* e *Drosophila* na filogenia de Grimaldi (1990). Estes resultados foram parcialmente recuperados por DeSalle (1992b), que apresenta *Zaprionus* como linhagem irmã do gênero *Drosophila* parafileticamente definido com relação ao gênero *Liodrosophila*. DeSalle & Grimaldi (1992) também favorecem este cenário. Throckmorton (1975), entretanto, defendia que *Zaprionus* ocupa um



posicionamento derivado na filogenia de *Drosophila*, fazendo parte da radiação *immigrans-Hirtodrosophila*. Posteriormente, *Zaprionus* foi apresentado, ainda, como estando entre os primeiros grupos a divergir depois de *Sophophora* (Pélandakis & Solignac, 1993; Kwiatowski *et al.*, 1997; Rensen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000), ou até mesmo como sendo o primeiro (Thomas & Hunt, 1993; Russo *et al.*, 1995; Yotoko *et al.*, 2003; DaLage *et al.*, 2007). Neste caso, a afinidade filogenética entre *Zaprionus* e os membros da radiação *immigrans-Hirtodrosophila* foi fracamente suportada nos trabalhos de Kwiatowski & Ayala (1999), Tatarenkov *et al.* (2001) e Katoh *et al.* (2007), embora valores de bootstrap maiores tenham sido encontrados na análise executada por Tatarenkov *et al.* (1999a) com base nas posições nucleotídicas 1 + 2 dos genes *Ddc*, *Sod* e *Adh*. Rensen & O'Grady (2002), entretanto, posicionam *Zaprionus* como táxon irmão do subgênero *Dorsilopha*, sendo que o clado assim constituído agrupa-se, neste trabalho, ao clado formado pelo subgênero *Sophophora* e pelos gêneros *Scaptodrosophila* e *Chymomyza*.

### 3.3 – Posicionamento de *Hirtodrosophila*

Throckmorton (1975) propôs que *Hirtodrosophila*, juntamente com o subgênero *Dorsilopha*, são proximamente relacionados entre si, de modo a compor uma ramificação da radiação *immigrans-Hirtodrosophila*. Grimaldi (1990), por outro lado, posiciona *Hirtodrosophila* externamente ao clado composto pelos gêneros *Drosophila*, *Zaprionus* e *Scaptomyza*, mas junto aos grupos hawaianos de *Drosophila* e outros gêneros micófagos, de modo a propor sua elevação taxonômica à categoria de gênero.

Com relação a resultados obtidos a partir de dados moleculares, DeSalle (1992a) situa *Hirtodrosophila* como grupo irmão do clado constituído por representantes dos subgêneros *Sophophora* e *Drosophila* (incluindo drosofilídeos hawaianos), enquanto

DeSalle (1992b), mais precisamente, situa *Hirtodrosophila* como a primeira linhagem a divergir depois de *Scaptodrosophila*, definindo-a como basal com relação aos gêneros *Chymomyza*, *Zaprionus* e *Drosophila*. Utilizando parte dos marcadores utilizados por DeSalle (1992a e b), Remsen & DeSalle (1998) apresentam *Hirtodrosophila* como grupo-irmão do subgênero *Drosophila*, definido inclusivamente mediante a presença dos drosofilídeos hawaianos. Resultados semelhantes foram ainda obtidos por Tatarenkov *et al.* (1999a e 2001), que apresentam *Hirtodrosophila* como um dos primeiros ramos a divergir depois de *Sophophora*. Kwiatowski *et al.* (1997), por outro lado, favoreceram a topologia, primeiramente apresentada por Throckmorton (1975), de que *Hirtodrosophila* e *Dorsilopha* compõem táxons irmãos entre si, e filogeneticamente aparentados ao subgênero *Drosophila*. Com base neste resultado, estes autores defendem que *Hirtodrosophila* mantenha o status de subgênero pertencente ao gênero *Drosophila*. A designação de *Hirtodrosophila* como subgênero de *Drosophila* foi ainda defendida por Kwiatowski & Ayala (1999), segundo os quais este táxon é proximoamente relacionado ao grupo *immigrans*, visão esta compartilhada por Katoh *et al.* (2007).

Mais recentemente, entretanto, Remsen & O'Grady (2002) recuperaram o agrupamento de *Hirtodrosophila* junto a outros gêneros micófagos, tais como *Mycodrosophila* e *Paramycodrosophila*, bem como a *Samoaia*, de modo a compor uma ramificação basal da subfamília Drosophilinae. *Mycodrosophila* e *Hirtodrosophila* também constituíram linhagens irmãs nos trabalhos de Katoh *et al.* (2000) e de DaLage *et al.* (2007), embora nestes casos, o clado assim estabelecido tenha se posicionado junto à base do subgênero *Drosophila*.

### 3.4 – Posicionamento de *Liodrosophila*

Segundo Throckmorton (1975), *Liodrosophila* situa-se na base da radiação *Drosophila*, de forma a constituir uma das primeiras linhagens a divergir depois de

*Sophophora*. Semelhantemente, DeSalle (1992b) defende que *Liodrosophila* compõe a linhagem irmã do subgênero *Drosophila*, monofileticamente definido. Tatarenkov *et al.* (1999a e 2001), por outro lado, apresenta *Liodrosophila* como uma ramificação interna ao subgênero *Drosophila*, embora o suporte para esta relação seja bastante reduzido.

Este cenário é inteiramente discordante àquele apresentado por Grimaldi (1990), segundo o qual *Liodrosophila* agrupa-se a diversos outros gêneros, tais como *Sphaerogastrella* e *Dettopsomyia*, sendo que estes, uma vez acrescentados do clado composto por *Samoaia* e *Zaprionus*, constituem parte do clado irmão do gênero *Drosophila* monofileticamente constituído. O agrupamento de *Liodrosophila* com *Sphaerogastrella* e *Dettopsomyia*, de modo a compor o clado *Formosa*, foi novamente recuperado por Remsen & O'Grady (2002). Neste caso, entretanto, o clado assim constituído compôs o grupo-externo direto de um gênero *Drosophila* inteiramente parafilético, amplamente misturado a membros de outros gêneros como *Chymomyza*, *Scaptodrosophila*, *Zaprionus* e *Scaptomyza*.

Contrariamente a este cenário, entretanto, DaLage *et al.* (2007) apresentaram *Liodrosophila* agrupado aos gêneros *Hirtodrosophila* e *Mycodrosophila* como primeira linhagem a divergir depois de *Sophophora* e *Zaprionus*.

### 3.5 – Posicionamento de *Samoaia*

Assim como para *Liodrosophila*, a designação genérica de *Samoaia* foi suportada por Grimaldi (1990), que apresentou este táxon como parte do grupo genérico *Zaprionus*, que juntamente com o grupo genérico *Styloptera* compunha o clado-irmão do gênero *Drosophila*, monofileticamente constituído. Throckmorton (1975), por outro lado, apesar de também apresentar *Samoaia* como proximamente relacionado a

*Zaprionus*, apresenta ambos como parte de sua radiação *immigrans-Hirtodrosophila*, resultado este parcialmente recuperado por Pélandakis & Solignac (1993).

Mais recentemente, um cenário semelhante ao originalmente proposto por Throckmorton (1975) foi suportado por Tatarenkov *et al.* (1999a e 2001). Estes trabalhos, entretanto, diferem quanto ao posicionamento exato de *Samoaia* dentro da radiação *immigrans-Hirtodrosophila*, possivelmente devido a diferenças de amostragem. Em Tatarenkov *et al.* (1999a), *Samoaia* agrupa com *D. immigrans*, com baixo suporte, enquanto que em Tatarenkov *et al.* (2001), *Samoaia* constitui a espécie irmã do clado constituído pelos grupos *immigrans*, *funebri*, *quinaria*, *testaceae* e *tripunctata*, de modo a compor uma ramificação basal da radiação *immigrans-tripunctata*. Em ambos os casos, o cenário defendido apresenta *Zaprionus* como grupo externo do subgênero *Drosophila*.

O trabalho de Remsen & O'Grady (2002), entretanto, trouxe novas incongruências a esta questão. Neste caso, *Samoaia* agrupou-se com membros do grupo genérico *Zygothrica* de Grimaldi (1990) (representados pelos táxons micófagos *Hirtodrosophila*, *Mycodrosophila* e *Paramycodrosophila*) junto à base do clado maior que constitui a subfamília Drosophilinae. Segundo os autores, este relacionamento é bastante surpreendente, uma vez que diversos representantes do grupo genérico *Zygothrica* são também endêmicos da Ilha de Samoa.

### 3.6 – Posicionamento de *Dorsilopha*

Throckmorton (1975) posicionou o subgênero *Dorsilopha* internamente à radiação *immigrans-Hirtodrosophila*. Posteriormente, Grimaldi (1990) favoreceu a noção de que *Dorsilopha* compõe a linhagem-irmã do clado formado pelos subgêneros *Drosophila* e *Sophophora*.

Mais recentemente, Kwiatowski *et al.* (1997) relembrou aspectos do cenário proposto por Throckmorton (1975) ao favorecer *Dorsilopha* e *Hirtodrosophila* como linhagens irmãs. Neste caso, embora o clado assim constituído corresponda à linhagem irmã do subgênero *Drosophila*, e não do clado *immigrans-tripunctata*, é importante lembrar que apenas representantes da seção *virilis-repleta* foram amostrados dentre os representantes do subgênero *Drosophila*. Kwiatowski & Ayala (1999), Tatarenkov *et al.* (1999a e 2001) e Katoh *et al.* (2007) também apresentaram *Dorsilopha* como um grupo externo ao subgênero *Drosophila*, sendo que nestes trabalhos o mesmo apresenta-se, geralmente, como um dos primeiros, ou até mesmo o primeiro táxon a divergir após *Sophophora*. Em Remsen & O'Grady (2002), entretanto, *Dorsilopha* agrupa-se ao gênero *Zaprionus* de modo a constituir a linhagem irmã do clado formado por *Chymomyza*, *Scaptodrosophila* e pelo subgênero *Sophophora*.

### 3.7 – Monofilia e posicionamento do subgênero *Sophophora*

Segundo Throckmorton (1975), o subgênero *Sophophora* posiciona-se na base da filogenia de *Drosophila*, sendo este filogeneticamente mais próximo de diversos outros gêneros e subgêneros (*Dorsilopha*, *Hirtodrosophila*, *Samoaia*, *Liodrosophila*, *Scaptomyza*, *Engiscaptomyza*, *Zaprionus*, etc.) do que do próprio subgênero *Sophophora*. De acordo com sua visão, a radiação *Sophophora* emerge a partir da radiação *Scaptodrosophila*, e compreende quatro linhagens principais: o grupo *obscura*; o grupo *melanogaster*; os grupos *saltans* e *willistoni*; e os gêneros *Chymomyza* e *Neotanygastrella*. Em seu trabalho de 1962, Throckmorton apresenta, ainda, a possibilidade de que os grupos *obscura* e *melanogaster* tenham se originado a partir de uma única população *proto-melanogaster* ancestral.

No cladograma de Grimaldi (1990), entretanto, o subgênero *Sophophora* constitui o grupo-irmão do subgênero *Drosophila*, sendo que estes, uma vez acrescidos

do subgênero *Dorsilopha* constituiriam o gênero *Drosophila* em sua totalidade. Este último resultado foi, em parte, corroborado pelos achados de DeSalle (1992b), que apresentam o subgênero *Sophophora* como táxon irmão do clado constituído pelo subgênero *Drosophila* e pelo gênero *Liodrosophila*.

Posteriormente, entretanto, a maior parte dos dados moleculares tem favorecido a hipótese desenvolvida por Throckmorton (1975), de que o subgênero *Sophophora* constitui o primeiro ramo a divergir de todas as demais espécies do gênero *Drosophila*, bem como de outros gêneros (Thomas & Hunt, 1993; Russo *et al.*, 1995; Kwiatowski *et al.*, 1997; Remsen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000 e 2007; DaLage *et al.*, 2007). Russo *et al.* (1995) estimaram que esta separação se deu há aproximadamente 40 milhões de anos. No trabalho de Yotoko *et al.* (2003), porém, o subgênero *Sophophora* constituiu o segundo grupo a divergir, ramificando apenas após a separação de *Hirtodrosophila*. Na análise de Remsen & O'Grady (2002), por outro lado, o subgênero *Sophophora*, monofileticamente definido, agrupa-se ao clado constituído pelos gêneros *Scaptodrosophila* e *Chymomyza*, de modo a constituir o grupo irmão do cluster *Zaprionus/Dorsilopha*. Neste caso, o clado assim estabelecido forma o grupo-irmão do subgênero *Drosophila*, parafileticamente definido com relação ao gênero *Scaptomyza* e ao subgênero *Siphlodora*.

Outra pendência importante emergiu com os achados de Pélandakis *et al.* (1991), Pélandakis & Solignac (1993) e Kwiatowski *et al.* (1997) que ressuscitaram a questão, primeiramente levantada por Throckmorton (1975), de que a linhagem *saltans/willistoni* pode não formar um grupo monofilético com a linhagem *melanogaster/obscura*, de forma que o subgênero *Sophophora* constitui um táxon parafiletico com relação ao subgênero *Drosophila*. Além disso, Katoh *et al.* (2000) sugerem que o *Sophophora* é

parafilético com relação a alguns membros do gênero *Lordiphosa*, que constituem o clado-irmão da linhagem Neotropical formada pelos grupos *saltans* e *willistoni*.

A monofilia e o posicionamento basal do subgênero *Sophophora* foi, porém, confirmada pelo estudo de Tatarenkov *et al.* (1999a), onde todas as espécies de *Sophophora* apresentaram uma deleção de um único códon na região codificadora do gene *Ddc*, não encontrada em qualquer outra espécie do gênero *Drosophila*. Os problemas no correto posicionamento dos grupos *saltans* e *willistoni* são atribuídos, por diversos autores (Rodriguez-Trelles *et al.*, 1999b, 2000a e 2000b; Tatarenkov *et al.*, 1999a; Tarrío *et al.*, 2000; Powell *et al.*, 2003; DaLage *et al.*, 2007) à evolução molecular atípica apresentada por estes grupos, caracterizados por altas taxas de substituição não-sinônima (ao menos para o gene *Xdh*) e baixa proporção GC.

### 3.8 – Filogenia do subgênero *Drosophila*

Segundo Throckmorton (1975), o subgênero *Drosophila* consiste de dois cladogramas principais: a radiação *virilis-repleta* e a radiação *immigrans-Hirtodrosophila* (da qual deriva a radiação *tripunctata*), além dos táxons *Liodrosophila* e *D. funebris*, basais em relação a ambas radiações segundo o mesmo autor. Estimativas recentes do tempo de divergência destas radiações sugerem que as mesmas tenham se originado a partir de um ancestral comum existente há mais de 33 milhões de anos atrás (Russo *et al.*, 1995). Ainda segundo Throckmorton, o subgênero *Drosophila* dera origem, em tempos remotos, a diversos outros gêneros e subgêneros, de forma que não constitui um agrupamento monofilético. Grimaldi (1990), DeSalle (1992a e b) e Powell & DeSalle (1995), entretanto, favoreceram a monofilia do subgênero *Drosophila*, sendo que estes três últimos trabalhos o situam como grupo-irmão aos drosofilídeos hawaianos.

Uma série de trabalhos posteriores, entretanto, têm indicado que o subgênero *Drosophila* é, na verdade, parafilético (Russo *et al.*, 1995; Remsen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000 e 2007; Remsen & O’Grady, 2002). No trabalho de Kwiatowski & Ayala (1999), por exemplo, *D. immigrans* aparece como mais proximamente relacionada ao subgênero *Hirtodrosophila* do que às demais espécies amostradas do subgênero *Drosophila*. Este parecer, na verdade, apenas confirma a visão de Throckmorton (1975). Recentemente, entretanto, Yotoko *et al.* (2003) contradisse esta visão ao apresentar um subgênero *Drosophila* monofilético para as espécies consideradas, e estatisticamente distante ao subgênero *Hirtodrosophila*.

### 3.8.1 – Filogenia dos membros da radiação *virilis-repleta*

As relações filogenéticas dos grupos e espécies na radiação *virilis/repleta* encontram-se bastante mal resolvidas na abordagem de Throckmorton (1975). Durando *et al.* (2000) e O’Grady *et al.* (2001) sugerem que o grupo *repleta* é, na verdade, parafilético com relação aos grupos *dreyfusi*, *canalinae* e *mesophragmatica*, embora no primeiro destes trabalhos os autores ressaltem que suas conclusões encontram-se afetadas pela baixa resolução apresentada na base de sua árvore. Ainda em Durando *et al.* (2000), dos cinco subgrupos pertencentes ao grupo *repleta*, apenas *mercatorum* e *hydei* são vistos como monofiléticos (o subgrupo *mulleri* revela-se polifilético enquanto que os representantes do subgrupo *repleta* aparecem como parafiléticos com relação ao subgrupo *mercatorum*). Entretanto, nas análises de Tatarenkov & Ayala (2001) e DaLage *et al.* (2007), o grupo *repleta* é monofileticamente recuperado e, com exceção do grupo *tumiditarsus*, todos os demais membros da radiação *virilis-repleta* formam um *cluster* monofilético, proximamente relacionado aos drosofilídeos hawaianos no primeiro destes trabalhos e na análise bayesiana do segundo. A parte mais derivada da



filogenia apresentada por Tatarenkov & Ayala (2001) contém todos os representantes do grupo *repleta*, cujas relações são altamente congruentes com os estudos prévios de Throckmorton (1975) e Wassermann (1982), com exceção da inconsistência acerca do posicionamento correto dos subgrupos *mercatorum* e *repleta*, apresentados como grupos-irmãos nas filogenias destes últimos. Wang *et al.* (2006), entretanto, apresentam o grupo *repleta* como uma ramificação basal da seção *virilis-repleta*, sendo que neste caso o clado composto por este grupo, juntamente com os grupos *annulimana* e *nannoptera*, compõe a segunda linhagem a divergir, perdendo apenas para o grupo *polychaeta*.

*D. gaucha* (grupo *mesophragmatica*), *D. camargoi* (grupo *dreyfusi*) e *D. canalinea* (grupo *canalinea*) são os grupos mais próximos ao grupo *repleta*, de modo a constituir uma radiação *repleta* monofilética (Throckmorton, 1975; Tatarenkov & Ayala, 2001). Neste caso, Tatarenkov & Ayala (2001) apresentam os grupos *mesophragmatica* e *repleta* como irmãos a partir da análise de seqüências do gene *Amd*, enquanto que na análise realizada com base no gene *Ddc*, bem como na análise combinada *Amd* + *Ddc*, os grupos *mesophragmatica* e *dreyfusi* aparecem como grupos irmãos. No cenário proposto por Throckmorton (1975), por outro lado, os grupos *canalinea* e *dreyfusi* apresentam-se mais proximamente relacionados entre si. Em Remsen & O'Grady (2002), por outro lado, *repleta* aparece como grupo irmão de *dreyfusi*, enquanto *mesophragmatica* aparece como grupo irmão de *bromeliae*, de forma que este trabalho não recupera uma radiação *repleta* monofilética. Mais recentemente, entretanto, Wang *et al.* (2006), apresentam o grupo *annulimana* como táxon irmão do grupo *repleta*, embora este resultado deva-se, provavelmente, à amostragem deficiente de outros membros da radiação, dada a ausência dos grupos *coffeata*, *dreyfusi*, *canalinae* e *mesophragmatica*. DaLage *et al.* (2007) também não incluiu representantes dos grupos

*coffeata* e *canalinae* em sua análise, mas recuperou o agrupamento de *repleta*, *mesophragmatica* e *dreyfusi* na análise bayesiana realizada com base em seqüências nucleotídicas do gene *Amyrel*, sendo que neste caso *repleta* e *mesophragmatica* constituem grupos irmãos entre si.

Outra incongruência com relação aos achados de Tatarenkov & Ayala (2001) é apresentada por Remsen & O'Grady (2002) quanto ao posicionamento do grupo *bromeliae*: enquanto o primeiro trabalho o situa junto ao grupo *nannoptera*, o segundo apresenta *bromeliae* como táxon irmão do grupo *mesophragmatica*. Ainda em Tatarenkov & Ayala (2001), o posicionamento do clado composto pelos grupos *bromeliae* e *nannoptera* como linhagem irmã da radiação *repleta* é fortemente suportado na maior parte das análises realizadas. Neste caso, o grupo *annulimana* constitui a próxima linhagem a divergir, posicionando-se como grupo-externo ao clado formado pela radiação *repleta* + *bromeliae/nannoptera*. Uma ordem de divergência diferente foi apresentada por Wang *et al.* (2006), segundo os quais *D. nannoptera* (grupo *nannoptera*) diverge antes de *D. aracataca* (grupo *annulimana*). Em DaLage *et al.* (2007), por outro lado, o grupo *bromeliae* apresenta um posicionamento inconstante mediante o uso de diferentes métodos de reconstrução filogenética, ora agrupando-se com os grupos *annulimana* e *dreyfusi*, de modo a constituir a linhagem irmã do clado formado pelos grupos *repleta* e *mesophragmatica* (máxima parcimônia), ora agrupando-se com o grupo *polychaeta* junto à base da radiação *virilis-repleta* (análise bayesiana). Em ambas as análises, entretanto, o grupo *annulimana* destaca-se como proximamente relacionado aos membros da radiação *repleta*.

Um clado congruente e bem definido, apresentado nos estudos de Tatarenkov & Ayala (2001), Remsen & O'Grady (2002) e Wang *et al.* (2006), é constituído pelos grupos *robusta* e *melanica*. Neste último trabalho, inclusive, estes dois grupos

apresentam uma composição parafilética com relação um ao outro (e com relação aos grupos *clefta* e *quadrissetata*, no que se refere apenas ao grupo *robusta*). Nos trabalhos anteriores de DeSalle (1992b), entretanto, representantes do grupo *melanica* agruparam-se com o clado constituído por *D. funebris*/*D. pinicola*, com *D. robusta* ramificando em uma posição intermediária da filogenia. O posicionamento correto do grupo *virilis* também permanece bastante confuso, sendo que nas análises de Powell & DeSalle (1995), Remsen & DeSalle (1998), Remsen & O'Grady (2002) e de DaLage *et al.* (2007), *D. virilis* posiciona-se mais próximo aos membros da radiação *repleta* do que a *D. melanica* e *D. robusta*, enquanto na análise de Katoh *et al.* (2000) e em Tatarenkov *et al.* (2001), o grupo *virilis* é agrupado com o grupo *robusta*. Tatarenkov & Ayala (2001) preferem deixar o posicionamento do grupo *virilis* em aberto, dado que o agrupamento do mesmo junto ao clado que constitui a radiação *repleta* é apenas fracamente suportado. No estudo de Wang *et al.* (2006), por outro lado, o grupo *virilis* apresenta-se como linhagem irmã do clado composto pelos grupos *robusta*, *melanica*, *quadrissetata* e *clefta*.

De qualquer forma, a maior parte dos trabalhos concorda quanto ao posicionamento do grupo *polychaeta* como ramificação basal da radiação *virilis-repleta* (Throckmorton, 1975; Tatarenkov & Ayala, 2001; Wang *et al.*, 2006, DaLage *et al.*, 2007).

### 3.8.2 - Posicionamento de *Scaptomyza* e de *Drosophila* hawaianas

Throckmorton (1975) considerou os grupos hawaianos de *Drosophila* e *Scaptomyza/Engyscaptomyza* como clados-irmãos e derivados com relação aos demais membros do subgênero *Drosophila*, agrupando-os junto à radiação *immigrans-Hirtodrosophila*. Grimaldi (1990), por outro lado, propôs que as espécies hawaianas de *Drosophila* representariam uma ramificação basal na subfamília Drosophilinae,

devendo ser elevados à categoria de gênero (*Idiomya*), enquanto *Scaptomyza* constituiria o grupo-irmão do clado formado pelos gêneros *Drosophila* e *Zaprionus*. Uma série de trabalhos posteriores (Thomas & Hunt, 1991 e 1993; DeSalle, 1992a; Powell & DeSalle, 1995), entretanto, descrevem as espécies hawaianas de *Scaptomyza* e *Drosophila* como uma ramificação basal e irmã do subgênero *Drosophila*.

Mais recentemente, diversos trabalhos (Russo *et al.*, 1995; Remsen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000; Remsen & O'Grady, 2002) tem confirmado a idéia de que as espécies hawaianas de *Drosophila* e *Scaptomyza/Engiscaptomyza* formam um grupo monofilético, que compõe a linhagem irmã da seção *virilis-repleta*. De modo geral, a visão é congruente com a hipótese, defendida por Throckmorton (1975), de que todos os drosofilídeos hawaianos originaram-se a partir de uma única espécie ancestral, mas é incongruente com sua visão de que esta linhagem constituiria parte da radiação *immigrans-Hirtodrosophila*.

Novas perspectivas foram, entretanto, acrescidas a este cenário com o trabalho de Katoh *et al.* (2007), que apresentou o clado constituído por *D. annulipes* (grupo *immigrans*) e *D. maculinotata* (grupo *funnebris*) como grupo irmão do clado formado por diversas espécies hawaianas de *Drosophila*. Neste caso, o clado assim definido constitui a linhagem irmã do gênero *Scaptomyza*, sendo que todas estas espécies agrupadas apresentam a seção *virilis-repleta* como grupo-externo mais próximo. Como *D. annulipes* e *D. maculinotata* apresentam uma distribuição restrita à Ásia, estes fatos sugerem que o ancestral das espécies hawaianas de *Drosophila* apresente uma origem asiática. Os autores destacam, entretanto, que este quadro é extremamente especulativo, sendo que novos estudos incluindo estas e outras espécies relacionadas tornam-se necessários para auxiliar na resolução da questão.

### 3.8.3 - Filogenia dos membros da radiação *immigrans-tripunctata*

Embora o grupo *tripunctata* represente um dos grupos de *Drosophila* mais especiosos da região Neotropical, onde apresenta 79 espécies identificadas (Bächli, 2008), sua sistemática tem sido largamente negligenciada. Throckmorton (1975) incluiu o grupo *tripunctata* dentro da radiação *tripunctata*, juntamente com outros grupos: *rubrifrons*, *guarani*, *calloptera*, *pallidipennis*, *cardini*, *sticta* e *macroptera*. Segundo este autor, este grupo é pobremente definido, e proximamente relacionado aos outros grupos de sua radiação, servindo ainda como base a partir da qual outros grupos surgem. Neste último ponto, Throckmorton concorda com Frota-Pessoa (1954), que dividira este grupo em outros quatro subgrupos, apontando que, possivelmente, os subgrupos I e III eram transicionais a outros grupos, como *guarani* e *cardini*, respectivamente, de forma que o grupo constituiria um conjunto parafilético de espécies. Throckmorton (1975) situa, desta forma, a radiação *tripunctata* como membro derivado da radiação *immigrans-Hirtodrosophila*, que apresenta os grupos *immigrans*, *testacea*, *quinaria*, além de outros gêneros e subgêneros, como ramificações basais. Como mencionado anteriormente, segundo Throckmorton (1975) o grupo *funnebris*, bem como o gênero *Liodrosophila* compunham divergências basais às radiações *immigrans-Hirtodrosophila* e *virilis-repleta*.

Na análise executada por Tatarenkov *et al.* (2001), *D. tripunctata* e *D. putrida* (grupo *testacea*) agrupam-se com *D. funnebris* e *D. phalerata* (grupo *quinaria*) com alto suporte estatístico, sendo *D. Immigrans*, *Samoia* e *Zaprionus* os grupos-irmãos mais próximos deste clado. Assim, neste estudo, a radiação *immigrans-Hirtodrosophila*, diferencialmente definida mediante a exclusão de *Dorsilopha* e *Hirtodrosophila*, e à inclusão de *D. funnebris*, revela-se monofilética. Resultados semelhantes foram ainda apresentados por Remsen & DeSalle (1998), onde *D. immigrans* destaca-se como

grupo-externo ao clado constituído pelos grupos irmãos *funebri* e *pinicola*, de modo a recuperar uma radiação *immigrans-tripunctata* monofilética para os representantes considerados, mas não uma radiação *immigrans-Hirtodrosophila*, dada a exclusão de *Hirtodrosophila* e *Zaprionus* do clado assim estabelecido. Também no trabalho de Katoh *et al.* (2000) a radiação *immigrans-Hirtodrosophila* não foi devidamente recuperada: embora estes autores confirmem a monofilia do clado composto por membros dos grupos *immigrans*, *funebri* e *quinaria*, o clado constituído por representantes dos gêneros *Hirtodrosophila* e *Mycodrosophila*, juntamente com alguns membros do gênero *Lordiphosa*, apresentou-se mais proximamente relacionado à linhagem *virilis-repleta*/drosofilídeos hawaianos do que aos demais membros da sua radiação. Da mesma forma, em Remsen & O'Grady (2002) e na análise de Yotoko *et al.* (2003) apenas a linhagem *immigrans-tripunctata* se constitui monofileticamente, de modo a formar o clado-irmão à linhagem formada pela radiação *virilis-repleta/Drosophilidae* Hawaianos, mas novamente mediante a exclusão de *Hirtodrosophila* e de *Zaprionus*. Perlman *et al.* (2003) também recuperaram a monofilia da radiação *immigrans-tripunctata*, em detrimento à fracamente suportada polifilia da radiação *tripunctata*. Todos estes resultados encontram-se em inteira discordância aos achados de Kwiatowski & Ayala (1999), que apresentam *D. immigrans* como espécie-irmã de *Hirtodrosophila*, sendo *Zaprionus* frequentemente recuperado como ramificação basal a este clado. Resultados semelhantes também são recuperados por Tatarenkov *et al.* (1999a), para a análise combinada das primeiras e segundas posições dos códons de *Ddc*, *Sod* e *Adh*. Embora um clado bastante semelhante à radiação *immigrans-Hirtodrosophila* original (Throckmorton, 1975) tenha sido ainda recentemente encontrado por Katoh *et al.* (2007) (dada a exclusão dos drosofilídeos hawaianos), o mesmo não apresenta maior suporte estatístico.

Resultados diferentes à todas análises mencionadas foram recentemente obtidos por DaLage *et al.* (2007), uma vez que neste caso não apenas os gêneros *Zaprionus* e *Hirtodrosophila*, mas também o grupo *immigrans*, situaram-se como grupos-externos ao clado constituído pelos demais representantes da radiação *immigrans-Hirtodrosophila* juntamente à radiação *virilis-repleta*. A radiação *immigrans-tripunctata* assim definida apresentou o grupo *quinaria*, agrupado a *funebri* e *testacea* como clado basal, de modo a recuperar uma radiação *tripunctata* monofilética, composta pelos grupos *cardini*, *guarani*, *pallidipennis* e *tripunctata*. Yotoko *et al.* (2003), por outro lado, recuperam uma radiação *tripunctata* parafilética, dada a inclusão de *D. funebri* e *D. putrida* (grupo *testaceae*) dentro do *cluster* composto majoritariamente por membros daquela radiação, que apresenta os grupos *quinaria* e *immigrans* como grupos-externos. Já em Remsen & O'Grady (2002), o grupo *tripunctata* agrupa-se diretamente com o grupo *testaceae*, enquanto os grupos *funebri/pinicola* constituem o grupo-irmão do clado formado pelos grupos *guarani/calloptera/cardini* e pelos grupos *bizonata/macroptera/rubrifrons/quinaria*, o que acaba por constituir uma radiação *tripunctata* polifilética. Assim, de modo geral, todos estes resultados contradizem Throckmorton (1975), segundo o qual o grupo *testaceae* constituiria, juntamente com os grupos *quinaria* e *immigrans*, a parte basal da radiação *immigrans-Hirtodrosophila*, enquanto o grupo *funebri* constituiria uma ramificação basal da radiação *Drosophila*. Por outro lado, alguns destes achados confirmam parcialmente os resultados de DeSalle (1992b), ao conferir uma posição derivada ao grupo *funebri*. Neste trabalho, o grupo *funebri* agrupa-se com *D. pinicola*, sob alto suporte estatístico, embora os demais representantes das radiações *immigrans-tripunctata* e *virilis-repleta* encontrem-se inteiramente misturados topologicamente.

O trabalho recentemente publicado por Yotoko *et al.* (2003) também se destaca como o primeiro a incluir diversos membros do grupo *tripunctata* em uma abordagem

molecular. Neste caso, o grupo *tripunctata* apresentou uma composição parafilética e altamente descontínua ao longo da filogenia, na qual alguns de seus membros mesclam-se com representantes dos grupos *calloptera*, *funebri*, *testaceae*, *guarani* e *pallidipennis*, sendo impossível identificar qualquer grupo original, provavelmente devido a episódios de especiação rápida. Da mesma forma, dentre estes grupos, aqueles em que mais de uma espécie fora considerada (*calloptera* e *guarani*), revelaram-se polifiléticos, com exceção do grupo *cardini*, monofileticamente definido. Ao incluir três representantes do grupo *tripunctata*, DaLage *et al.* (2007) também recuperam seu caráter merofilético.

Por tudo isso, a maior parte dos trabalhos considerados tem indicado que o gênero *Drosophila* é, na verdade, um grupo parafilético, que reúne representantes de diversos outros gêneros de Drosophilinae. Assim, embora Grimaldi (1990) tenha defendido a hipótese de que o gênero *Drosophila* forme um grupo monofilético, derivado e irmão de outros gêneros de Drosophilinae, a visão de Throckmorton (1975), apresentando outros subgêneros e gêneros surgindo a partir do gênero *Drosophila* (que ocupa uma posição basal dentro da subfamília Drosophilinae) ainda parece ser a mais confiável.



## OBJETIVOS

---

Dadas as circunstâncias anteriormente mencionadas, a presente tese tem como objetivo geral contribuir para o entendimento da história evolutiva do gênero *Drosophila* e outros gêneros afins, pelo delineamento dos seguintes objetivos específicos:

### Capítulos II e III:

1) Resolver questões relacionadas ao correto posicionamento filogenético dos gêneros *Hirtodrosophila*, *Liodrosophila*, *Samoaia*, *Scaptomyza* e *Zaprionus*, bem como dos subgêneros *Dorsilopha* e *Sophophora*, com relação ao gênero e ao subgênero *Drosophila*;

2) Analisar a monofilia do gênero e do subgênero *Drosophila*, bem como seus posicionamentos dentro da filogenia de Drosophilidae;

3) Construir uma filogenia robusta para o subgênero *Drosophila*, buscando solucionar aspectos relacionados à sua subdivisão interna e às relações filogenéticas entre e dentro de grupos de espécies tradicionalmente alocados às suas duas maiores radiações: grupos *cardini*, *calloptera*, *guarani*, *guaramunu*, *immigrans*, *pallidipennis*, *quinaria*, *testacea* e *tripunctata*, incluídos na radiação *immigrans-tripunctata*; grupos *annulimana*, *bromeliae*, *canalinae*, *dreyfusi*, *flavopilosa*, *melanica*, *mesophragmatica*, *nannoptera*, *polychaeta*, *repleta*, *robusta* e *virilis*, pertencentes à radiação *virilis-repleta*;

4) Acessar a monofilia e o posicionamento dos drosofilídeos hawaianos com relação ao subgênero *Drosophila* (apenas para **Capítulo III**);

5) Datar eventos de divergência, bem como de origem e diversificação de clados pertencentes ao subgênero *Drosophila* (apenas para **Capítulo III**).

#### **Capítulo IV:**

1) Contribuir para o entendimento das relações filogenéticas entre seis das 13 espécies tradicionalmente situadas no grupo *mesophragmatica* de *Drosophila*;

2) Inferir os padrões espaço-temporais de evolução do grupo *mesophragmatica*, a partir da construção de sua filogenia;

3) Analisar aspectos relacionados à genealogia e evolução de caracteres previamente descritos para o grupo ao longo da filogenia obtida;

4) Fazer novas proposições acerca da taxonomia do grupo.

#### **Capítulo V:**

1) Acessar as relações filogenéticas entre espécies, subespécies e semiespécies pertencentes ao subgrupo *willistoni*;

2) Propor um cenário evolutivo congruente com os padrões filogenéticos encontrados;

3) Datar alguns eventos de divergência, origem e diversificação dentro do subgrupo *willistoni*;

4) Testar os níveis de divergência molecular apresentados pelos diferentes táxons que compõem o subgrupo *willistoni*, analisando sua correlação com os graus de isolamento reprodutivo.

## CAPÍTULO II

---

Trabalho publicado na revista *Molecular Phylogenetics and Evolution*

### **Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: A nuclear versus mitochondrial gene approach<sup>10</sup>**

Lizandra J. Robe<sup>1</sup>, Vera L. S. Valente<sup>1</sup>, Myriam Budnik<sup>2</sup>, Elgion L. S. Loreto<sup>3\*</sup>

1 – Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil

2 – Instituto de Ciências Biomédicas, Faculdade de Medicina, Universidade do Chile, Santiago, Chile

3 – Departamento de Biologia, Universidade Federal de Santa Maria (UFSM), Camobi Km 09, CP5050, Santa Maria, Rio Grande do Sul, Brazil, 97105-900

\* Corresponding author: Phone (55)(55) 220-8912; Fax (55)(55) 220-8020

E-mail addresses: [lizbiogen@pop.com.br](mailto:lizbiogen@pop.com.br) (L. J. Robe), [vera.valente@bol.com.br](mailto:vera.valente@bol.com.br) (V. L. S. Valente), [mbudnik@med.uchile.cl](mailto:mbudnik@med.uchile.cl) (M. Budnik), [elgion@base.ufsm.br](mailto:elgion@base.ufsm.br) (É. L. S. Loreto)

---

<sup>10</sup> A parte inicial deste trabalho, referente à coleta dos dados de *Amd*, foi apresentada na dissertação de mestrado da doutoranda, Lizandra Jaqueline Robe.

# Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: A nuclear versus mitochondrial gene approach

Lizandra J. Robe<sup>a</sup>, Vera L.S. Valente<sup>a</sup>, Myriam Budnik<sup>b</sup>, Élgion L.S. Loreto<sup>c,\*</sup>

<sup>a</sup> Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil

<sup>b</sup> Instituto de Ciências Biomédicas, Faculdade de Medicina, Universidade do Chile, Santiago, Chile

<sup>c</sup> Departamento de Biologia, Universidade Federal de Santa Maria (UFSM), Camobi Km 09, CP5050, Santa Maria, Rio Grande do Sul 97105-900, Brazil

Received 25 August 2004; revised 26 April 2005

Available online 20 June 2005

## Abstract

The genus *Drosophila* has played an essential role in many biological studies during the last 100 years but much controversy and many incompletely addressed issues still remain to be elucidated regarding the phylogeny of this genus. Because information on the Neotropical species contained in the subgenus *Drosophila* is particularly incomplete, with this taxonomic group being underrepresented in many studies, we designed a study to answer some evolutionary questions related to these species. We subjected at least 41 Drosophilidae taxa to a phylogenetic analysis using a 516-base pair (bp) fragment of the  $\alpha$ -methyl-dopa (*Amd*) nuclear gene and a 672 bp fragment of the mitochondrial cytochrome oxidase subunit II (*COII*) gene both individually and in combination. We found that the subgenus *Drosophila* is paraphyletic and subdivided into two main clusters: the first containing species traditionally placed in the *virilis-repleta* radiation and the second assembling species of the *immigrans-Hirtodrosophila* radiation. Inside the first of these clusters we could detect the monophyly of both the *flavopilosa* (the sister-clade of the *annulimana* group) and the *mesophragmatica* (closely related to the *repleta* group) species groups. Concerning the *immigrans-Hirtodrosophila* lineage, *Zaprionus*, *Liodrosophila*, *Samoaia*, and *Hirtodrosophila* were the early offshoots, followed by the *immigrans*, *quinaria*, *testacea*, and *funebis* species groups. The *tripunctata* radiation appears to be a derived clade, composed of a paraphyletic *tripunctata* group, intimately interposed with members of the *cardini*, *guarani*, and *guaramunu* species groups. Overall, the *COII* gene yielded a poor phylogenetic performance when compared to the *Amd* gene, the evolutionary hypothesis of which agreed with the total evidence tree. This phenomenon can be explained by the fast saturation of transitional substitutions in *COII*, due to strong biases in both base composition and substitution patterns, as also by its great among-site rate variation heterogeneity.

© 2005 Elsevier Inc. All rights reserved.

**Keywords:** *Amd* gene; *cardini* group; *COII* gene; *Drosophila*; *flavopilosa* group; *guaramunu* group; *guarani* group; *mesophragmatica* group; Neotropical species; *tripunctata* group

## 1. Introduction

The family Drosophilidae (Diptera) includes about 3000 species distributed over more than 60 genera, with

about half of the total number of species as yet described belonging to the genus *Drosophila* (Wheeler, 1982). According to this author, the genus *Drosophila* has a wide and scattered distribution, being composed of 15

\* Corresponding author. Fax: +55 55 220 8020.

E-mail addresses: [lizbiogen@pop.com.br](mailto:lizbiogen@pop.com.br) (L.J. Robe), [vera.valente@bol.com.br](mailto:vera.valente@bol.com.br) (V.L.S. Valente), [mbudnik@med.uchile.cl](mailto:mbudnik@med.uchile.cl) (M. Budnik), [elgion@base.ufsm.br](mailto:elgion@base.ufsm.br) (É.L.S. Loreto).

subgenera. Among these we distinguish the speciose subgenera *Drosophila* and *Sophophora*, whose systematics in most cases still remains controversial and incomplete.

The largest group of Drosophilidae inhabiting the Neotropics is the subgenus *Drosophila* (Val et al., 1981), where it includes at least 25 species groups. Malogolowkin (1953) and Throckmorton (1975) subdivided these species into two main lineages: the *virilis-repleta* section (or radiation) containing 15 species groups; and the *quinaria-tripunctata* section (also known as the *immigrans-Hirtodrosophila* radiation) containing 9 Neotropical species groups. Besides this, the subgenus *Drosophila* is also composed of the *funebria* group, basal to both radiations according to Throckmorton (1975). This author also supported the thesis that the subgenus *Drosophila* forms a paraphyletic clade from which a series of other genera and subgenera originated ‘at different times and levels.’ Recently, this hypothesis has been substantiated by several studies (Kwiatowski and Ayala, 1999; Remsen and DeSalle, 1998; Remsen and O’Grady, 2002; Russo et al., 1995; Tatarenkov et al., 1999, 2001; Thomas and Hunt, 1993) but it has been negated by others (DeSalle, 1992; Grimaldi, 1990; Powell and DeSalle, 1995). These uncertainties still present in *Drosophila* phylogenetic studies are compounded by the fact that Neotropical members of the subgenus *Drosophila* are poorly represented in many of these studies.

One of the subgenus *Drosophila* species groups that are underrepresented in phylogenetic studies is the *tripunctata* group, a member of the *quinaria-tripunctata* section of Malogolowkin (1953) and Throckmorton (1975). This group is composed of 64 species (Vilela and Bächli, 2000) but its internal phylogeny has rarely been tested, except by the morphological work of Frota-Pessoa (1954) and the recent molecular analysis carried out by Yotoko et al. (2003). According to Throckmorton (1975), the *tripunctata* species group constitutes a paraphyletic clade, situated in his ‘*tripunctata* radiation’ along with other species groups such as *pallidipennis*, *rubrifrons*, *sticta*, *macroptera*, *calloptera*, *cardini*, and *guarani*. Kastritsis (1969) suggested that some members of the *guarani* group were phylogenetically closer to the *tripunctata* group than to other members of its own group, which has led to the subdivision of the former *guarani* species group into the *guarani* and *guaramunu* groups.

Conversely, the *virilis-repleta* section of the subgenus *Drosophila* has been more widely considered in molecular phylogenetic studies (Durando et al., 2000; Nurminsky et al., 1996; Tatarenkov and Ayala, 2001) but still contains a series of gaps concerning the relationships of some of its groups. For example, the Andean *mesophragmatica* species group has not as yet had its essentially conflicting inter-group relationships elucidated (Durando et al., 2000; Remsen and O’Grady, 2002; Tatarenkov and Ayala, 2001; Throckmorton, 1975).

Besides that, its internal subdivision has never been studied using a molecular approach, the only phylogenetic studies to date having been based on morphological (Nacur, 1958), cytogenetic (Brncic et al., 1971) or isozyme variations (Nair et al., 1971). Another example is the fact that even though Throckmorton (1975) questioned the exact derivation of the *flavopilosa* species group no subsequent study has been carried out to clarify this point.

Based on the considerations outlined above, this study was principally designed to elucidate the phylogenetic relationships within and between Neotropical species groups traditionally placed in the subgenus *Drosophila*, i.e., the *immigrans*, *tripunctata*, *cardini*, *guarani*, and *guaramunu* species groups included in the *quinaria-tripunctata* section and the *mesophragmatica* and *flavopilosa* species groups assembled in the *virilis-repleta* section. Besides that, we also wished to shed more light on some questions related to the monophyly of the genus and subgenus *Drosophila* and on the taxonomic status of some of their radiations. We also addressed some questions concerning the correct position of the taxa *Zaprionus*, *Samoia*, *Liodrosophila*, *Hirtodrosophila*, and *Dorsilopha* within the family Drosophilidae. The molecular markers used in this study were the  $\alpha$ -methyl dopa (*Amd*) nuclear gene sequence and the mitochondrial cytochrome oxidase subunit II (*COII*) gene sequences, the partitions being used both separately and combined. These markers were used because they had previously proved to be informative at this level of divergence (Tatarenkov and Ayala, 2001; Tatarenkov et al., 2001; Yotoko et al., 2003). Combining a nuclear and a mitochondrial gene is also a good choice because these two types of data are unlinked and evolving under different evolutionary constraints (Lin and Danforth, 2004).

## 2. Materials and methods

### 2.1. Species

The Drosophilidae taxa used in this study (49 for *Amd*, 71 for *COII*, and 41 in the combined analysis) as well as their respective *Amd* and *COII* sequence GenBank accession numbers at National Center for Biotechnology Information (NCBI) site are given in Table 1, where the new sequences obtained in this study are underlined. The genus *Scaptodrosophila* was utilized as outgroup because there is ample evidence that this is the case with respect to the genus *Drosophila* (DeSalle, 1992; Grimaldi, 1990; Kwiatowski et al., 1994, 1997; Kwiatowski and Ayala, 1999; Pélandakis and Solignac, 1993; Russo et al., 1995; Tatarenkov et al., 1999, 2001; Throckmorton, 1975). Most of the focused groups have a Neotropical restricted distribution.

Table 1

List of *Drosophila* species and other Drosophilidae used in this study, with their GenBank accession numbers for the *Amd* and *COII* sequences

Genus	Subgenus	Section	Group	Species	Amd Acc. Nos.	COII Acc. Nos.
<i>Drosophila</i>	<i>Drosophila</i>	<i>quinaria-tripunctata</i>	<i>guarani</i>	<b><i>D. ornatifrons</i> 1</b>	<u>AY699250</u>	AY162977 <sup>h</sup>
				<i>D. ornatifrons</i> 2	—	AY162978 <sup>h</sup>
				<b><i>D. subbadia</i></b>	<u>AY699251</u>	<u>AY847772</u>
				<i>D. guaru</i>	—	<u>AY847763</u>
			<i>guaramunu</i>	<b><i>D. griseolineata</i></b>	<u>AY699257</u>	AF478424 <sup>f</sup>
				<i>D. maculifrons</i> 1	—	<u>AY847766</u>
				<i>D. maculifrons</i> 2	—	AY162979 <sup>h</sup>
				<i>D. maculifrons</i> 3	—	AY162980 <sup>h</sup>
			<i>tripunctata</i>	<b><i>D. angustibucca</i></b> <sup>A</sup>	<u>AY699248</u>	AY162983 <sup>h</sup>
				<i>D. angustibucca</i> 2 <sup>A</sup>	<u>AY699249</u>	—
				<i>D. mediosignata</i>	—	AY162985 <sup>h</sup>
				<i>D. paraguayensis</i> 1	—	AY162986 <sup>h</sup>
				<i>D. paraguayensis</i> 2	—	AY162987 <sup>h</sup>
				<i>D. cuaso</i> 1	—	AY162984 <sup>h</sup>
				<i>D. cuaso</i> 2	—	AY170441 <sup>h</sup>
				<i>D. metzii</i>	—	AY162992 <sup>h</sup>
				<b><i>D. mediopunctata</i> 1</b>	<u>AY699254</u>	AY162988 <sup>h</sup>
				<b><i>D. mediopunctata</i> 2</b>	<u>AY699255</u>	AY162989 <sup>h</sup>
				<b><i>D. crocina</i></b> =	<u>AY699252</u>	<u>AY847759</u>
				<i>D. mediotriata</i> (Vilela and Bächli, 1990)		
				<b><i>D. mediotriata</i></b>	<u>AY699253</u>	<u>AY847767</u>
				<i>D. paramediotriata</i> 1	—	AY162995 <sup>h</sup>
				<i>D. paramediotriata</i> 2	—	AY162996 <sup>h</sup>
				<b><i>D. bandeirantorum</i> 1</b>	<u>AY699256</u>	AY162990 <sup>h</sup>
				<i>D. bandeirantorum</i> 2	—	AY162991 <sup>h</sup>
				<b><i>D. tripunctata</i> 1</b>	AF293709 <sup>a</sup>	AF478432 <sup>f</sup>
				<i>D. tripunctata</i> 2	—	AF519343 <sup>g</sup>
				<i>D. mediopicta</i>	—	<u>AY847768</u>
				<i>D. medioimpressa</i>	—	AY162994 <sup>h</sup>
			<i>cardini</i>	<i>D. cardini</i> 1	—	AY162974 <sup>h</sup>
				<i>D. cardini</i> 2	—	AF519319 <sup>g</sup>
				<b><i>D. cardinoides</i></b>	<u>AY699258</u>	AY162975 <sup>h</sup>
				<b><i>D. neocardini</i></b>	<u>AY699260</u>	<u>AY847770</u>
				<b><i>D. polymorpha</i></b>	<u>AY699259</u>	AY162976 <sup>h</sup>
			<i>pallidipennis</i>	<i>D. pallidipennis</i> 1	—	AY162981 <sup>h</sup>
				<i>D. pallidipennis</i> 2	—	AY162982 <sup>h</sup>
			<i>calloptera</i>	<i>D. calloptera</i>	—	AF478419 <sup>f</sup>
			<i>testacea</i>	<b><i>D. putrida</i> 1</b>	AF293723 <sup>a</sup>	AF478431 <sup>f</sup>
				<i>D. putrida</i> 2	—	AF519335 <sup>g</sup>
			<i>quinaria</i>	<i>D. quinaria</i>	—	AF478428 <sup>f</sup>
				<b><i>D. phalerata</i></b>	AF293721 <sup>a</sup>	AF147115 <sup>c</sup>
			<i>immigrans</i>	<b><i>D. immigrans</i> 1</b>	<u>AY699261</u>	AF478424 <sup>f</sup>
				<b><i>D. immigrans</i> 2</b>	AF293713 <sup>a</sup>	AY162993 <sup>h</sup>
				<i>D. immigrans</i> 3	—	AF519324 <sup>g</sup>
			<i>funnebris</i>	<b><i>D. funnebris</i></b>	AF293709 <sup>a</sup>	AF478422 <sup>f</sup>
			<i>virilis-repleta</i>	<b><i>D. gasici</i></b> AR	<u>AY699239</u>	<u>AY847760</u>
				<b><i>D. gasici</i></b> CA	<u>AY699240</u>	<u>AY847761</u>
				<b><i>D. gasici</i></b> CO	<u>AY699241</u>	<u>AY847762</u>
				<b><i>D. brncici</i></b>	<u>AY699238</u>	<u>AY847757</u>
				<b><i>D. mesophragmatica</i></b>	<u>AY699242</u>	<u>AY847769</u>
				<b><i>D. gaucha</i></b>	AF324955 <sup>b</sup>	AF478423 <sup>f</sup>
				<b><i>D. pavani</i></b>	<u>AY699245</u>	<u>AY847771</u>
				<b><i>D. viracochi</i></b>	<u>AY699244</u>	<u>AY847773</u>
			<i>repleta</i>	<b><i>D. hydei</i></b>	AF293712 <sup>a</sup>	AF478429 <sup>f</sup>
				<i>D. eohydei</i>	—	AF145889 <sup>d</sup>
				<i>D. nigrohhydei</i>	—	AF145890 <sup>d</sup>
				<b><i>D. buzzatii</i></b>	AF324947 <sup>b</sup>	AF146169 <sup>b</sup>
				<i>D. borborema</i>	—	AF146171 <sup>b</sup>
				<i>D. serido</i>	—	AF146173 <sup>b</sup>
				<i>D. mercatorum</i>	AF324957 <sup>b</sup>	—

(continued on next page)

Table 1 (continued)

Genus	Subgenus	Section	Group	Species	Amd Acc. Nos.	COII Acc. Nos.	
				<i>dreyfusi</i>	<i>D. camargoi</i>	AF324951 <sup>b</sup>	AF478421 <sup>f</sup>
				<i>canalina</i>	<i>D. canalinea</i>	AF324952 <sup>b</sup>	—
				<i>flavopilosa</i>	<i>D. cestri</i>	<u>AY699246</u>	<u>AY847758</u>
					<i>D. incompta</i>	<u>AY699247</u>	<u>AY847764</u>
				<i>annulimana</i>	<i>D. aracatacas</i> <sup>B</sup>	AF324949 <sup>b</sup>	<u>AY847756</u>
				<i>virilis</i>	<i>D. virilis</i>	AF293729 <sup>a</sup>	—
				<i>robusta</i>	<i>D. robusta</i>	AF293724 <sup>a</sup>	—
				<i>nannoptera</i>	<i>D. nannoptera</i>	AF324959 <sup>b</sup>	AF478425 <sup>f</sup>
	<i>Sophophora</i>			<i>melanogaster</i>	<i>D. simulans</i> 1	AY699262	AF474082 <sup>e</sup>
					<i>D. simulans</i> 2	AF293726 <sup>a</sup>	—
					<i>D. erecta</i>	AF293708 <sup>a</sup>	—
					<i>D. yakuba</i>	—	X03240 <sup>a</sup>
			<i>obscura</i>	<i>D. pseudoobscura</i>	AF293722 <sup>a</sup>	—	—
	<i>Dorsilophia</i>			<i>D. busckii</i>	AF293707 <sup>a</sup>	AF519347 <sup>g</sup>	—
<i>Hirtodrosophila</i>				<i>H. pictiventris</i>	AF293711 <sup>a</sup>	AF478434 <sup>f</sup>	—
<i>Liodrosophila</i>				<i>L. aerea</i>	AF293715 <sup>a</sup>	AF478435 <sup>f</sup>	—
<i>Samoaia</i>				<i>S. leonensis</i>	AF293725 <sup>a</sup>	AF478438 <sup>f</sup>	—
<i>Zaprionus</i>				<i>Z. indianus</i>	<u>AY699263</u>	<u>AY847774</u>	—
				<i>Z. tuberculatus</i>	AF293731 <sup>a</sup>	AF478440 <sup>f</sup>	—
<i>Scaptodrosophila</i>				<i>S. latifasciaeformis</i>	<u>AY699264</u>	<u>AY847765</u>	—

(1) The classification proposed by Grimaldi (1990) was followed in this paper. (2) More than one *Drosophila* population was sampled for some of the species listed. This is illustrated by the three *D. gasci* lineages, where population AR was from Arica (Chile), population CA was from Cochabamba (Bolivia), and population CO was from Bogota (Colombia). For the other species the different populations are demarcated by different terminal numbers. (3) Only the species listed in boldface were included on the simultaneous analysis. (4) The sequences newly obtained in this study are underlined.

<sup>A</sup> Species recently redefined as *D. nappae* by Vilela et al. (2004).

<sup>B</sup> For *COII* the *annulimana* species group was represented by *D. annulimana*.

<sup>a</sup> *Amd* sequences from Tatarenkov et al. (2001) and *COII* sequences from Clary et al. (1982).

<sup>b</sup> *Amd* sequences from Tatarenkov and Ayala (2001) and *COII* sequences from Spicer (1995).

<sup>c</sup> *COII* sequences from Spicer and Jaenike (1996).

<sup>d</sup> *COII* sequences from Spicer and Pitnick (1996).

<sup>e</sup> *COII* sequences from O'Grady and Kidwell (2002).

<sup>f</sup> *COII* sequences from Remsen and O'Grady (2002).

<sup>g</sup> *COII* sequences from Perlam et al. (2003).

<sup>h</sup> *COII* sequences from Yotoko et al. (2003).

## 2.2. DNA manipulation

The total DNA of each of the isolinesages established was extracted according to the protocol described by Jowett (1986), with some modifications. The degenerate *Amd-un2* and *Amd-bw* primers (Tatarenkov et al., 2001) were used to amplify a variable length fragment of the *Amd* gene (Fig. 1A) while the TL2-J-3037 and TK-N-3785 primers (Simon et al., 1994) were used to amplify a 684-base pair (bp) fragment of the *COII* gene (Fig. 1B). In each case the final components of the 50  $\mu$ l reaction mixture were: 5  $\mu$ l reaction buffer, 2.5 mM MgCl<sub>2</sub>, 0.25 mM of each dNTP, 20 pmol of each primer, 1 U of *Taq* Polymerase (all from Invitrogen), and 100 ng DNA. The *Amd* amplification parameters varied according to the group under study, following local optimization criteria, while the *COII* amplification parameters were more or less constant: denaturation at 94 °C for 7 min, followed by 35 cycles of denaturation for 1 min at 94 °C, annealing for 1 min at 55 °C, and extension for 75 s at 72 °C, after which the reaction mixture was additionally held at 72 °C for 5 min to complete the extension. The *Amd* amplicons obtained

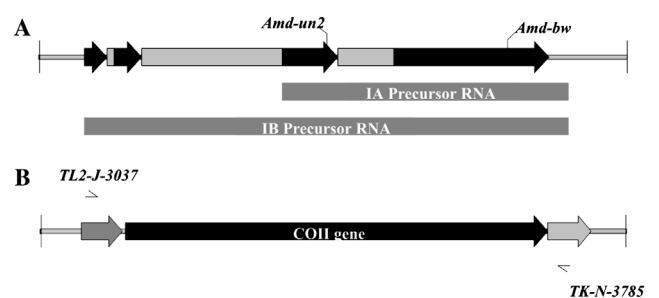


Fig. 1. Genomic maps for the *Amd* and *COII* genes. The arrows denote exons and the clear gray rectangles represent the introns. (A) The *Amd-un2* and *Amd-bw* primers were used to amplify a region of the *Amd* gene encompassing 92 bp (of 471 bp) from the third exon, 991 bp (of 1316 bp) from the fourth exon, and a variable length intron. This gene gives rise to two different precursor RNAs (indicated by the dark gray rectangles below its genomic region) which code for the two *Amd* isoforms. Note that for the B isoform the third exon is spliced off (Adams et al., 2000). (B) The TL2-J-3037 primer anneals to the tRNA-Leu coding region while the TK-N-3785 primer anneals to the tRNA-Lys coding regions (both indicated by the gray arrows), these two primers being used to amplify the entire 684 bp coding region of the *COII* gene.

were separated on 0.8% agarose gel and fragments of the expected size were excised and purified using the Qiaquick DNA-Gel Extraction System (Qiagen), according to the supplier's specifications. The more specific *COII* amplicons were directly purified by incubation at 37 °C for 30 min with exonuclease I and shrimp alkaline phosphatase (SAP) (both from Amersham) followed by a 15 min inactivation step at 80 °C. DNA sequencing was performed directly using the purified amplicons and a MegaBACE 500 automatic sequencer. The dideoxy chain-termination reaction was implemented using the DYEnamic ET (Amersham) kit and the *Amd-bw* primer for *Amd* and the TL2-J-3034 or TK-N-3785 primers for *COII*.

### 2.3. Phylogenetic analysis

The nucleotide sequences had their identity initially confirmed using the BlastN program available at the NCBI site. Inspection and correction of automated sequences were subsequently accomplished using the Staden Package Gap 4 program (Staden, 1996), wherefrom a confidence consensus sequence was obtained for each taxon. These sequences were aligned using the ClustalX 1.81 program (Jeanmougin et al., 1998) according to the system default parameters. This alignment was subjected to a saturation test using the DAMBE program (Xia and Xie, 2001) and to a permutation test probability (PTP) performed on PAUP 4.0b10 (Swofford, 2003).

The individual and combined phylogenetic analyses were executed using four main methods: (1) neighbor joining (NJ) (Saitou and Nei, 1987) using MEGA 2.1 (Kumar et al., 2001); (2) maximum parsimony (MP) using an heuristic search with 100 replicates of random addition of taxa performed using the equal weight criteria as contained in PAUP 4.0b10; (3) maximum likelihood (ML) using PAUP 4.0b10 following the model proposed by the Akaike information criterion (AIC) test (Akaike, 1974) executed using the ModelTest program (Posada and Crandall, 1998); (4) Bayesian analysis performed using the MrBayes 3.0b4 (Huelsenbeck and Ronquist, 2001) according to the model proposed by the ModelTest, with the evaluation of at least 1,000,000 generations and a 'burn-in' region of 2000 trees. The confidence values of each clade for the first three analyses were measured by the bootstrap test (Felsenstein, 1985) with 1000 replications (except for the ML analysis, where only 500 replications were considered). For the Bayesian analysis the posterior probability of each clade on the 50% majority rule consensus tree was calculated (Hall, 2001). Alternative topologies were compared using the Templeton test (Templeton, 1983) performed under the parsimony criteria and Shimodaira–Hasegawa test (Shimodaira and Hasegawa, 1999) performed under the likelihood criteria.

To test for congruence between the two data partitions, 1000 replicates of the partition homogeneity (PHT) test (essentially the incongruence length difference (ILD) test of Farris et al., 1994, 1995), as contained in PAUP 4.0b10, were applied. The partition Bremer support (PBS) scores method (Baker et al., 1998; Bremer, 1988, 1994) was also used to measure the amount of support provided by *Amd* and *COII* to each node on the total evidence phylogeny. These PBS values were calculated using PAUP 4.0b10 and TreeRot.v2 (Sorenson, 1999) and then standardized by dividing the total Bremer support of each gene by the minimum number of steps for that gene (Durando et al., 2000).

The Bayesian framework and the ModelTest results were also used for investigating and characterizing the substitution patterns in our two datasets.

## 3. Results and discussion

### 3.1. Individual analysis of the *Amd* gene

The multiple alignment of the 49 sequences of the *Amd* gene displayed 264 (of 516 bp) variable sites, of which 237 were parsimoniously informative. The average transition/transversion ratio was 1.3, i.e., more than 2.5 times the value expected considering a random substitution distribution. The general frequencies of the four nucleotides T, C, A, and G were, respectively, 25.2, 24.3, 20.8, and 29.8. This slight bias toward the G + C content increased when only the third codon positions were considered, reaching a threshold of 75.4 in *Drosophila griseolineata*. A saturation test (Fig. 2A), performed in order to evaluate the phylogenetic information presented by the sequences (Schneider, 2003), revealed that in most cases both transitions and transversions remained informative. But a portion of the curve for transitions appears to become partially saturated above the 16% divergence value, indicating the occurrence of multiple substitutions and homoplasy in this mutation class. As the transitions always remained higher than transversions, instead of simply ruling out those substitutions, we decided to use more complex evolutionary models which could account for multiple substitutions and homoplasy as well as for the transition/transversion ratio and the G + C content.

The Tamura 3-Parameter (Tamura, 1992) model was chosen for the NJ analysis. In this case, the distances obtained varied from 0.00 (for same species populations) to 0.345 (between *D. pseudoobscura* and *S. leonensis*) with a mean of 0.227. As in most comparisons the assigned values did not exceed the 0.30 threshold (suggested by Russo (2001) as the limit above which the differences are excessively underestimated), we chose not to use the amino acid sequences. The general time reversible (GTR) model using the ratio of invariable sites (I)



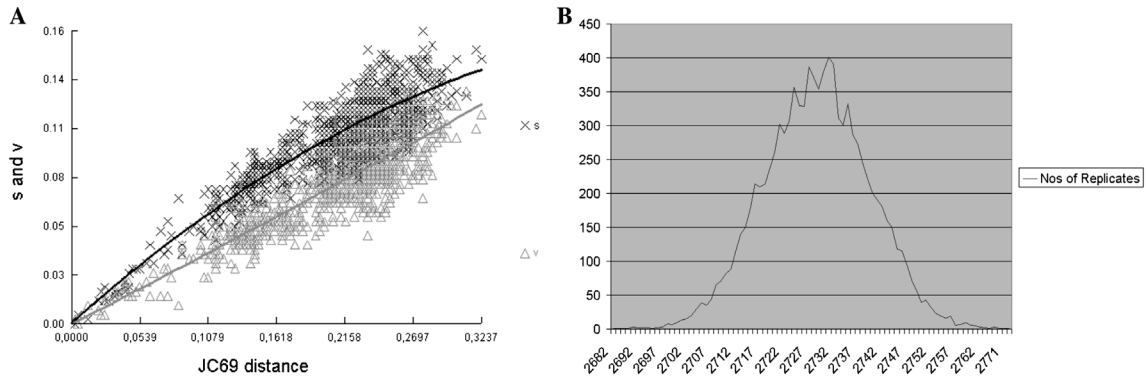


Fig. 2. Graphs showing: (A) *Amd* transition (s) and transversion (v) rates plotted against the JC distances (Jukes and Cantor, 1969) for different *Drosophilidae* species; (B) distribution of the scores of the randomly generated maximum parsimony (MP) trees constructed using *Amd* sequences. The score of the most parsimonious *Amd* tree was 1730.

and the gamma distribution of the variable sites (G) was the model indicated by the AIC test for likelihood analyses. For the MP analysis the length of the best tree (1730) laid well outside the scores presented by the permuted data (Fig. 2B) in the PTP test ( $p = 0.0001$ ), attesting to the strong *Amd* phylogenetic signal. This tree presented a consistency index (ci) of 0.287 and a retention index (ri) of 0.582.

Bayesian analysis constructed 1,000,000 trees of which 6696 different trees were finally considered in the construction of the consensus tree (Fig. 3) as well as in the estimation of the posterior probabilities of each clade. The phylogenies obtained by the other methods yielded essentially similar results, since both the Templeton and the Shimodaira–Hasegawa tests detected no statistically significant differences between any of them. It was particularly interesting to see that our Bayesian topology performed better than the ML one under the likelihood criterion of the Shimodaira–Hasegawa test, although this difference was not statistically significant ( $p = 0.572$ ). As compared with the 44 clades of the Bayesian tree the ML tree was 91% similar (40 out of 44 clades), the MP tree 82% similar (36 out of 44 clades), and the NJ tree 79% similar (35 out of 44 clades) (data not shown). So it appears that although some clusterings differ among the different trees (particularly those of more inclusive levels, especially within the *immigrans*–*Hirtodrosophila* radiation) there are only a few divergent clades among them. This reiteration among different inference methods can be taken as an indication of the strong internal consistency and congruency of the data.

When *Scaptodrosophila latifasciaeformis* was used as the outgroup the subgenus *Sophophora* appeared at the base of the genus *Drosophila* phylogeny, where it was preceded only by the subgenus *Dorsilopha* (Fig. 3), here represented by *D. busckii*, which composed a basal polytomy with *S. latifasciaeformis*. This early offshoot of the subgenus *Sophophora* agrees with a series of previous studies (Kwiatowski and Ayala, 1999; Remsen and

O’Grady, 2002; Russo et al., 1995; Tatarenkov et al., 1999, 2001; Throckmorton, 1975) but contrasts with the findings of Grimaldi (1990) and DeSalle (1992). On the other hand, the basal position of *Dorsilopha* was unexpected because most studies have placed this subgenus among the derived clades inside the genus *Drosophila* (Grimaldi, 1990; Kwiatowski and Ayala, 1999; Tatarenkov et al., 1999, 2001; Throckmorton, 1975).

The subgenus *Drosophila* was the next to branch off and appeared to be phylogenetically closer to other *Drosophilidae* genera (*Zaprionus*, *Samoia*, *Hirtodrosophila*, and *Liodrosophila*) than to the subgenus *Sophophora*. The subgenus *Drosophila* was also paraphyletically defined with respect to these taxa, which at least partially agrees with other published work (Kwiatowski and Ayala, 1999; Tatarenkov et al., 1999, 2001; Throckmorton, 1975; Yotoko et al., 2003). In contrast with these, Remsen and O’Grady (2002) presented *Hirtodrosophila*, *Samoia*, and *Liodrosophila* as basal clades within *Drosophilidae*, a result that is also supported by Grimaldi’s analysis (Grimaldi, 1990), where *Zaprionus* also appeared as sister to a monophyletic *Drosophila* genus.

The subgenus *Drosophila* was further subdivided into two main clades, one composed of members of the *virilis-repleta* radiation with a constitution essentially similar to that presented by Throckmorton (1975) and another composed of the traditional *immigrans*–*Hirtodrosophila* radiation of Throckmorton (1975) plus the species *L. aerea* and *D. funebris*. Nevertheless only the *virilis-repleta* cluster presented good support values (1.00).

As regards the *virilis-repleta* radiation, the confirmation of the *mesophragmatica* group monophyly is important given its high confidence (support value = 1.00). This group was internally subdivided into three main clades, one consisting of *D. gasici*, *D. brncici*, and *D. mesophragmatica*, another composed of the sibling species *D. pavani* and *D. gaucha* and a third basal

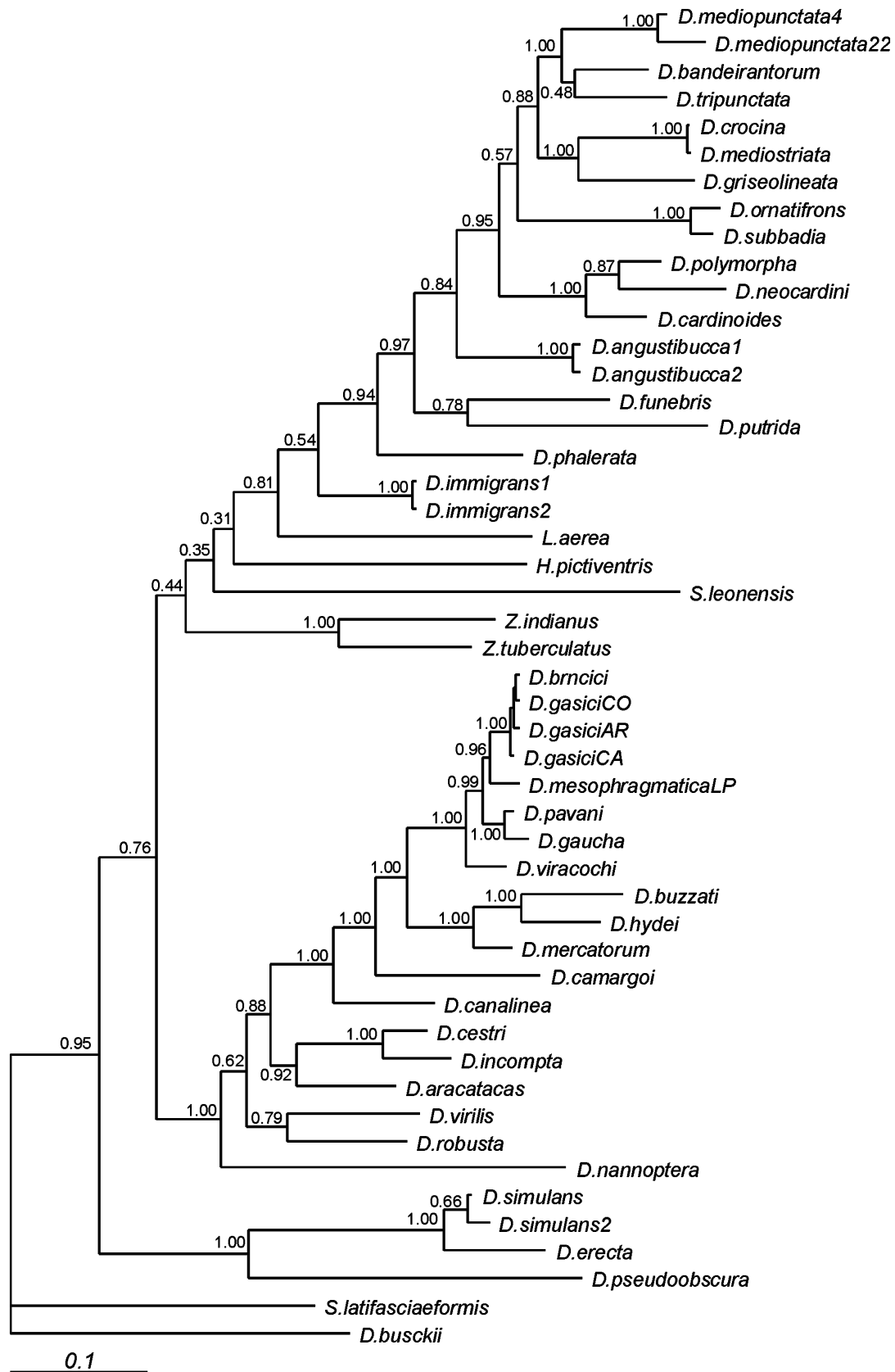


Fig. 3. The *Amd* consensus phylogeny obtained using Bayesian analysis and the GTR + I + G model as proposed by the AIC test for the 49 nucleotide sequences. The posterior probability of each clade is indicated beside its respective internal branch.

and monotypic clade containing *D. viracochi*. Such grouping is widely consistent with the subgroups obtained by Brncic et al. (1971) through cytogenetic analysis, as is also the case for the paraphyly of *D. gasici*, which was probably due to high ancestral polymorphism. The *repleta* group was the *mesophragmatica* group sister-clade, which partially agrees with some previous published work (Durando et al., 2000; Tatarenkov and Ayala, 2001) but which is at variance with the findings of Throckmorton (1975) and Remsen and O'Grady (2002). To this *mesophragmaticarepleta* clade it follows the basal branching of *D. camargoi* (*dreyfusi* group) and *D. canalinea* (*canalinea* group), defining a monophyletic *repleta* radiation as previously also detected by Tatarenkov and Ayala (2001). Unexpectedly, the clade formed by the *flavopilosa* species group and *D. aracatacas* (*annulimana* species group) clustered as the sister-clade of the *repleta* radiation. Throckmorton (1975) placed these groups as basal to the other *virilis-repleta* radiation members but they occupy a more derived position in this work. *D. nannoptera* (*nannoptera* group) and a clade made up of *D. virilis* (*virilis* group) and *D. robusta* (*robusta* group) were actually the first split inside the *virilis-repleta* radiation, although the first of these branching-points was only weakly supported (support value = 0.62).

Within the *immigrans-Hirtodrosophila* radiation the genus *Zaprionus* (support value = 0.35) as well as the species *S. leonensis* (0.31), *H. pictiventris* (0.81), and *L. aerea* (0.54) constitute the early offshoot, although the values in parentheses show that the statistical support for this was low. The next to branch off was *D. immigrans* (*immigrans* group), another taxon indicated by Throckmorton as composing the basis of this radiation, which is followed by *D. phalerata* (*quinaria* group) and by the clade joining *D. putrida* (*testacea* group) and *D. funebris* (*funebris* group). This last group, together with *Liodrosophila*, was proposed by Throckmorton (1975) as being basal as compared to the *virilis-repleta* and *immigrans-Hirtodrosophila* radiations, but this was not confirmed either by our work or by other published studies (Remsen and O'Grady, 2002; Tatarenkov et al., 2001; Yotoko et al., 2003), especially concerning the position of *D. funebris*. Remaining inside this major clade representing the *immigrans-Hirtodrosophila* radiation there was a monophyletic *tripunctata* radiation composed of a *tripunctata* group entirely paraphyletic in respect to the *cardini*, *guarani*, and *guaramunu* species groups (support value = 0.84). This paraphyly of the *tripunctata* group has previously been noted by several other workers (Frota-Pessoa, 1954; Throckmorton, 1975; Yotoko et al., 2003), although Yotoko et al. (2003) pointed out that the *tripunctata* radiation was also paraphyletic concerning the *funebris* and *testacea* species groups, i.e., the clades which make up the *tripunctata* radiation sister-group in our present study.

It was possible to subdivide the *tripunctata* radiation cluster into five main clades which we list here in an ascending order beginning with the most primitive: monotypic clade 1 containing *D. angustibucca* a member of *tripunctata* group subgroup I (Frota-Pessoa, 1954); clade 2 presenting a monophyletic *cardini* group whose internal composition agrees with the clusters proposed for the *cardini* subgroup by Heed and Russel (1971); clade 3 containing the two *guarani* group species, *D. ornatifrons* and *D. subbadia*; clade 4 made up of the *guaramunu* group species *D. griseolineata* along with *D. mediotriata* (synonym *D. crocina*, according to Vilela and Bächli, 1990) which is a member of subgroup III of the *tripunctata* group; clade 5 containing three members of the *tripunctata* group, *D. mediopunctata* (subgroup II), *D. bandeirantorum* (subgroup III), and *D. tripunctata* (subgroup IV). The relationships between the five clusters described above present a good overall resolution for the Bayesian tree, although for the remainder analysis performed the bootstrap values obtained for them were generally very low (results not shown). As can be seen by the arrangement of the subgroups and groups above, it was not generally possible to detect similarities between the morphological groupings obtained by Frota-Pessoa (1954) and the internal composition of the *tripunctata* radiation cluster presented here.

### 3.2. Individual analysis of the COII gene

For the *COII* gene the multiple alignment of the 71 sequences showed 330 (of 672 bp) variable sites of which 241 were parsimoniously informative. The medium transition/transversion ratio was 1.0 but this increased to 6.5 when only first codon positions were considered. The frequencies of the four nucleotides T, C, A, and G were also significantly biased, presenting an average of 39.9, 14.2, 32.4, and 13.5, respectively. This bias toward A + T content increased even more when only the third codon positions were maintained reaching an average of 90.6. The saturation test (Fig. 4A) revealed that, in contrast to our results for the *Amd* gene, the transition/transversion ratios for *COII* decreased significantly with the increasing of the distance, a clear sign of fast saturation of transitional substitutions due to the strong biases in both base composition and substitution patterns. However, because we wanted to solve some important questions at the intra-group level where this mutation class remains informative we decided to adopt an equal weight strategy inside a complex evolutionary model.

The Tamura–Nei model (Tamura and Nei, 1993) with a gamma distribution was chosen here for the NJ analysis. In this case, all the distances obtained varied from 0.00 (for populations of the same species but also between *D. brncici*, *D. gasiciAR*, and *D. gasiciCO*) to 0.247 (between *D. nannoptera* and *D. mediopunctata*), with a mean of 0.15. The AIC test indicated that the

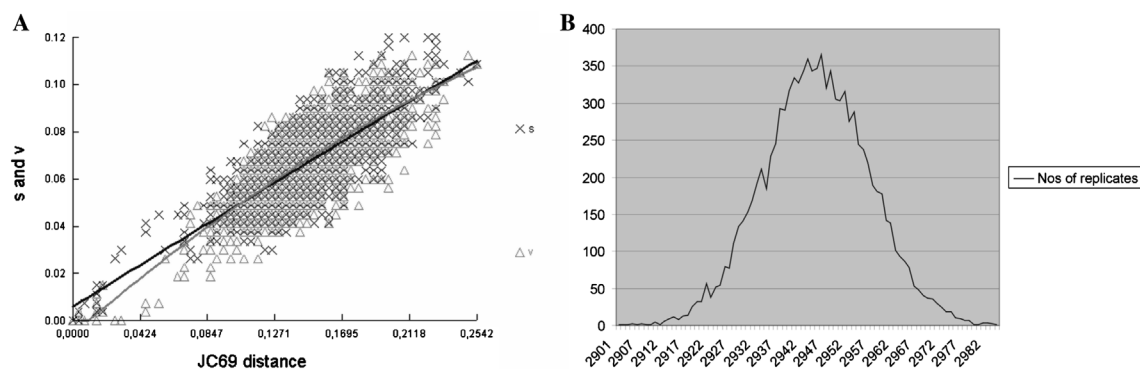


Fig. 4. Graphs showing: (A) *COII* transition (s) and transversion (v) rates plotted against the JC distances (Jukes and Cantor, 1969) for different *Drosophilidae* species; (B) distribution of the scores of the randomly generated MP trees constructed using *COII* sequences. The score of the most parsimonious *COII* tree was 1937.

GTR + I + G model was again the best for the likelihood analyses. The MP analysis produced a best tree whose length (1937) laid outside the scores presented by the permuted data (Fig. 4B) in the PTP test ( $p=0.0001$ ), which also attested to the strong *COII* phylogenetic signal. This tree presented a ci of 0.256 and a ri of 0.568.

The Bayesian analysis for *COII* evaluated 2,000,000 generations from which 15,243 different trees were used to construct the 50% majority rule consensus tree (Fig. 5). The phylogeny constructed by the ML method was 87.5% similar to the Bayesian tree, but only 59% of the Bayesian tree clades were also supported by the MP analysis (38 out of 64) and 64% by the NJ analysis (41 out of 64). The Templeton test performed under a parsimony criterion regarded the MP hypothesis as significantly better than the others, while the Shimodaira–Hasegawa test performed under a likelihood criterion detected no significant differences between the trees. In general, those clusters involving inter-species relationships among members of the *immigrans-tripunctata* radiation were the major source of this inconsistency, these clades also being only weakly supported by the Bayesian posterior probabilities.

As the *COII* gene is a rapidly evolving mitochondrial locus it produced some unexpected results at the species group-level analysis (Fig. 5). This is the case, for example, for the polyphyletic *repleta* group and for the “floating” placement of *D. nanoptera*, *D. simulans*, and *D. yakuba*, probably led by the phenomenon of long branch attraction, given the high saturation presented by this gene. The *COII* gene was also unable to detect the monophyly of either the *flavopilosa* group or the *repleta virilis-repleta* radiations, although it presented a monophyletic *mesophragmatica* group, whose sister-group relationship with the *annulimana* group was not previously reported. The *mesophragmatica* group internal branching pattern was different from that obtained by *Amd* in that *D. viracochi* clustered with *D. brncici*, *D. gasici*, and *D. mesophragmatica* instead of being the

early offshoot inside its group. This result also disagrees with the studies of Nacur (1958) and Brncic et al. (1971) but it is very similar to the relationships presented by Nair et al. (1971) using isozyme variation analysis, except that our study branched *D. mesophragmatica* earlier than *D. viracochi*.

The most derived part of the *COII* tree again consisted of representatives of the *immigrans-Hirtodrosophila* radiation plus *D. funebris* and *L. aerea*, whose most general constitution was similar to that presented by *Amd*. The branching pattern among the subgenus *Dorsilopha* and the genera *Hirtodrosophila*, *Samoaia*, *Liodrosophila*, and *Zaprionus* remained obscure, given its weak confidence values, but these taxa appeared to be the early offshoot inside that major clade, leaving, again, the genus and subgenus *Drosophila* entirely paraphyletic. After these splitting, all the remaining species belonged to the subgenus *Drosophila* (support value = 0.54) but their branching pattern was also somewhat confusing. We found that the *funebris* and *quinaria* groups split off first, leaving a clade containing species of the *tripunctata* radiation plus *D. nanoptera*, *D. immigrans* (*immigrans* group), and *D. putrida* (*testacea* group) with a very weak confidence value (0.21). Thus, in contrast to the results produced with the *Amd* gene, the *tripunctata* radiation was here paraphyletic concerning these three species, a result that agrees with the findings of Yotoko et al. (2003).

Inside this major cluster that composes the *immigrans-Hirtodrosophila* radiation there were six main clades, but their external relationships do not have any statistical support. These clades are listed below in an ascendant order, beginning with the most primitive one:

- Clade 1 grouped *D. immigrans*, *D. pallidipennis*, and *D. nanoptera* (support value = 0.28) presenting them as basal inside the *tripunctata* radiation. This composes the sister-clade of a paraphyletic *tripunctata* group, presented on the remaining clades.

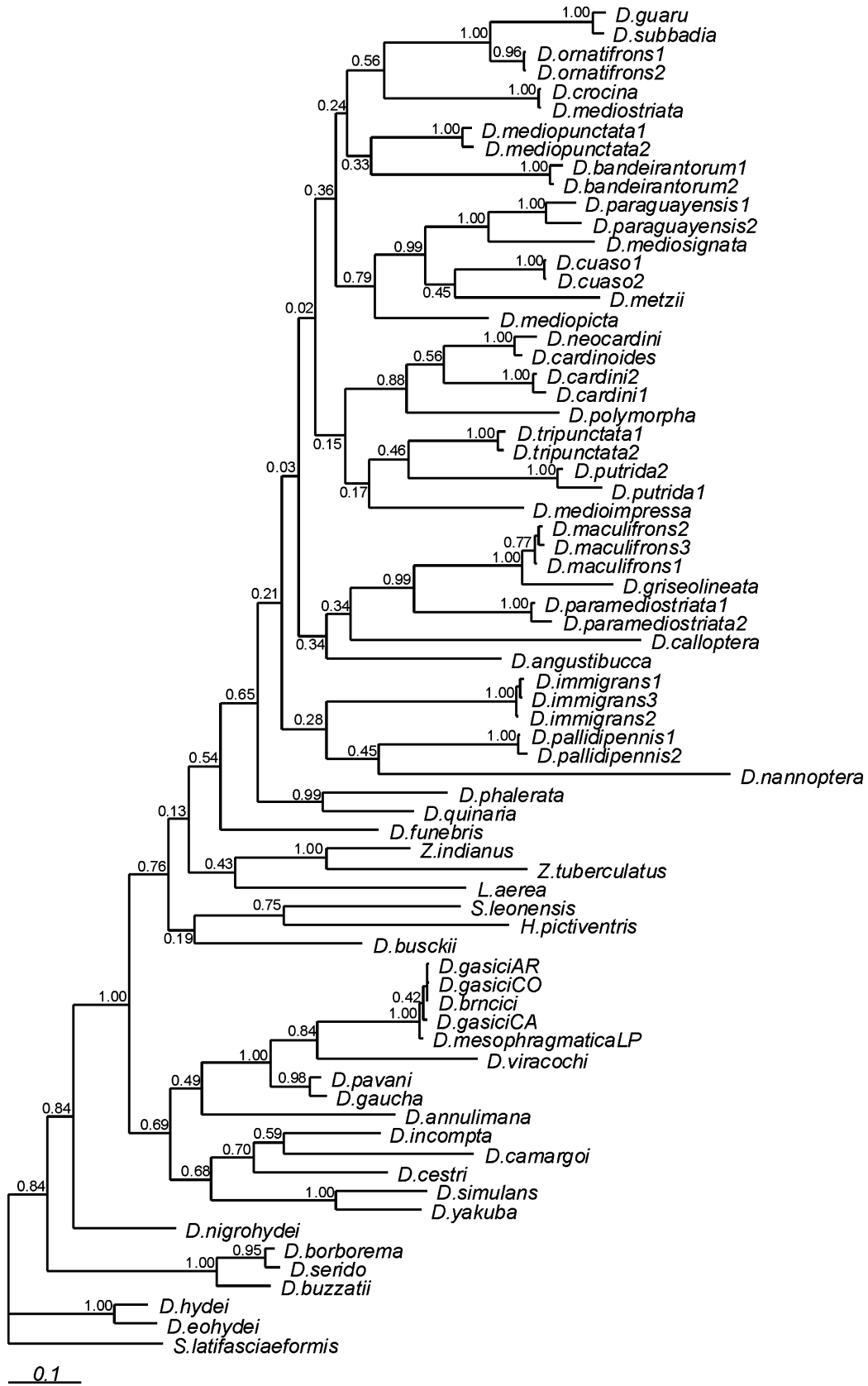


Fig. 5. The *COII* consensus phylogeny obtained using Bayesian analysis and the GTR+I+G model, as proposed by the AIC test for the 71 nucleotide sequences. The posterior probability of each clade is indicated beside its respective internal branch.

- Clade 2, partially congruent with Yotoko et al. (2003), grouped a monophyletic *guaramunu* group with *D. paramediostriata*, a member of subgroup III of the *tripunctata* group (Frota-Pessoa, 1954) (support value = 0.99) and these with *D. calloptera* (*calloptera* group) (support value = 0.34). *D. angustibuca* (*tripunctata* group, subgroup I) constituted the early offshoot inside this clade (support value = 0.34).
- Clade 3 presented the cluster of *D. putrida* (*testacea* group), *D. tripunctata* (*tripunctata* group, subgroup IV), and *D. medioimpressa* (*tripunctata* group, subgroup II) (support value = 0.17) as the sister-clade of a monophyletic *cardini* group (support value = 0.88), with a confidence value of 0.15. The *cardini* group had internal relationships which entirely conflicted with that presented by Heed and Russel (1971) and by our individual analysis of the *Amd* gene because *COII* clustered *D. neocardini* with *D. cardinoides* and *D. cardini*, leaving *D. polymorpha* on a separate branch (probability of 0.56).
- Clade 4 showed better support values joining only species of the *tripunctata* group. It presented *D. mediopicta* (subgroup III) as the sister of a cluster previously reported by Yotoko et al. (2003) that grouped *D. paraguayensis* and *D. mediosignata* (subgroup II) with *D. cuaso* (subgroup II) and *D. metzii* (subgroup IV). Bächli et al. (2000) described *D. paraguayensis*, *D. mediosignata*, and *D. cuaso* as three cryptic species.
- Clade 5 was also produced by the *Amd* analysis and clustered two representatives of the *tripunctata* group, *D. mediopunctata* (subgroup II) and *D. bandeirantorum* (subgroup III), but with no statistical support (support value = 0.33).
- Clade 6, again with a low support value (0.56), produced a new result by presenting *D. mediostriata* (*tripunctata* group, subgroup III) as the sister species of a monophyletic *guarani* group which clustered *D. guaru* with *D. subbadia* and these with *D. ornatifrons* (support value = 1.00).

### 3.3. Simultaneous analysis of *Amd* and *COII*

The PHT test performed to evaluate the congruence between the *Amd* and *COII* data partitions detected a significant conflict between these two datasets, as indicated by the *p* value obtained (0.001). Even so we decided to combine the two partitions, since that previous analyses made by Remsen and DeSalle (1998) and Remsen and O'Grady (2002) showed that interesting results can be produced from simultaneous analysis even if the constituent data partitions display heterogeneity.

For this new dataset we included only the 41 taxa for which the two gene sequences were available. The total number of characters analyzed was 1188 bp (516 from *Amd* and 672 from *COII*). The Tamura–Nei

model with a gamma distribution was again chosen for the NJ analysis and the GTR + I + G model was again indicated by the AIC test for use in the likelihood analyses.

Bayesian analysis evaluated 2,000,000 generations from which 4111 different trees were considered in the production of the consensus tree (Fig. 6). The low number of different trees obtained is a good indicator that the 'top of the hill' was reached. Bayesian analysis again performed better than the ML one under the likelihood criterion of the Shimodaira–Hasegawa test, although the difference was not statistically significant ( $p = 0.683$ ). The phylogenies achieved by the other methods were also overall very similar to the Bayesian topology, since neither the Templeton test nor the Shimodaira–Hasegawa test detected significant differences between them. Table 2 presents the confidence values for the Bayesian clades numbered in Fig. 6 under the different phylogenetic reconstruction methods. In general, although many of the clusters were not statistically significant (particularly those of more inclusive level) there were only a few divergent clades among the different trees, most of which again referred to the *immigrans*–*Hirtodrosophila* inter-group relationships. Nevertheless, the reinforcement of results obtained by different inference methods can be taken as an indication of the strong internal consistency and congruency of the data overall.

As with the *Amd* analysis, following the separation of *S. latifasciaeformis* and *D. busckii*, *D. simulans* emerged as the sister-clade of all the remaining *Drosophilidae* species (Fig. 6). This early offshoot of the subgenus *Sophophora* left a paraphyletic subgenus *Drosophila*, phylogenetically closer to the genera *Hirtodrosophila*, *Samoia*, *Zaprionus*, and *Liodrosophila* than to the subgenus *Sophophora*. Our total evidence tree therefore clearly showed a well-supported (probability of 0.97) paraphyletic subgenus *Drosophila* in accordance with the *Amd* data and several studies by other authors (Kwiatowski and Ayala, 1999; Remsen and O'Grady, 2002; Russo et al., 1995; Tatarenkov et al., 1999, 2001; Thomas and Hunt, 1993; Throckmorton, 1975; Yotoko et al., 2003).

There was also strong support for the *immigrans*–*Hirtodrosophila* (support value = 0.85) and *virilis-repleta* (support value = 0.96) radiations within the subgenus *Drosophila*, although the first of these clusters did not display the same composition as that presented by Throckmorton (1975) because of the inclusion of *D. funebris* and *L. aerea*. This last result was also noted by Tatarenkov et al. (2001) concerning the two species and by Remsen and O'Grady (2002) and Yotoko et al. (2003) concerning only *D. funebris*. The *virilis-repleta* radiation branching pattern was extremely similar to that obtained with the *Amd* gene individual analysis.

In the major *immigrans*–*tripunctata* radiation cluster, a clade joining the two *Zaprionus* species with *L. aerea*

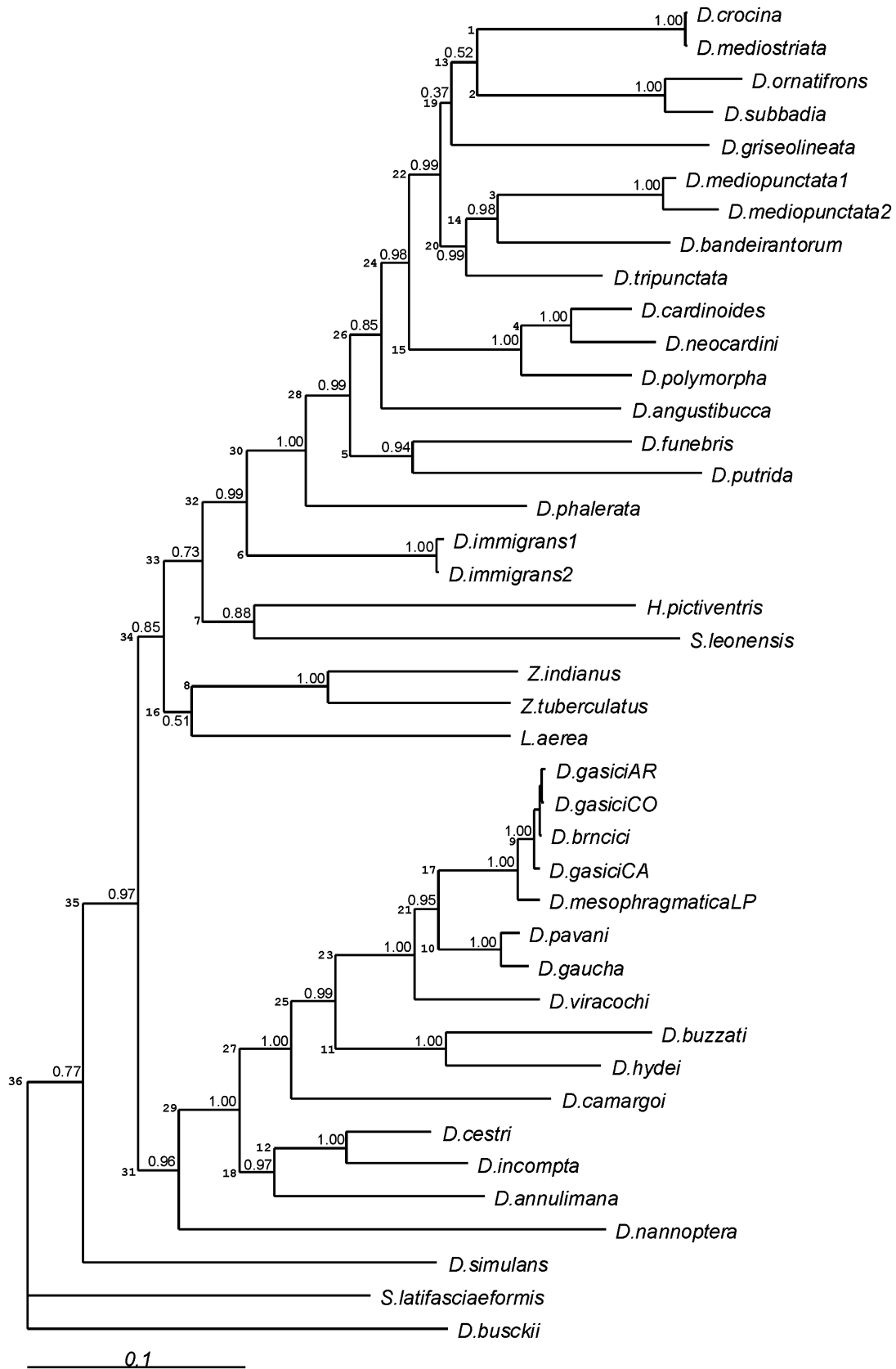


Fig. 6. Total evidence consensus phylogeny obtained using Bayesian analysis and the GTR + I + G model, as proposed by the AIC test for the 41 nucleotide sequences. The posterior probability of each clade is indicated beside its respective internal branch. At the front of the internal branches there is an arbitrarily defined number representing the clade that follows and which can be used to interpret Tables 2 and 3.

Table 2

Bootstrap values presented for each of the clades shown in Fig. 6 by the trees constructed using different phylogenetic reconstruction methods

Clade	Phylogenetic reconstruction method		
	NJ	MP	ML
1	100	100	100
2	100	100	100
3	100	100	100
4	77	91	93
5	42	33	—
6	100	100	100
7	—	—	27
8	100	98	99
9	99	98	100
10	100	100	99
11	98	96	98
12	100	97	98
13	—	—	—
14	38	38	49
15	99	99	99
16	—	—	—
17	100	100	100
18	58	61	53
19	—	—	—
20	40	17	29
21	—	52	71
22	11	—	18
23	100	100	99
24	—	—	16
25	70	48	65
26	36	39	22
27	50	40	63
28	27	39	48
29	83	50	71
30	50	57	55
31	67	26	45
32	37	35	44
33	—	—	—
34	—	—	11
35	—	—	18
36	32	37	36

constituted an early split (support value = 0.73), followed by the clade grouping *H. pictiventris* and *S. leonensis* (support value = 0.99). These species also constituted the early offshoot of the *immigrans-Hirtodrosophila* radiation in our individual *Amd* and *COII* analyses. Similar results have also been reported by Throckmorton (1975), by Remsen and DeSalle (1998) nuclear combined analysis for *Zaprionus* and *Hirtodrosophila*, by Tatarenkov et al. (2001) for *Liodrosophila*, *Zaprionus*, and *Samoaia*, and by Yotoko et al. (2003) for *Zaprionus* only. Other studies, however, have pointed to these genera as being outgroups for all *Drosophila* subgenera other than *Sophophora* (Russo et al., 1995; Tatarenkov et al., 1999) and even as the outgroups to the genus *Drosophila* (Grimaldi, 1990; DeSalle, 1992 concerning *Zaprionus* and *Hirtodrosophila*; Remsen and O'Grady, 2002 concerning *Samoaia* and *Hirtodrosophila*; Yotoko et al., 2003 concerning *Hirtodrosophila*).

Therefore, more studies are necessary in order to reach a more reliable phylogeny.

After the separation of these taxa, all the remaining species belong to the subgenus *Drosophila*. In this combined analysis, the basal split order was essentially similar to that of the individual *Amd* analysis, with *D. immigrans* branching off first followed by *D. phalerata* and by the *D. putrida/D. funebris* clade. In this analysis the *tripunctata* radiation was also monophyletic and had a good confidence value (0.85).

The *tripunctata* radiation cluster produced by our total evidence data was subdivided into four main clades, whose differences with those clades obtained individually by *Amd* referred to the internal composition of the *cardini* group, which more closely resembled that obtained with *COII* analysis, and the placement of the *guarani* clade inside the clustering of *D. griseolineata* and *D. mediotriata*, although with a low support value (0.37). This *D. mediotriata/guarani/D. griseolineata* cluster can be assigned to the mixture of the *COII* gene information joining *D. mediotriata* with the *guarani* group and the *Amd* gene signal grouping *D. mediotriata* with *D. griseolineata*. The branching order and inter-clade relationships nevertheless remained essentially similar to that found for *Amd* individually.

To assess the relative contribution of the two gene regions to our simultaneous analysis (SA) hypothesis the PBS scores for each gene on each node were calculated. This test only confirmed the pattern described above, i.e., the *Amd* gene provided almost twice the *COII* support and contributed much more to our simultaneous analysis (Table 3). The breakdown of the PBS values also indicated substantial conflict between these two partitions, because 11 of the 36 nodes resolved in our SA consensus tree had conflicting PBS values. This pattern may have been the result of either differences in internal homoplasy within each gene or conflicting signals between the genes.

The different MP and Bayesian topologies inferred using the three distinct datasets (reduced by the number of superposed taxa) were also statistically tested among each other by the Templeton and Shimodaira–Hasegawa tests. Overall, the *Amd* gene data did not reject the SA trees while the *COII* topologies were statistically worse than the former under all tests performed. Otherwise, the *Amd* and SA trees were significantly rejected by the *COII* partition. Under the combined dataset only the *COII* topologies were significantly rejected by both tests, while the *Amd* topologies were not statistically different from the SA topologies. Therefore, the phylogenetic trees based on the *COII* gene were again very inconsistent with those based on the *Amd* gene and on the combined dataset, but, overall, the *Amd* and SA topologies appear not to have been significantly different.



Table 3  
Partitioned Bremer support (PBS) scores across the simultaneous analysis Bayesian tree for each gene partition in each of the clades presented in Fig. 6

Clade	<i>Amd</i>	<i>COII</i>
1	20.2	24.8
2	15	12
3	12	18
4	-2	10
5	0	3
6	14.2	31.8
7	0	0
8	15.7	3.3
9	6	0
10	5	11
11	15.7	-1.7
12	14	-4
13	0	0
14	1	-1
15	8	3
16	0	0
17	2	15
18	1.6	4.4
19	0	0
20	1	-1
21	1.6	0.4
22	0	0
23	8	4
24	0	0
25	11	-7
26	3.3	1.7
27	11	-7
28	-2	7
29	19	-13
30	7	0
31	0	0
32	-3	8
33	0	0
34	0	0
35	0	0
36	5.7	-0.7
Total PBS	191	122
Min steps	497	496
Total PBS/Min steps <sup>a</sup>	<b>0.38</b>	<b>0.24</b>

<sup>a</sup> PBS values summed across the tree and standardized by the minimum number of steps for each partition.

### 3.4. *Amd* and *COII* molecular evolution patterns

Given the detected conflict between the *Amd* and *COII* partitions, it is important to identify the source of this disagreement. By using the Modeltest results and performing Bayesian analysis using the GTR + site-specific rates (SSR) model it was possible to quantitatively compare the substitution patterns presented by the nuclear *Amd* and the mitochondrial *COII* genes. It is well known that nuclear genes generally evolve more slowly than mitochondrial genes, making them better markers for deep divergences (Simon et al., 1994). However, when we consider our data alone this pattern is not seen because the relative rate presented by *Amd* (1.11)

was somewhat greater than that presented by *COII* (0.91). A better understanding of rate variation can be obtained by looking at the rates among codon positions within genes, from which it can be seen that the first, second, and third codon positions of the *Amd* gene evolved faster than the *COII* codon positions (Fig. 7A). Indeed, this pattern could already be expected because the median distance values obtained for *Amd* (0.227) were greater than those for *COII* (0.15). It should not be forgotten, however, that these results could be a by-product of the saturation presented by our mitochondrial gene.

It was previously seen that the *COII* gene showed greater base compositional bias than the *Amd* gene, but other patterns of nucleotide substitution, such as the Q matrix of transformation, may also be important. The instantaneous rate matrix for *COII* was more asymmetrical relative to that presented by *Amd* and was also more skewed towards one type of change over another (Fig. 7B), although in both cases there was a higher overall rate of transitions. Another contrasting point that can be seen in Fig. 7B is that while the *Amd* nuclear gene showed very similar overall rates of transversions, the *COII* mitochondrial dataset had a rate of CG transversions which was almost 14 times higher than the rate of AT transversions. The obvious consequence of this highly skewed transformation rate matrix is greater levels of homoplasy, which are not easily corrected for.

Another parameter that often differs between nuclear and mitochondrial genes is the shape of the gamma distribution, as given by the  $\alpha$  value, describing the among-site rate variation. As can be seen from Fig. 7C, the *Amd* gene showed a more homogeneous pattern of among-site rate variation as compared to the *COII* gene, which had lower  $\alpha$  values. The *Amd* gene also presented a higher proportion of invariable sites (Pi values), a parameter which is positively correlated with  $\alpha$  in coding regions according to Lin and Danforth (2004). These authors also showed a positive correlation between  $\alpha$  and ci, the consistency index, which suggest that data partitions with more heterogeneous substitution rates show a higher level of homoplasy. This correlation was also detected by us, with the *Amd* gene having a higher ci value (0.287) than the *COII* gene (0.256).

The combination of all these factors may together explain the overall poor performance of the mitochondrial *COII* gene relative to the nuclear *Amd* gene, given that all these properties should lead to high levels of homoplasy and, therefore, saturation. The choice of complex models that account for these biased substitution patterns appeared not to solve this problem.

## 4. Conclusion

Our study reinforces the work previously published by Goto and Kimura (2001) and Lin and Danforth (2004) in

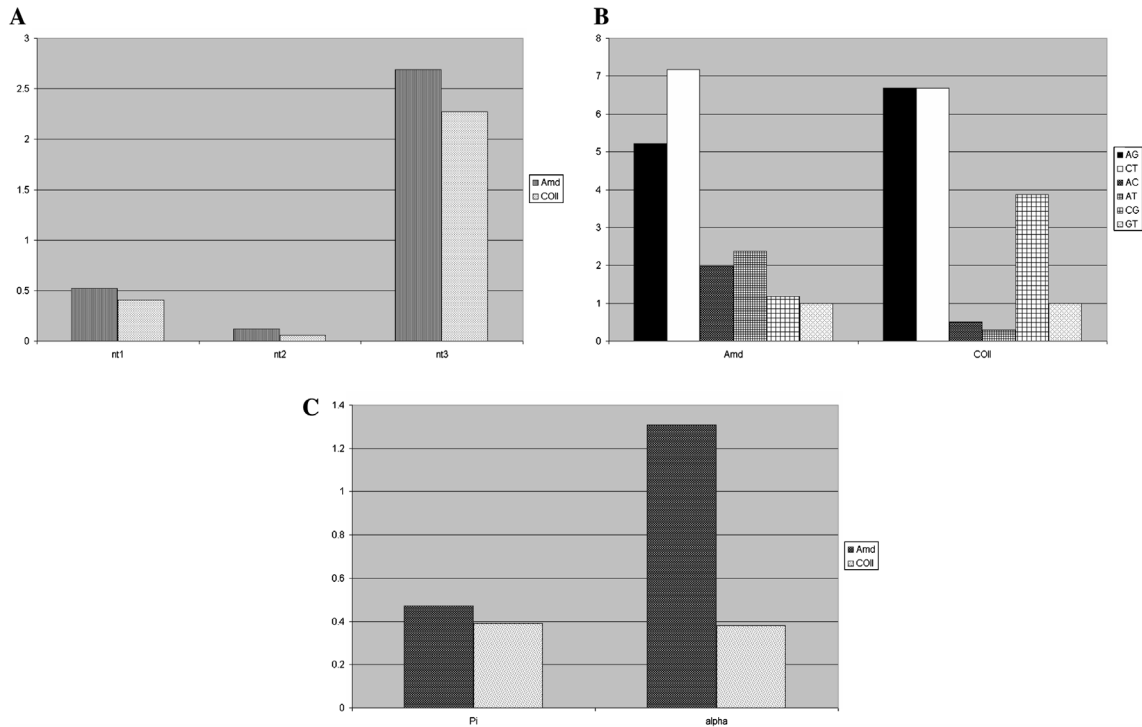


Fig. 7. The *Amd* and *COII* nucleotide substitution patterns: (A) relative rates of nucleotide substitution based on the GTR + SSR model, with sites partitioned by gene and by codon position; (B) *Amd* and *COII* transformation rate matrices obtained with AIC test performed on ModelTest; (C) *Amd* and *COII* gamma shape and proportion of invariable sites (Pi) obtained with AIC Test.

that it shows that nuclear genes have a slight advantage over mitochondrial genes when used for phylogenetic analyses, especially these of more inclusive levels. The *Amd* gene had higher consistency index (ci) values and almost twice the partition Bremer support (PBS) scores as compared to the *COII* gene. The partition homogeneity test (PHT) revealed significant incongruence between *Amd* and *COII* and the Templeton and the Shimodaira–Hasegawa tests indicated that the maximum parsimony (MP) and the Bayesian trees (the two types of trees tested) constructed using the *COII* gene were significantly rejected by the *Amd* and the combined dataset. In addition, the *COII* support values were rather lower than the *Amd* support values, especially for between-group comparisons. The poor phylogenetic performance of the *COII* gene can be attributed to its strong A/T bias associated with its highly asymmetrical rate matrices and high among-site rate heterogeneity. These properties appear to have resulted in the *COII* mitochondrial gene having high levels of homoplasy which are not easily corrected for by the use of complex evolutionary models. Lin and Danforth (2004) state that the properties that characterize genes suitable for phylogenetic studies appear to be the opposite to those which characterize the *COII* gene, i.e., genes better suited for phylogenetic analysis have a more even base composition, higher  $\alpha$  values, and a less skewed transformation rate matrix. Otherwise, although some saturation occurs in the *Amd* gene, at least as far as transitions are concerned, the phylogenetic

utility of this gene can be recovered by the use of complex evolutionary models.

In our study, the phylogenetic trees based on the *COII* gene were significantly different from those based on the *Amd* gene and the combined dataset and we consider our total evidence tree (TET) (Fig. 6) as the best phylogenetic hypothesis to now available for these species. Except for minor differences, our TET was essentially similar to the phylogenetic scheme produced by Throckmorton (1975). In this sense, it proposes a paraphyletic genus *Drosophila* concerning the genera *Hirtodrosophila*, *Liodrosophila*, *Zaprionus*, and *Samoia* and a paraphyletic subgenus *Drosophila* concerning these taxa but not the subgenus *Dorsilopha* which constitutes a basal polytomy with *S. latifasciaeformis*. The subgenus *Drosophila* also constitutes a derived split within the genus *Drosophila*, while the subgenus *Sophophora* is the early offshoot. The major division of species in the subgenus *Drosophila* species occurs between the *virilis-repleta* and *immigrans-Hirtodrosophila* radiations proposed by Throckmorton (1975), but in our study the last radiation yielded a somewhat different composition because it included *Liodrosophila* and *D. funebris* as basal species within the *immigrans-Hirtodrosophila* radiation while Throckmorton (1975) assigned them as basal species within the subgenus *Drosophila*. The paraphyly of the *tripunctata* group pointed out by various authors (Frota-Pessoa, 1954; Throckmorton, 1975; Yotoko et al., 2003) was confirmed by us, although we found that the

*tripunctata* radiation was monophyletic. Even so it was not possible to find more similarities between the morphological subdivision of Frota-Pessoa (1954) and the internal composition of our *tripunctata* radiation. This was also the most inconsistent region of our trees, with a low confidence level and incongruent results in regard to the different phylogenetic reconstruction methods and to the different datasets. A possible reason for this is the fact that the internal branches concerning the *tripunctata* radiation internal relationships were overall short. All these patterns support the scenario originally presented by Throckmorton (1975) and also stated by Yotoko et al. (2003) in which these taxa emerged by means of fast and multiple substitution episodes. If these results come to be confirmed, a large set of combined and informative datasets need to be accumulated in order to reach a more reliable phylogeny for these species.

Finally, it is worth reiterating that, taxonomically, the family Drosophilidae remains an agglomeration of merophyletic taxa (merophyly designates the absence of monophyly according to Bernardi (1981)), beginning with the genus *Drosophila* and its most diverse subgenus, the subgenus *Drosophila*. As our results show, genera emerging from subgenera are a normal feature of the taxonomy of these taxa. Nevertheless, despite the accumulation of evidence, such as that presented here, few effective attempts have been made to review the nomenclature of the family and reconcile its taxonomy with recent systematic studies.

## Acknowledgments

We are grateful to the following people: Marco Silva Gottschalk (M.Sc.) for kind collaboration on the identification of the species cited in this paper; Liliana Essi (M.Sc.) and Paulo Marcos Pinto (M.Sc.) for the valuable help in difficult moments; Marcos de Oliveira Carvalho for assistance with bioinformatics; Adriana Ludwig for providing the *flavopilosa* group DNA; Dr. Robert Thomas and Mr. Félix Nonnemacher for revision of English. We also thank the Brazilian agencies CNPq for providing grants and fellowships, and CNPq, FAPERGS and PROPESQ-UFRGS research funding.

## References

- Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19, 716–723.
- Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Wortman, J.R., Yandell, M.D., Zhang, Q., Chen, L.X., Brandon, R.C., Rogers, Y.H., Blazej, R.G., Champe, M., Pfeiffer, B.D., Wan, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Gabor, G.L., Abril, J.F., Agbayani, A., An, H.J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y., Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M., Dodson, K., Doup, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A.E., Garg, N.S., Gelbart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G.H., Ke, Z., Kennison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A.A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Pacle, J.M., Palazzolo, M., Pittman, G.S., Pan, S., Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Svirskaas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wassarman, D.A., Weinstock, G.M., Weissenbach, J., Williams, S.M., Woodage, T., Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M., Venter, J.C., 2000. The genome sequence of *Drosophila melanogaster*. Science 287, 2185–2195.
- Bächli, G., Vilela, C.R., Ratcov, V., 2000. Morphological differences among *Drosophila paraguayensis* Duda 1927 and its close relatives (Diptera, Drosophilidae). Bull. Soc. Entomol. Suisse 73, 67–92.
- Baker, R.H., Yu, X., DeSalle, R., 1998. Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. Mol. Phylogenet. Evol. 9, 427–436.
- Bernardi, N., 1981. Phylogenetic relationships, monophyletic group and related concepts. Rev. Bras. Entomol. 25 (4), 323–326.
- Bremer, K., 1988. The limits of amino acid sequence data in angiosperm phylogenetics reconstruction. Evolution 42, 795–803.
- Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.
- Brcic, D., Nair, P.S., Wheeler, M.R., 1971. Cytotaxonomic relationships within the *mesophragmatica* species group of *Drosophila*. Univ. Texas Publ. 7103, 1–16.
- Clary, D.O., Goddard, J.M., Martin, S.C., Fauron, C.M., Wolstenholme, D.R., 1982. *Drosophila* mitochondrial DNA: a novel gene order. Nucleic Acids Res. 10 (21), 6619–6637.
- DeSalle, R., 1992. The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. Mol. Phylogenet. Evol. 1, 31–40.
- 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. Mol. Phylogenet. Evol. 16, 296–307.
- Farris, J.S., Källersjö, M., Kluge, A., Bult, C., 1994. Testing significance of congruence. Cladistics 10, 315–320.
- Farris, J.S., Källersjö, M., Kluge, A., Bult, C., 1995. Constructing a significance test for incongruence. Syst. Biol. 44, 570–572.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approaching using bootstrap. Evolution 39, 783–791.
- Frota-Pessoa, O., 1954. Revision of the *tripunctata* group of *Drosophila* with description of fifteen new species. Arq. Mus. Paranaense 10, 253–304.

- Goto, S.G., Kimura, M.T., 2001. Phylogenetic utility of mitochondrial *COI* and nuclear *Gpdh* genes in *Drosophila*. *Mol. Phylogenet. Evol.* 18 (3), 404–422.
- Grimaldi, D.A., 1990. A phylogenetic, revised classification of the genera in the Drosophilidae (Diptera). *Bull. Am. Mus. Natl. Hist.* 197, 1–139.
- Hall, B.G., 2001. *Phylogenetic Trees Made Easy—a How to Manual for Molecular Biologists*, second ed. Sinauer Associates, Massachusetts.
- Heed, W.B., Russell, J.S., 1971. Phylogeny and population structure in island and continental species on the *cardini* group of *Drosophila* studied by inversion analysis. *Univ. Texas Publ.* 7103, 91–130.
- Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.
- Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., Gibson, T.J., 1998. Multiple sequence alignment with ClustalX. *Trends Biochem. Sci.* 23, 403–405.
- Jowett, T., 1986. Preparation of nucleic acids. In: Roberts, D.B. (Ed.), *Drosophila: a Practical Approach*. IRL Press, Oxford, pp. 275–276.
- Jukes, T.H., Cantor, C.R., 1969. Evolution of protein molecules. In: Munro, H.M. (Ed.), *Mammalian Protein Metabolism*. Academic Press, New York, pp. 21–132.
- Kastritis, C.D., 1969. The chromosomes of some species of the *guarani* group of *Drosophila*. *J. Hered.* 60, 50–57.
- Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA: Molecular Evolutionary Genetics Analysis Software. Version 2. Pennsylvania State University, University Park, and Arizona State University, Tempe.
- Kwiatowski, J., Sharecky, D., Bailey, K., Ayala, F.J., 1994. Phylogeny of *Drosophila* and related genera inferred from nucleotide sequence of the Cu, Zn *Sod* gene. *J. Mol. Evol.* 38, 443–454.
- Kwiatowski, J., Krawczyk, M., Jaworsky, M., Sharecky, D., Ayala, F.J., 1997. Erratic evolution of glycerol-3-phosphate dehydrogenase in *Drosophila*, *Chymomyza* and *Ceratitis*. *J. Mol. Evol.* 44, 9–22.
- Kwiatowski, J., Ayala, F.J., 1999. Phylogeny of *Drosophila* and related genera: conflict between molecular and anatomical analyses. *Mol. Phylogenet. Evol.* 13, 319–328.
- Lin, C.-P., Danforth, B.N., 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analysis of combined datasets. *Mol. Phylogenet. Evol.* 30, 686–702.
- Malogolowkin, C., 1953. Sobre a genitália dos drosofilídeos. IV. A genitália masculina no subgênero *Drosophila* (Diptera, Drosophilidae). *Rev. Bras. Biol.* 13, 245–264.
- Nacur, J., 1958. Genitália masculina de *Drosophila* do grupo *mesophragmatica* (Diptera). *Rev. Bras. Biol.* 18, 243–249.
- Nair, P.S., Brncic, D., Kojima, K.-I., 1971. II. Isozyme variations and evolutionary relationships in the *mesophragmatica* species group of *Drosophila*. *Stud. Genet.* VI, 17–28.
- 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. *Mol. Biol. Evol.* 13, 132–149.
- O'Grady, P.M., Kidwell, M.G., 2002. Phylogeny of the subgenus *Sophophora* (Diptera: Drosophilidae) based on combined analysis of nuclear and mitochondrial sequences. *Mol. Phylogenet. Evol.* 22 (3), 443–453.
- Pélandakis, M., Solignac, M., 1993. Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. *J. Mol. Evol.* 37, 525–543.
- Perlam, S.J., Spicer, G.S., Shoemaker, D.D., Jaenike, J., 2003. Associations between mycophagous *Drosophila* and their *Howardula* nematode parasites: a worldwide phylogenetic shuffle. *Mol. Ecol.* 12 (1), 237–249.
- Posada, C., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Powell, J.R., DeSalle, R., 1995. *Drosophila* molecular phylogenies and their uses. In: Hecht, M.K. (Ed.), *Evolutionary Biology*. Plenum, New York, pp. 87–139.
- Remsen, J., DeSalle, R., 1998. Character congruence of multiple data partitions and the origin of the Hawaiian Drosophilidae. *Mol. Phylogenet. Evol.* 9 (2), 225–235.
- Remsen, J., O'Grady, P., 2002. Phylogeny of Drosophilinae (Diptera: Drosophilidae), with comments on combined analysis and character support. *Mol. Phylogenet. Evol.* 24, 249–264.
- Russo, C.A.M., Takezaki, N., Nei, M., 1995. Molecular phylogeny and divergence times of Drosophilid species. *Mol. Biol. Evol.* 12, 391–404.
- Russo, C.A.M.de, 2001. Como escolher genes para problemas filogenéticos específicos. In: Matioli, S.R. (Ed.), *Biologia molecular e evolução*. Holos, Ribeirão Preto.
- Saitou, N., Nei, M., 1987. The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 9, 406–425.
- Schneider, H., 2003. Métodos de análise filogenética—Um guia prático. Holos and SBG, Ribeirão Preto.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116.
- Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* 87, 651–701.
- Sorenson, M.D., 1999. TreeRot, version 2. Boston University, Boston, MA.
- Spicer, G.S., 1995. Phylogenetic utility of the mitochondrial cytochrome oxidase gene: molecular evolution of the *Drosophila buzzatii* species complex. *J. Mol. Evol.* 41 (6), 749–759.
- Spicer, G.S., Jaenike, J., 1996. Phylogenetic analysis of breeding site use and amanitin tolerance within the *Drosophila quinaria* species group. *Evolution* 50, 2328–2337.
- Spicer, G.S., Pitnick, S., 1996. Molecular systematics fo the *Drosophila hydei* subgroup as inferred from mitochondrial DNA sequences. *J. Mol. Evol.* 43 (3), 281–286.
- Staden, R., 1996. The Staden sequence analysis package. *Mol. Biotechnol.* 5, 233–241.
- Swofford, D.L., 2003. PAUP: phylogenetic analysis using parsimony (and other methods). Version 4. Sinauer Associates, Massachusetts.
- Tamura, K., 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C biases. *Mol. Biol. Evol.* 9, 678–687.
- Tamura, K., Nei, M., 1993. Estimation of the number of base nucleotide substitution in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526.
- Tatarenkov, A., Kwiatowski, J., Skarecky, D., Barrio, E., Ayala, F.J., 1999. On the evolution of *Dopa decarboxylase (Ddc)* and *Drosophila* systematics. *J. Mol. Evol.* (48), 445–462.
- Tatarenkov, A., Ayala, F.J., 2001. Phylogenetic relationships among species groups of the *virilis-repleta* radiation of *Drosophila*. *Mol. Phylogenet. Evol.* 21, 327–331.
- Tatarenkov, A., Zurovcová, M., Ayala, F.J., 2001. *Ddc* and *Amd* sequences resolve phylogenetic relationships of *Drosophila*. *Mol. Phylogenet. Evol.* 20, 321–325.
- Templeton, A.R., 1983. Convergent evolution and non-parametric inferences from restriction fragment and DNA sequence data. In: Weir, B. (Ed.), *Statistical Analysis of DNA Sequence Data*. Marcel Dekker, New York, pp. 151–179.
- Thomas, R.H., Hunt, J.A., 1993. Phylogenetic relationships in *Drosophila*: a conflict between molecular and morphological data. *Mol. Biol. Evol.* 10 (2), 362–374.
- Throckmorton, L.H., 1975. The phylogeny, ecology and geography of *Drosophila*. In: King, R.C. (Ed.), *Handbook of Genetics*. Plenum, New York, pp. 421–469.
- Val, F.C., Marques, M.D., Vilela, C.R., 1981. Drosophilidae of Neotropical region. In: Ashburner, M., Carson, H.L., Thompson, J.N.

- (Eds.), The Genetics and Biology of *Drosophila*. Academic Press, Orlando, pp. 123–168.
- Vilela, C.R., Bächli, G., 1990. Taxonomic studies on Neotropical species of seven genera of (Drosophilidae Diptera). Bull. Soc. Entomol. Suisse 63 (Suppl.), 1–332.
- Vilela, C.R., Bächli, G., 2000. Five new species of Neotropical *Drosophila* (Diptera, Drosophilidae). Bull. Soc. Entomol. Suisse 73, 49–65.
- Vilela, C.R., Valente, V.L.S., Basso-da-Silva, L., 2004. *Drosophila angustibucca* Duda *sensu* Frota-Pessoa is an undescribed species (Diptera, Drosophilidae). Rev. Bras. Entomol. 48, 233–238.
- Xia, X., Xie, Z., 2001. DAMBE: data analysis in molecular biology and evolution. J. Hered. 92, 371–373.
- Yotoko, K.S.C., Medeiros, H.F., Solferini, V.N., Klaczko, L.B., 2003. A molecular study of the systematics of the *Drosophila tripunctata* group and the *tripunctata* radiation. Mol. Phylogenet. Evol. 28, 614–619.
- Wheeler, M.R., 1982. The Drosophilidae: a taxonomic overview. In: Ashburner, M., Carson, H.L., Thompson, J.N. (Eds.), The Genetics and Biology of *Drosophila*. Academic Press, Orlando, pp. 1–97.

### CAPÍTULO III

---

Trabalho submetido à revista *Molecular Phylogenetics and Evolution*

**On the phylogenetic relationships of the subgenus *Drosophila* (Diptera, Drosophilidae): New evidences from the *Amd*, *Ddc* and *hunchback* gene partitions**

Lizandra J. Robe<sup>1\*</sup>, Elgion L. S. Loreto<sup>1,2</sup>, Vera L. S. Valente<sup>1</sup>

1 – Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul (UFRGS), CP 15053, Porto Alegre, Rio Grande do Sul, Brazil, 91501-970

2 – Departamento de Biologia, Universidade Federal de Santa Maria (UFSM), Camobi Km 09, Santa Maria, Rio Grande do Sul, Brazil, 97105-900

\* Corresponding author: Phone (55)(51) 3308-6713; Fax (55)(51) 3308-7311

E-mail addresses: [lizbiogen@yahoo.com.br](mailto:lizbiogen@yahoo.com.br) (L. J. Robe), [elgion@base.ufsm.br](mailto:elgion@base.ufsm.br) (E. L. S. Loreto), [vera.gaiesky@ufrgs.br](mailto:vera.gaiesky@ufrgs.br) (V. L. S. Valente).

## Abstract

Despite the central role that the genus *Drosophila* has played in the origin and subsequent development of genetical and evolutionary research, the establishment of its phylogenetic relationships is not completely reached until now. The phylogeny of the subgenus *Drosophila* presents the worst picture, probably due to its evolutionary history, postulated to have involved fast and successive radiations. In order to shed light on different aspects related to the subgenus *Drosophila* evolution, nucleotide sequences from the *alpha methyl-dopa* (*Amd*) and *dopa decarboxylase* (*Ddc*) genes for at least 80 different taxa were analysed both individually and simultaneously through a Bayesian framework. Besides, rapidly evolving sequences from the nuclear *hunchback* (*hb*) and the mitochondrial *cytochrome oxidase II* (*COII*) were joined to the previous ones with the aim of refining the *immigrans-tripunctata* radiation evolutionary history. In general, our results agree with previous evidence in that they recover the paraphyly of the genus *Drosophila*, which presents the subgenus *Sophophora* as basal to all other species. Hence, also the subgenus *Drosophila* is paraphyletically defined, and its *immigrans-tripunctata*, *virilis-repleta* and Hawaiian *Drosophilidae* radiations are confidently recovered. It is striking that Bayesian dating estimates that the basal branchings within these three radiations occurred along Oligocene and Early Miocene epochs, providing an exception to the idea that most speciation events within the *Drosophila* genus are related to fragmentations in paleoclimatic cooling periods. At least some of these divergences could be associated to the invasion of new areas, as represented by the New World environment.

Keywords: *Amd* gene; *Ddc* gene; genus *Drosophila*; Hawaiian *Drosophilidae* radiation; *Hb* gene; *immigrans-tripunctata* radiation; *repleta* radiation; subgenus *Drosophila*; *tripunctata* radiation; *virilis-repleta* radiation.

## 1. Introduction

Over the past century, the genus *Drosophila* has played a central role in the origin and subsequent development of genetical, evolutionary, developmental and even ecological research. In this sense, it can be stated that deciphering its phylogenetic relationships is of fundamental importance (Powell and DeSalle, 1995), given that: (1) it allows the analysis of the evolutionary history of a series of characteristics already studied for the group; (2) it supplies the basis from which to obtain higher-level conclusions, starting from the analysis of the phylogenetic context applicable to the results; (3) it is of assistance to define sample strategies. These phylogenetic studies became even more essential with the availability of the full genomes of 12 *Drosophila* species (Clark et al., 2007), which certainly will enhance comparative studies that require subjacent working phylogenies. However, as this genus is a very speciose one, the establishment of its systematics and phylogenetic relationships is no easy task, having been aimed by many researchers since Sturtevant (1939 and 1942) but not completely achieved so far.

The subgenus *Drosophila* is the largest between the eight subgenera that nowadays comprise the genus *Drosophila* (Bächli, 2008). It encompasses over 720 described species, more than half of the total number of *Drosophila* known at present. Markow and O'Grady (2006), based on an update of Throckmorton (1975) assumptions, subdivided the subgenus *Drosophila* into three main lineages: (1) the *virilis-repleta* radiation, that comprises the *angor*, *annulimana*, *bromeliae*, *canalinae*, *carbonaria*, *carsoni*, *coffeata*, *dreyfusi*, *flavopilosa*, *melanica*, *mesophragmatica*, *nannoptera*, *peruviana*, *polychaeta*, *quadrisetata*, *repleta*, *robusta*, *tumiditarsus*, *virilis* and *xanthopallescens* groups; (2) the *immigrans-tripunctata* radiation, containing the *calloptera*, *cardini*, *guarani*, *guaramunu*, *immigrans*, *macroptera*, *pallidipennis*, *pinicola*, *quinaria*, *rubrifrons*, *sticta*, *testacea* and *tripunctata* groups; (3) the Hawaiian *Drosophilidae* radiation, that includes the Hawaiian *Drosophila* and *Scaptomyza*. Additional support for this kind of subdivision can be found specially at Kwiatowski and Ayala (1999), Remsen and DeSalle (1998), Remsen and O'Grady (2002) and Tatarenkov et al. (2001), but similar results were also presented by DaLage et al. (2007), Robe et al. (2005), Russo et al. (1995), Tatarenkov et al. (1999a) and Yotoko et al. (2003). There is also a great deal of evidence showing that this subgenus is paraphyletic with regard to other *Drosophilidae* genera and subgenera (Kato et al., 2000; Kwiatowski and Ayala, 1999; Remsen and DeSalle, 1998; Remsen and O'Grady, 2002; Robe et al., 2005; Russo et al., 1995; Tatarenkov et al., 1999a and 2001). However, these relationships are still largely controversial.



According to Throckmorton (1962), the evolution of the subgenus *Drosophila* branch has been in marked contrast with the other species pattern, given that it occurred with the perpetuation and even with an increase of the ancestral genetic variability, which provided their descendants the required versatility to exploit new, rich and unexploited environments more fully. That author also states that, after subdivision, these large, complex and polymorphic ancestral populations appear to have suffered a recent and explosive adaptive radiation in the New World, after which populations rapidly expanded. In agreement with this, Powell and DeSalle (1995) also pointed out that the subgenus *Drosophila* represents a complex picture, given that its major lineages are quite old and its phylogenetic resolution is often not great. These hypotheses may explain the fact that the phylogeny of this subgenus has a much more brushy appearance, with deep nodes and poor resolution (DaLage et al., 2007; Robe et al., 2005; Yotoko et al., 2003).

The possible scenario depicted above, added to the fact that many previous molecular phylogenetic studies have encompassed a relatively narrow taxonomic range within the subgenus *Drosophila*, render many poorly understood aspects to the phylogeny of the group. Considering these uncertainties, the aims of this study were: (1) to assess the monophyly of the subgenus *Drosophila*, concerning other *Drosophila* subgenera, as *Sophophora* and *Dorsilopha*, and also other Drosophilidae genera, as *Hirtodrosophila*, *Liodrosophila*, *Samoaia*, *Scaptomyza* and *Zaprionus*; (2) to evaluate the taxonomic validity of the several radiations proposed by Throckmorton (1975) for the subgenus, as the *virilis-repleta*, the *repleta*, the *immigrans-Hirtodrosophila* and the *tripunctata* radiations; (3) to establish the phylogenetic relationships within and between some of the species groups traditionally placed at this subgenus, which are: the *cardini*, *calloptera*, *guarani*, *guaramunu*, *immigrans*, *pallidipennis*, *quinaria*, *testacea* and *tripunctata* groups, included in the *immigrans-tripunctata* radiation; and the *annulimana*, *bromeliae*, *dreyfusi*, *canalina*, *flavopilosa*, *melanica*, *mesophragmatica*, *nannoptera*, *polychaeta*, *repleta*, *robusta* and *virilis* groups, assembled in the *virilis-repleta* radiation; and (4) to address the placement of the Hawaiian Drosophilidae radiation within the subgenus. For these purposes, about 87 different taxa were analyzed for the two linked paralogous nuclear genes *alpha methyl dopa* (*Amd*) and *dopa decarboxylase* (*Ddc*) (Eveleth and Marsh, 1986), both individually and simultaneously. Although these analyses provided a reliable description of the subgenus *Drosophila* evolution, sequences from the rapidly evolving nuclear *hunchback* (*hb*) and mitochondrial *cytochrome oxidase subunit II* (*COII*) genes were also obtained in order to better estimate the *immigrans-tripunctata* lineage phylogeny. The data as a whole support the notion of three different radiations within

the subgenus *Drosophila* evolution, supplying many new and interesting insights concerning their evolutionary history.

## 2. Materials and methods

### 2.1. Species

The 87 taxa studied are listed in Table 1, which also includes their source and respective *Amd*, *Ddc*, *Hb* and *COII* GenBank Accession Numbers at the National Center for Biotechnology Information (NCBI). As referred to in that Table, the flies were either collected from the field by specific collaborators or obtained from the Tucson *Drosophila* Stock Center, after which isolineages were established and maintained in a 17°C ± 1°C chamber. Two species of the genus *Scaptodrosophila*, *S. latifasciaeformis* and *S. lebanonensis* were utilized as outgroups in the subgenus *Drosophila* analyses, as previously suggested (DaLage et al., 2007; DeSalle, 1992a; Grimaldi, 1990; Katoh et al., 2000; Kwiatowski et al., 1994, 1997; Pélandakis and Solignac, 1993; Russo et al., 1995; Tatarenkov et al., 1999a; Throckmorton, 1975). For the analysis involving only the *immigrans-tripunctata* radiation, however, *Zaprionus indianus* and *Z. tuberculatus* were the outgroup, as there is ample evidence that this is an early offshoot with respect to the *immigrans-tripunctata* radiation clade (Robe et al., 2005; Throckmorton, 1975) or at least with respect to the subgenus *Drosophila* (DaLage et al., 2007; Kwiatowski and Ayala, 1999; Remsen and DeSalle, 1998; Remsen and O'Grady, 2002; Russo et al., 1995; Tatarenkov and Ayala, 1999a and 2001; Yotoko et al., 2003). More than one population from a same species were included whenever possible to gauge the intraspecific variations and evidence cases of shared ancestral polymorphisms.

### 2.2. DNA manipulation

Total genomic DNA of each lineage was extracted according to Sassi et al. (2005). The amplification and sequencing strategies applied to the three nuclear markers *Amd*, *Ddc* and *Hb* are illustrated in Figure 1. For *Amd*, a new forward primer, *AmdEx4F* (5' GTB GTB GTN ATG GAY TGG YTV GCN AA 3'), was designed and utilized with the reverse *Amd-bw* (Tatarenkov et al., 2001). For *Ddc* and *Hb*, the primers used in PCR amplifications were the same as those employed in Mota et al. (in press). As a broad taxonomic range was sampled, PCR parameters for *Amd*, *Ddc* and *Hb* varied and were individually optimized, while the PCR conditions followed the general procedure described in Robe et al. (2005). All *COII* nucleotide sequences were obtained from GenBank (NCBI site).

*Hb* PCR products were purified with a solution of PEG 13% and NaCl 1.6M, and both strands of the PCR fragments were directly sequenced in a MegaBACE 500 automatic sequencer, with the use of the DYEnamic ET® Sequencing Kit (Amersham) according to the manufacturer's protocol. In this case, the two amplification primers were also used for sequencing.

For *Amd* and *Ddc*, nevertheless, the only way to obtain sequences of high quality was by cloning the amplified fragments. In this way, PCR products amplified with the use of Platinum® Taq Polymerase High Fidelity (Invitrogen) were gel-purified with QIAquick Gel Extraction Kit (Qiagen) and cloned into pCR®4 vector from the TOPO-TA Cloning Kit (Invitrogen), using *E. coli* DH10β as host. Sequences were determined in both directions using M13 standard primers and the BigDye Terminator Sequencing Kit (Applied Biosystems) in an ABI-PRISM 3100 Genetic Analyzer.

### 2.3. Data analysis

The resulting sequence electropherograms were assembled, inspected and corrected with the use of the Staden Package Gap 4 program (Staden, 1996). The contigs from *Amd*, *Ddc* and *COII* were aligned through the use of the ClustalW algorithm as implemented in Mega 3.1 (Kumar et al., 2004). *Hb* alignment, however, was far from trivial, requiring the use of a more accurate alignment program, the MUSCLE software (Edgar, 2004). Anyway, the conflicting and doubtful areas of the *Hb* provided alignment were later removed, so that the matrix initially containing 725 bp was reduced to only 507 bp.

All phylogenetic analyses were conducted in two different stages: (1) the phylogeny of the subgenus *Drosophila* as a whole was assessed through the *Amd* and *Ddc* partitions, which were analysed both individually and simultaneously; (2) derived relationships within the *immigrans-tripunctata* radiation were focused through a combined analysis encompassing not only *Amd* and *Ddc*, but also a rapidly evolving region of the nuclear *Hb* gene and the mitochondrial *COII* gene. In the *immigrans-tripunctata* radiation combined matrix, particularly, some taxa for which no character data were available for one of the genes were coded as missing for that partition. Also in this case, a placeholder, *D. calloptera*, was utilized in substitution to *D. ornatipennis* for the *COII* data set.

In order to determine whether nucleotide composition differences occur among taxa, the chi-square test implemented in PAUP version 4.0b10 (Swofford, 2003) and in TreePuzzle 5.2 (Schmidt et al., 2002), and the disparity index test (Kumar and Gadagkar, 2001) implemented in Mega 3.1 were performed on the data sets. Mega 3.1 program, together with ModelTest 3.7 (Posada and Crandall, 1998) and BaseML of

PAML 3.14 (Yang, 1997) were used to calculate some descriptive parameters (Table 2 and 3), so as to characterize the subjacent substitution patterns.

Several different methods were used to reconstruct the phylogenetic relationships between the sampled OTU's. Neighbor Joining (NJ) and Maximum Parsimony (MP) were implemented in PAUP 4.0b10. Maximum Likelihood (ML) inferences were carried out on PhyML 2.4.4 (Guindon and Gascuel, 2003), whereas Bayesian Analyses (BA) were conducted on MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001) and on BayesPhylogenies (Pagel and Meade, 2004). NJ analyses were conducted through the use of the LogDet correction (Lockhart et al., 1994), since significant base composition heterogeneity was detected among the taxa. MP analyses were performed adopting both an un-weighted and a weighted strategy. In the last case, third codon positions were down-weighted with regard to the first and second codon positions (2:2:1) in order to reduce the effects of non-stationarity. All the MP trees were obtained through heuristic search, with Tree Bisection-Reconnection (TBR) branch swapping applied to 100 random stepwise addition starting trees. ML inferences were performed using a tree generated by a NJ heuristic search as input (BIONJ option) and applying the model selected by the Akaike Information Criterion (AIC) Test (Akaike, 1974) as performed on ModelTest 3.7, whose parameters were individually estimated and optimized by the PhyML program. For the BA implemented in MrBayes 3.1.2 at least 3,000,000 generations of the Markov Chain Monte Carlo (MCMC) search were conducted, according to the model proposed by the AIC test as implemented in MrModeltest 2.2 (Nylander, 2004). In general, the BA trees were sampled every 1,000 generations, and a burnin of approximately 20% was applied since runs were stopped only after the average deviation of split frequencies fell well below 0.01. For the Bayesian analysis of combined data sets, individual models and unlinked parameters were assigned to each partition before running MrBayes 3.1.2. Additionally, the combined matrix was also submitted to the novel approach presented by the software BayesPhylogenies, which allows *a posteriori* partitioning of the datasets. In this case, the specified number of partitions reflected the number of different data sets introduced in the concatenated matrix. TreePuzzle 5.2 was utilized to build a 50% majority rule consensus tree with estimated branch lengths from the BayesPhylogenies output trees. Support for each clade in the NJ, MP and ML trees was assessed by the bootstrap test (Felsenstein, 1985) with 1,000 replications, whereas for the BA, the posterior probability (PP) of each clade on the 50% majority rule consensus tree was measured.

Before concatenation, the partitions were tested for congruence through the use of the Incongruence Length Difference (ILD) test (Farris et al., 1994 and 1995), which is referred to as the Partition Homogeneity Test (PHT) in PAUP. This test was

implemented under parsimony criterion with 1,000 replications, for each of which 100 starting trees were generated by random stepwise addition followed by TBR in order to obtain the null distribution. In the combined analyses, the average Partition Bremer Support (PBS) scores (Bremer, 1988, 1994) presented by each partition on each node were assessed using PAUP 4.0b10 and TreeRot 3.0 (Sorenson and Franzosa, 2007). The standardization method previously reported in Robe et al. (2005) was also used here. Furthermore, to measure the amount of incongruence between alternative topologies, these were compared using the Templeton Test (Templeton, 1983), under parsimony criterion and the Shimodaira-Hasegawa (SH) Test (Shimodaira and Hasegawa, 1999), under likelihood criterion, as performed in PAUP 4.0b10. For the SH Test the null distribution was generated through 1,000 nonparametric bootstrap replicates using the resampling estimated log-likelihood (RELL) strategy.

Divergence times were also estimated through the use of a Bayesian dating method implemented in MultiDivTime (Thorne et al., 1998; Thorne and Kishino, 2002), which enables inferences concerning divergence time using a relaxed molecular clock procedure. In this case, the Bayesian tree obtained through the combined *Amd* and *Ddc* analysis was used as the input tree file for estimating the model parameters in BaseML (part of PAML), and also to estimate the maximum likelihood of the branch lengths and its variance-covariance matrix in EstBranches, according to Rutshmann (2005). The split between the Hawaiian *Scaptomyza* and *Drosophila* species as occurred about  $30.5 \pm 6.6$  My (Tamura et al., 2004) was used as time calibration point. This genomic molecular clock estimation is in broad agreement with fossil records, since *Scaptomyza* fossils were found in Dominican amber with a minimal estimated age of 23 My (Grimaldi, 1987). The time set between ingroup root and tips was 63 My, with a standard deviation of 12.4 My (Tamura et al., 2004).

### **3. Results**

#### **3.1. Data partitions**

Empirical and estimated characteristics of each of the data sets analysed are presented in Table 2, whereas their optimal substitution models and parameters, as selected by the AIC Test executed by ModelTest 3.7, are indicated in Table 3. As can be seen in these tables, all the matrixes may be affected by some levels of saturation, brought about not only by multiple substitutions, but also by homoplasies related to the great transition/transversion ratio, to the compositional bias, to the assymmetric matrix of rate substitution, and to the significant heterogeneity in evolutionary rates across sites.

Moreover, the chi-square test implemented in PAUP 4.0b10 detected significant base composition heterogeneity between taxa in all data sets (Table 2). The disparity index ( $I_D$ ), which measures the observed difference of substitution patterns between a pair of sequences (Kumar and Gadagkar, 2001) revealed that *D. mediopicta* presents a huge deviation from the other taxa base frequencies. This species showed total  $I_D$  values as extreme as 16.075 (in the comparison with *D. ornatifrons\_1* for *Amd*) and 16.640 (in the comparison with *D. ornatifrons\_1* for *Ddc*). It is not surprising that the single matrix that did not reach the 0.01 significance level for the chi-square test was *Hb*, which does not possess *D. mediopicta* sequences. Additionally, performing the Tree-Puzzle chi-square test, which compares the nucleotide composition of each sequence to the frequency distribution assumed in the maximum likelihood model, *D. ornatifrons*, *D. paramediostriata*, *D. maculifrons* and *D. griseolineata* also detached as possessing significant shifts in their GC content for the *Amd*, *Ddc* and *Hb* partitions. Among the members of the *virilis-repleta* section, only *D. robusta* presented significant base composition heterogeneity for the *Amd* and *Ddc* data sets.

To account for these biases, likelihood-based methods were emphasized, since these are known to be more robust to assumptions violations (Swofford et al., 1996) and seem to perform better when there is significant base composition heterogeneity between taxa (Galtier and Gouy, 1995 and 1998).

### 3.2 – The subgenus *Drosophila* phylogeny

#### 3.2.1 – The *Amd* partition

Figure 2 shows the Bayesian phylogeny obtained with the *Amd* data partition. Rooting this phylogeny with *Scaptodrosophila* yields a genus *Drosophila* entirely paraphyletic with respect to the genera *Hirtodrosophila*, *Liodrosophila*, *Samoaia*, *Scaptomyza* and *Zaprionus*. Within this major cluster that comprises the genus *Drosophila*, the subgenus *Sophophora* represents an early offshoot (PP of 0.84). *D. busckii*, a member of the *Dorsilopha* subgenus, is the next to branch off, leaving a well supported clade (PP of 0.95) that, although poorly resolved at the base, splits into four main lineages: (1) the first, monoespecific, composed by *H. pictiventris*; (2) the second grouping *Zaprionus* and *Liodrosophila* as sister genera, although with low support (PP of 0.60); (3) the third confidently clustering a monophyletic *virilis-repleta* radiation with the Hawaiian Drosophilidae radiation clade (PP of 1.00); (IV) and the fourth presenting *S. leonensis* at the base of a well defined *immigrans-tripunctata* radiation (PP of 1.00).

Nested within the *virilis-repleta* radiation cluster there is a monophyletic *repleta* radiation (PP of 0.87), that presents the *mesophragmatica* and *repleta* groups as sister clades (PP of 1.00), more closely related to each other than to the *dreyfusi* and

*canalinae* groups. A clade comprised by the sister species groups *bromeliae* and *nannoptera* represent the *repleta* radiation sister taxon, with a high support value (PP of 1.00). The monophyletic *annulimana* and *flavopilosa* groups detach as the closest relatives to the lineage depicted above (PP of 0.97), although their exact relationships between each other and that clade remain unsolved. Otherwise, at the base of the *virilis-repleta* radiation clade, *D. polychaeta*, a representative of the *polychaeta* group, is the early offshoot (PP of 0.81), followed in successive progression by the *D. robusta* and *D. melanica* sister groups clade (PP of 1.00) and by the *D. virilis* group (PP of 0.97), which appears as phylogenetically closer to the clade encompassing the *repleta* radiation than to the *robusta/melanica* lineage.

As regards the *immigrans-tripunctata* radiation clade, the *immigrans*, *phalerata*, *funnebris* and *testacea* groups are the deepest, whereas the monophyletic *tripunctata* radiation occupies a derived position with strong confidence (PP of 1.00). This lineage branches off into five well supported clades (specified in an ascending order beginning with the most basal): (1) the first grouping *D. bandeirantorum* (*tripunctata* group) and *D. pallidipennis* (*pallidipennis* group) as sister to a cluster encompassing other six species of the *tripunctata* group (*D. cuaso*, *D. paraguayensis*, *D. tripunctata*, *D. mediopunctata*, *D. unipunctata* and *D. roehrae*) (PP of 1.00); (2) the second comprising three representatives of the *guarani* group (*D. ornatifrons*, *D. guaru* and *D. subbadia*) (PP of 1.00); (3) the third encompassing a monophyletic *cardini* group, at least as concerned to the *D. cardini*, *D. cardinoides*, *D. polymorpha*, *D. neocardini*, *D. nigrodunni* and *D. arawakana* species (PP of 1.00); (4) the fourth joining *D. nappae* (*tripunctata* group) with *D. ornatipennis* (*calloptera* group) (PP of 0.99); (5) and the fifth presenting *D. maculifrons* and *D. griseolineata* (*guaramunu* group) as intermediate species that connect two clusters composed by members of the *tripunctata* group: the *D. crocina* and *D. mediodistriata* lineage and the *D. mediodiffusa*, *D. mediopictoides*, *D. paramediodistriata* and *D. mediopicta* lineage (PP of 1.00). Hence, the *Amd* Bayesian phylogeny recovers a *tripunctata* group paraphyletic with respect to other members of the *tripunctata* radiation, that is, the *calloptera*, *cardini*, *guarani*, *guaramunu* and *pallidipennis* species groups.

Concerning the trees obtained by the different methods, ML yielded an extremely similar topology, except for some branches presented as polytomous or poorly supported (PP < 0.70) in the BA phylogeny. Besides that, there are incongruencies as regards some relationships within the *immigrans-tripunctata* radiation, although in general these conflicting clades are only weakly supported by the ML tree. Thus, ML favours the clustering of *D. funnebris* and *D. putrida* with a tenuous bootstrap support (38%) at the base of the *tripunctata* radiation whose internal

composition differ from the Bayesian one concerning the relationships between the five major clades depicted above, and regarding also the relationships between species within some of them. An important difference between the BA and ML topologies refers to the branching position of *D. ornatipennis* (*calloptera* group), which clusters with *D. nappae* in the BA, but with the *cardini* group in the ML analysis (68% bootstrap). On the other hand, the MP topologies were very polytomous at the base, and the only conflict with the BA refer to the MP weighted tree positioning of *D. mediopicta* as an early offshoot (55% bootstrap) in the clade that joins the *guaramunu* group species, *D. griseolineata* and *D. maculifrons*, with the *tripunctata* group species *D. paramediostriata*, *D. mediodiffusa* and *D. mediopictoides*. This configuration is also supported by the NJ bootstrap tree, that besides being very unresolved at the base, presented some unexpected results, as the *repleta* group polyphyly and the subdivision of the *virilis-repleta* radiation into two clades: one composed by the *repleta* radiation plus the *bromeliae/nannoptera* cluster (68% bootstrap); the other presenting the *annulimana*, *flavopilosa*, *virilis*, *robusta* and *melanica* species groups (82% bootstrap). Furthermore, NJ favors the grouping of *D. nappae* with the clade composed by members of the *guarani* group (51% bootstrap), and presents different relationships between the species of the *D. cardini* group. Comparing the bootstrap topologies obtained by the different methods under the Shimodaira-Hasegawa and Templeton Tests revealed that the BA and ML trees performed significantly better than the MP and NJ ones not only under maximum likelihood but also under the maximum parsimony criteria. This result was probably a consequence of the poor resolution presented at the base of the MP and NJ bootstrap trees.

### 3.2.2 – The *Ddc* partition

The *Ddc* Bayesian phylogeny, however, presents some important differences concerning the *Amd* one (Figure 3). In this case, the early offshoot within the genus *Drosophila* is represented not only by the subgenus *Sophophora*, but also by the subgenus *Dorsilopha*, represented by *D. busckii* (PP of 0.98). After that divergence, the clade defining a paraphyletic *Drosophila* subgenus splits into two major lineages: the one comprising a *virilis-repleta* radiation, differentially composed to the exclusion of *D. polychaeta* (PP of 0.95); and the one assembling the members of the traditional *immigrans-Hirtodrosophila* radiation (Throckmorton, 1975), given the inclusion of *L. aerea* and *D. funebris*, and the exclusion of *D. immigrans* and *D. busckii* (PP of 0.80). According to this scenario the Hawaiian Drosophilidae are embedded within this *immigrans-Hirtodrosophila* radiation clade.

As regards the *virilis-repleta* radiation relationships, the *Ddc* Bayesian phylogeny defines the *repleta* radiation as paraphyletic with respect to the *bromeliae*



and *nannoptera* sister groups (PP of 0.62). Furthermore, the *repleta* and *mesophragmatica* sister groups appear as phylogenetically closer to *D. canalinea* (*canalinea* group) and to the *bromeliae/nannoptera* clade than to *D. camargoi* (*dreyfusi* group) (PP of 1.00), a branching order opposed to that presented in the *Amd* phylogenies. On the other hand, the poorly supported positioning of the *annulimana* and *flavopilosa* groups in the *Amd* Bayesian topology is somewhat resolved by *Ddc*, that presents these as sister taxa (PP of 1.00). Another source of ambiguity between *Amd* and *Ddc* concerns the position of *D. virilis* (*virilis* group), placed as the *annulimana/flavopilosa* clade sister species by *Ddc* (PP of 1.00). Given the absence of *D. polychaeta* at the *Ddc virilis-repleta* clade, the clade joining *D. robusta* (*robusta* group) and *D. melanica* (*melanica* group) detach as the oldest lineage (PP of 0.70).

Within the *immigrans-Hirtodrosophila* radiation clade, the clustering of *H. pictiventris* and *S. leonensis* added to the branching of the Hawaiian Drosophilidae define the oldest lineages, although support for this is weak (PP of 0.53). The clade joining the genera *Zaprionus* and *Liodrosophila* with a strong support (PP of 0.99) is the next to branch off, and its sister clade splits into two major lineages: one giving rise to the ancestor from which *D. phalerata* (*quinaria* group) and *D. funebris* (*funebris* group) diverged (PP of 1.00); and the other encompassing a monophyletic *tripunctata* radiation (PP of 1.00). The five major lineages defined by the *Amd* Bayesian tree are also confidently recovered by the *Ddc* partition, although some of their internal relationships differ among the two partitions. On the other hand, the general configuration of these five lineages across the tree is very similar between the *Amd* and *Ddc* Bayesian phylogenies.

Although again both Shimodaira-Hasegawa and Templeton Tests rejected MP and NJ bootstrap topologies as significantly worse than the ML and BA ones (Templeton also rejected the BA), there is not a great deal of conflict between them. The *Ddc* ML tree presented some interesting incongruencies as regards the BA one, but as for *Amd*, these conflicts are always poorly supported. In this way, ML presents *D. camargoi* as the sister species of the *repleta/mesophragmatica* cluster (34% bootstrap), although the *repleta* radiation continues paraphyletic with respect to the *nannoptera/bromeliae* clade (20% bootstrap). The *melanica/robusta* clade also altered its position in the ML analysis, clustering with the *virilis/annulimana/flavopilosa* lineage (50% bootstrap). Furthermore, the ML topology also recovered the Hawaiian Drosophilidae as the *virilis-repleta* radiation sister lineage (18% bootstrap), defining a clade that present the *Samoaia/Hirtodrosophila* branch as the closest outgroup. In the ML tree *immigrans-Hirtodrosophila* radiation cluster, the *tripunctata* radiation is paraphyletic as regards the *funebris/quinaria* clade, which joins the lineage composed

by *D. tripunctata*, *D. mediopunctata*, *D. unipunctata*, *D. paraguayensis*, *D. cuaso*, *D. bandeirantium* and *D. pallidipennis* with a bootstrap support of 40%. There are also some tenuous conflicts within the *tripunctata* radiation clade concerning the relationships between the five major lineages above depicted, whose internal composition highly resemble the BA one. The MP weighted and unweighted strategies resulted in poorly resolved bootstrap topologies, whose major conflict with the BA concerns the position of *D. mediopicta* as the deepest (61% and 64% bootstrap, respectively) in the lineage that comprises *D. griseolineata* and *D. maculifrons* (*guaramunu* group) besides *D. mediodiffusa*, *D. mediopictoides* and *D. paramediostriata* (*tripunctata* group). NJ also supports this relationship (100% bootstrap), and it is the unique method that recovers a monophyletic *repleta* radiation (76% bootstrap) with the *Ddc* partition. Besides that, it presents the grouping of the *D. crocina*/*D. mediostriata* clade with *D. nappae* (68% bootstrap), breaking one of the above depicted *tripunctata* radiation lineages, but conserving the others.

### 3.2.3 – The *Amd* + *Ddc* total evidence tree

The PHT results revealed that the *Amd* and *Ddc* partitions are not significantly heterogeneous ( $P = 0.166$ ). In addition, the rank correlation coefficient presented by these in the MultiDivTime output was near 0.78 ( $p = 0.001$ ), leaving to a rejection of the null hypothesis that this pair of genes change rate independently. Therefore, as *Amd* and *Ddc* also share extremely similar parameters for the AIC selected GTR + I + G model, a concatenated analysis was considered suitable. Figure 4 shows the Bayesian phylogeny generated from the *Amd* and *Ddc* data sets analyzed simultaneously. This hypothesis is broadly congruent with the *Amd* and *Ddc* individual ones, except for some polytomous or poorly supported basal relationships and for some derived aspects concerning the *immigrans-tripunctata* radiation clade. Nevertheless, some unexpected results obtained with the *Ddc* partition for the *virilis-repleta* radiation, as for example the branching positions of *D. camargoi* and *D. virilis*, were not here further supported.

The subgenus *Drosophila* total evidence tree (TET) recovers, although with a tenuous support (PP of 0.71), the monophyly of the clade encompassing all *Drosophila* other than *Sophophora*. In this case, *D. busckii* (*Dorsilopha* subgenus) and *S. leonensis* (*Samoaia* genus) are placed as the first to branch off next to *Sophophora*, so that the subgenus *Drosophila* becomes paraphyletic in respect to the genera *Hirtodrosophila*, *Liodrosophila*, *Scaptomyza* and *Zaprionus* (PP of 0.93). The clade so defined furthest splits off to divide into three well supported main lineages: the Hawaiian Drosophilidae radiation clade (PP of 1.00); the *virilis-repleta* radiation clade (PP of 1.00); and the *immigrans-tripunctata* radiation clade (PP of 1.00). The Hawaiian

Drosophilidae comprise the sister taxon of the *virilis-repleta* radiation (PP of 1.00), and the clade so established presents *H. pictiventris* as closest outgroup (PP of 0.68). The *immigrans-tripunctata* radiation, on the other hand, presents the clustering of the genera *Zaprionus* and *Liodrosophila* as sister taxon (PP of 0.73).

The historic scenario presented within the *virilis-repleta* radiation clade in this TET is widely congruent with the hypothesis provided by the *Amd* data alone. The only exceptions are the clustering of the *flavopilosa* and *annulimana* groups as sister clades (PP of 1.00), and the relationships within the *D. repleta* species group, which better resemble *Ddc* results.

Within the TET *immigrans-tripunctata* radiation clade, *D. immigrans* (*immigrans* group) first separates from the other species (PP of 1.00), and then *D. phalerata* (*quinaria* group) and *D. funebris* (*funebris* group) independently leave the remaining group of species, so that the *tripunctata* radiation becomes monophyletic (PP of 1.00). This *tripunctata* radiation clade further divides into three main lineages: (1) the first and basal one presenting *D. pallidipennis* (*pallidipennis* group) embedded among members of the *tripunctata* group (*D. bandeirantorum*, *D. cuaso*, *D. mediopunctata*, *D. paraguayensis*, *D. tripunctata* and *D. unipunctata*) (PP of 1.00); (2) the second recovering the monophyletic *cardini* and *guarani* groups as sister lineages (PP of 0.50); (3) the third showing *D. nappae* (*tripunctata* group) as the oldest lineage, followed by the clade grouping *D. crocina* and *D. mediodstriata* (*tripunctata* group) as sister to the clade encompassing the monophyletic *guaramunu* group joined to *D. mediopicta* (*tripunctata* group) and further to *D. mediodiffusa*, *D. mediopictoides* and *D. paramediodstriata* (*tripunctata* group) (PP of 0.67).

In general, the BA and ML strategies performed significantly better than the unweighted and weighted MP and than the NJ bootstrap analyses, both under the Templeton and the Shimodaira-Hasegawa Tests. Although this result is mainly due to the poor resolution presented at the base of the MP and NJ topologies, some striking differences were encountered between these and the BA, regardless of the general concordance found between the MrBayes, the BayesPhylogenies and the PhyML derived trees. The ML main differences refer to the branching of *S. leonensis* as sister to *D. immigrans* (only 22% bootstrap), the clustering of *D. funebris* and *D. phalerata* (70% bootstrap), a change in the relative position of the *D. pallidipennis/D. bandeirantorum* clade, which is placed as the *D. paraguayensis/D. cuaso* clade sister lineage (47% bootstrap), and the shifting relationships of *D. nappae*, which appears as sister to the monophyletic *guarani* group (45% bootstrap). Some of these relationships, particularly those which refer to the *D. pallidipennis/D. bandeirantorum* clade (PP of 0.95) and to *D. nappae* (PP of 0.60), were also recovered by the BayesPhylogenies

methodology of *a posteriori* partitioning the data sets. The MP unweighted and weighted topologies, besides clustering *D. phalerata* and *D. funebris* (83% and 78% bootstrap, respectively), presented *D. mediopicta* as the early offshoot within the clade that joins the *guaramunu* group with *D. mediodiffusa*, *D. mediopictoides* and *D. paramediostriata* (60% and 59% bootstrap, respectively), whereas *D. nappae* is placed within the *cardini/guarani* groups clade (59% and 51% bootstrap, respectively). The MP unweighted topology also presented some alternative results concerning the *virilis-repleta* radiation clade, since it grouped *D. virilis* and the *D. robusta/D. melanica* lineage with the *flavopilosa/annulimana* groups' clade (56% bootstrap). This placing of *D. virilis*, *D. robusta* and *D. melanica* was likewise recovered through NJ (96% bootstrap), which also presented some of the above depicted alternative relationships for *S. leonensis* (53% bootstrap), *D. phalerata* and *D. funebris* (100% bootstrap) and *D. mediopicta* (100% bootstrap). Within the clade that joins the *cardini* and *guarani* groups with *D. nappae*, NJ also places the *D. crocina/D. mediostriata* cluster (69% bootstrap).

Some striking results were also obtained by testing the three different Bayesian topologies obtained through the *Amd* and *Ddc* partitions taken individually and in combination under the Templeton and Shimodaira-Hasegawa Tests. Overall, the TET was not rejected by any of the three data partitions, although *Amd* rejected the *Ddc* topology while *Ddc* rejected the *Amd* one. The concatenated dataset, nevertheless, presented the *Amd* and *Ddc* topologies as significantly worse than the TET under the Templeton Test, but only the *Ddc* topology was regarded as significantly worse than the others under the Shimodaira-Hasegawa Test. Moreover, as assessed by the Partition Bremer Support (PBS) average values (Table 4), both *Amd* and *Ddc* partitions largely contribute to the combined analysis tree, although standardized support was twice as high for *Amd* as for *Ddc*. In this way, *Ddc* lends support to 43 out of the 75 nodes presented by the TET and conflicts with other 28, while *Amd* supports 67 nodes, conflicting with only four of the others. Nevertheless, 17 out of the 32 conflicting nodes present PBS values under 1.0, implying only a short amount of incongruence. So, in general, our results favor the TET as the best description so far available of the evolutionary history that led to the subgenus *Drosophila* present composition.

Utilizing this TET to estimate the divergences times within the subgenus *Drosophila* produces the results presented in Table 5. Although most of the credibility intervals are very large, the basal split between the subgenus *Sophophora* and the other species of the genus *Drosophila* falls near the K-T boundary, at about 55 My. The three major radiations that comprise the subgenus *Drosophila*, that is, the *immigrans-tripunctata*, the *virilis-repleta* and the Hawaiian *Drosophilidae* radiations, branched more recently, at about 42, 35 and 31 My, respectively.

### 3.3 – Refining the *immigrans-tripunctata* radiation relationships

In general the previous TET constituted a reasonable description of the evolution within the subgenus *Drosophila*, but it retained some incongruities and uncertainties concerning relationships within the *immigrans-tripunctata* radiation. In order to refine the conclusions about the evolution of this group of species, we decided to add new partitions to the analyses, specially evolutionary less restricted ones. The loci chosen were *Hb* and *COII*, which were joined to *Amd* and *Ddc* to construct a concatenated matrix.

Although the PHT results revealed some heterogeneity between the partitions (Table 6), specially as regards the comparisons involving the *COII* data set, a combined analysis was performed. The total evidence Bayesian tree derived through the inclusion of the four data partitions in a concatenated data set is depicted in Figure 5. This topology is widely similar to the trees recovered by *Amd* and *Ddc* individual and simultaneous analyses, yielding additional support to those conclusions.

In this way, the total evidence *immigrans-tripunctata* Bayesian phylogeny tightly recovers the *immigrans* group as the early offshoot (PP of 1.00), followed by the clade joining the sister groups *funebri* and *quinaria* (PP of 1.00). The monophyletic *tripunctata* radiation is the next to branch off, and it recovers a *tripunctata* group entirely paraphyletic with respect to the *pallidipennis*, *cardini*, *guarani*, *calloptera* and *guaramunu* groups (PP of 1.00). Within the *tripunctata* radiation cluster, three main independent lineages are recovered, whose general appearance concur with the concatenated *Amd* + *Ddc* hypothesis (Fig. 4). These lineages are named here according to the species described earlier that is included in each of them: the basal *tripunctata* lineage (PP of 1.00), and the sister *cardini* (PP of 0.84) and *mediostriata* (PP 1.00) lineages.

Within the *tripunctata* lineage *D. tripunctata* is the first to branch off (PP of 0.95), and its sister group branches off to divide into three different robust clades (each with a PP of 1.00), which are included in a nested succession: the *D. mediopunctata*/*D. unipunctata* clade, the *D. cuaso*/*D. paraguayensis* clade and the *D. bandeirantorum*/*D. pallidipennis* one. One recurrent result presented within this lineage is the inclusion of *D. pallidipennis*, a member of the *pallidipennis* group, which appears embedded among members of the *tripunctata* group subgroups II, III and IV. Concerning the results presented by the subgenus *Drosophila* Bayesian TET, the relative positions between the three major clades that comprise the *tripunctata* lineage, and between these and *D. tripunctata* are altered.

The *cardini* lineage, on the other hand, reflects better the taxonomic designations, since it recovers the *cardini* and *guarani* groups as two monophyletic sister taxa (each with a PP of 1.00). Within the *guarani* group *D. guaru* and *D. subbadia* are sister species, to the exclusion of *D. ornatifrons* (PP of 1.00). Nevertheless, within the *cardini* group the *cardini* subgroup is paraphyletically defined with respect to the *dunni* subgroup. In this way, *D. cardini* branches off early (PP of 1.00) while the remainder members of the group split off into two sister clades: one encompassing the two members of the *dunni* subgroup, that is, *D. arawakana* and *D. nigrodunni* (PP of 1.00); the other presenting *D. neocardini* closer to *D. cardinoides* than to *D. polymorpha* (PP of 0.89). So, the unstable positioning of *D. neocardini* underlines the major difference regarding the *cardini* lineage recovered here and that obtained through *Amd* and *Ddc* combined analyses.

In the *mediostriata* lineage, *D. nappae* (*tripunctata* group) is defined as the sister taxon of *D. ornatipennis* (*calloptera* group) (PP of 0.90), and this chimeric cluster was basal in relation to a major clade that joins representatives of the *tripunctata* group subgroups III and IV with the *guaramunu* group (PP of 1.00). Among these, the divergence of the clade grouping the cryptic *D. mediostriata* and *D. crocina* species is shown to predate that between the clade encompassing the *guaramunu* group plus *D. mediopicta* and the clade comprising *D. paramediostriata*, *D. mediodiffusa* and *D. mediopictoides* (PP of 1.00). Differently from the *Amd* and *Ddc* combined analyses, however, *D. paramediostriata* is monophyletic here.

The relationships presented by the ML tree fit perfectly the Bayesian hypothesis. The BayesPhylogenies topology also presents high levels of congruence with the MrBayes one, except for two minor differences encountered within the *mediostriata* lineage: *D. mediopicta* clusters with the *D. paramediostriata*, *D. mediodiffusa*, *D. mediopictoides* clade while *D. nappae* and *D. ornatipennis* are presented in a nested succession. Among the MP and NJ results, again only the relationships presented at the tips of the tree are resolved. In this way, the MP and NJ trees assign *D. mediopicta* as sister to the clade encompassing the *guaramunu* group and the *D. paramediostriata*, *D. mediodiffusa*, *D. mediopictoides* species. Besides this, NJ also places *D. polymorpha* as sister to the *dunni* subgroup representatives. As expected, these polytomous MP and NJ bootstrap topologies performed significantly worse than the others under the Templeton and Shimodaira-Hasegawa Tests.

Comparing the Bayesian topologies obtained for the *immigrans-tripunctata* radiation through *Amd*, *Ddc*, *COII* and *Hb* individual analyses with the total evidence one ensured that the combined tree is a reliable evolutionary hypothesis. This topology was not rejected by the Templeton Test or by the Shimodaira-Hasegawa Test under

any of these partitions, either taken individually or in combination. Although the Templeton and Shimodaira-Hasegawa Tests assigned the *Hb* Bayesian topology (Supplementary Material – Figure 1) as significantly different from the TET under all data sets except *Hb* alone, this partition holds support to 25 out of the 36 nodes presented by the *immigrans-tripunctata* radiation TET, while it presents negative PBS values for only two of its nodes (Table 7). Also *COII* lends a significant amount of support to this tree, agreeing with 21 and conflicting with only three of its nodes (Table 7). So, besides presenting relationships which are largely congruent with the *Amd* + *Ddc* simultaneous analysis for the entire subgenus *Drosophila*, the *immigrans-tripunctata* radiation TET is also widely supported by the *Hb* and *COII* gene partitions.

## 4. Discussion

### 4.1 – The best hypotheses

As significant deviations from stationarity were detected in our data sets, ML and BA derived hypotheses were considered as the best ones, since they are known to perform better under this kind of bias (Galtier and Gouy, 1995 and 1998). Besides, as combining data from several loci, even incongruent ones, is a way of attempting to amplify signal over noise (Baker and DeSalle, 1997; Remsen and DeSalle, 1998), total evidence hypotheses were also emphasized over individual ones.

Hence, the *Amd* and *Ddc* were combined in order to provide a reliable picture of the subgenus *Drosophila* evolution, since these genes have previously proved to be informative at this level of divergence (Robe et al., 2005; Tatarenkov and Ayala, 2001; Tatarenkov et al., 1999a and 2001). *Amd* and *Ddc* are paralogous genes, functionally related to catecholamine metabolism and physically separated by about 2.5 Kb, that have arisen from a duplication occurred long before the split between Lepidoptera and Diptera (Tatarenkov et al., 1999b). The PHT performed between *Amd* and *Ddc* was not able to detect significant heterogeneity, probably because of the fact that being linked genes, they share identical evolutionary histories. Furthermore, as *Amd* and *Ddc* were shown to share correlated evolutionary rates and model parameters, even their internal homoplasies should present a certain degree of correlation. Thus, neither *Amd* nor *Ddc* reject their simultaneous TET (Fig. 4), but instead they provide a significant amount of support to its nodes (Table 4). Indeed, this TET can be considered as the best phylogenetic hypothesis, so far rendered available for the subgenus *Drosophila* as a whole.

According to Zilversmit et al. (2002), enzyme phylogenies are mostly resolved at the base, while transcription factors derived trees show the most support for the

nodes at the tips. This suggests that the most effective approach consists in combining different data sets, namely enzyme and transcription factors ones. In this way, *Hb*, which is a transcription factor coding marker member of the gap class of segmentation genes (Treier et al., 1989) was joined to *Amd*, *Ddc*, and *COII* enzyme coding genes, in order to improve some less resolved results concerning the *immigrans-tripunctata* radiation relationships. In this case, although the PHT results reveal significant conflict between some of our data sets, none partition rejected the *immigrans-tripunctata* TET (Fig. 5). Apart from this, all of them provided significant support to most of the TET nodes. So, this combined hypothesis is justifiably considered as the best description so far available for the *immigrans-tripunctata* radiation phylogeny.

#### 4.2 – The subgenus *Drosophila* phylogeny

In general, this work adds to the existing evidences that point to a genus *Drosophila* entirely paraphyletic as regards other genera like *Hirtodrosophila*, *Liodrosophila*, *Samoaia*, *Scaptomyza* and *Zaprionus*. Throckmorton (1975) also supported the scenario in which these and other genera derive from the genus *Drosophila* basis, although Grimaldi (1990) did not agree with this notion. Further molecular evidences better reflected Throckmorton's (1975) results, agreeing with our data when recovering the genus *Drosophila* paraphyly with respect to different taxa as: *Hirtodrosophila* (DaLage et al., 2007; Kwiatowski et al., 1997; Kwiatowski and Ayala, 1999; Remsen and DeSalle, 1998; Robe et al., 2005; Tatarenkov et al., 1999a and 2001), *Liodrosophila* (DaLage et al., 2007; DeSalle, 1992a; Katoh et al., 2000; Robe et al., 2005; Tatarenkov et al., 1999a and 2001), *Samoaia* (Pélandakis and Solignac, 1993; Robe et al., 2005; Tatarenkov et al., 1999a e 2001), *Scaptomyza* (DaLage et al., 2007; DeSalle, 1992b; Katoh et al., 2000; Kwiatowski and Ayala, 1999; Remsen and DeSalle, 1998; Remsen and O'Grady, 2002; Russo et al., 1995; Tatarenkov et al., 1999a and 2001), and *Zaprionus* (DaLage et al., 2007; Katoh et al., 2000; Kwiatowski et al., 1997; Kwiatowski and Ayala, 1999; Pélandakis and Solignac, 1993; Remsen and DeSalle, 1998; Remsen and O'Grady, 2002; Robe et al., 2005; Russo et al., 1995; Tatarenkov et al., 1999a and 2001; Yotoko et al., 2003). All these data reflect the catchall property that the term *Drosophila* has acquired.

Our phylogenies also present strong support for *Sophophora* being basal to the other *Drosophila*, in agreement with Throckmorton (1975) and more recent molecular data (DaLage et al., 2007; Katoh et al., 2000; Kwiatowski et al., 1997; Kwiatowski and Ayala, 1999; Remsen and DeSalle, 1998; Remsen and O'Grady, 2002; Robe et al., 2005; Russo et al., 1995; Tatarenkov et al., 1999a and 2001), but contrasting with Grimaldi (1990) and DeSalle (1992a). According to our dating estimates, the split



between *Sophophora* and the other species of *Drosophila* occurred near the K-T boundary, about 55 Mya, in agreement with Tamura et al. (2004). Furthermore, the TET supports the subgenus *Dorsilopha* and the genus *Samoaia* as the next to branch off following *Sophophora*, a result incongruent with the derived branching presented for these by Throckmorton (1975), but congruent with Robe et al. (2005) and Tatarenkov et al. (1999a and 2001) as regards *Dorsilopha*, but not *Samoaia*. In this case, *Dorsilopha* diverged from the other species about 52.85 Mya, whereas the branching of *Samoaia* occurred at approximately 49 Mya.

Otherwise, *Hirtodrosophila*, *Scaptomyza*, *Liodrosophila* and *Zaprionus* occupy more derived positions in our phylogenies, branching inside a paraphyletic subgenus *Drosophila* clade, which further branches off to divide into three major lineages: (1) the Hawaiian Drosophilidae radiation, arisen about 30.75 Mya; (2) the *virilis-repleta* radiation, began at approximately 34.61 Mya; (3) the *immigrans-tripunctata* radiation, occurred about 41.74 Mya. In this case, the clade joining *Scaptomyza* with the Hawaiian species of *Drosophila* and the *Hirtodrosophila* lineage define the *virilis-repleta* radiation sister taxons, whereas *Liodrosophila* and *Zaprionus* are clustered as the *immigrans-tripunctata* radiation sister lineage. The separation between these two major clade is estimated to have occurred about 44.78 Mya.

Throckmorton (1975) recognized the existence of two major radiations within the subgenus *Drosophila*: the *immigrans-Hirtodrosophila* and the *virilis-repleta* ones. According to this author, *Hirtodrosophila*, the Hawaiian Drosophilidae and *Zaprionus* take part of the *immigrans-Hirtodrosophila* radiation, while *Liodrosophila* is basal in relation to the *virilis-repleta* and *immigrans-Hirtodrosophila* radiations. Grimaldi (1990), on the other hand, clusters *Zaprionus* and *Liodrosophila* as members of the genus *Drosophila* sister clade, while *Scaptomyza* and the clade grouping the genera *Hirtodrosophila* and *Idiomyia* (Hawaiian species of *Drosophila*) are placed even more distantly. DeSalle (1992b), nevertheless, suggests that *Scaptomyza* and the Hawaiian species of *Drosophila* cluster as the subgenus *Drosophila* sister clade, given that the so defined clade presents *Liodrosophila* as closest outgroup (DeSalle, 1992a). At the same time, however, DeSalle (1992a) presents *Hirtodrosophila* and *Zaprionus* in a nested succession outside the genus *Drosophila*.

Nowadays, most evidences agree with our data in that the Hawaiian species of *Drosophila* and *Scaptomyza* form a monophyletic group which is sister to the *virilis-repleta* radiation (Kato et al., 2000; Kwiatowski and Ayala, 1999; Remsen and DeSalle, 1998; Remsen and O'Grady, 2002; Russo et al., 1995; Tatarenkov et al., 1999a and 2001). The picture concerning the *Hirtodrosophila* positioning is nevertheless a little different. Although the clustering between this genus and the

members of the *virilis-repleta* radiation plus Hawaiian Drosophilidae had previously been reported by Tatarenkov et al. (1999a and 2001), this should be interpreted with caution, since neither these nor our topology present strong confidence for this relationship. Besides, among most of the evidences that support *Hirtodrosophila* as a derived taxon within *Drosophila*, it is either favored as one of the first to branch off next to *Sophophora* (DaLage et al., 2007; Remsen and DeSalle, 1998), or it is connected to the *immigrans-tripunctata* radiation species (Kwiatowski and Ayala, 1999; Robe et al., 2005; Throckmorton, 1975).

The clustering of *Liodrosophila* and *Zaprionus* is another striking result presented by this study. Although previously reported by Grimaldi (1990), Robe et al. (2005) and Tatarenkov et al. (2001), this relationship attained an unprecedented level of support in our TET. *Liodrosophila* was previously reported as one outgroup to the subgenus *Drosophila* (DaLage et al., 2007; DeSalle, 1992a; Throckmorton, 1975), and its positioning within the subgenus *Drosophila* was only weakly supported by Tatarenkov et al. (1999 and 2001). At the same time, the phylogenetic affinity between *Zaprionus* and the *immigrans-tripunctata* radiation species, as suggested by Throckmorton (1975), received only a tenuous molecular support till now in Robe et al. (2005) and in some topologies presented by Kwiatowski and Ayala (1999) and Tatarenkov et al. (1999a and 2001). Most evidences place *Zaprionus* as one of the subgenus *Drosophila* closest outgroups (DaLage et al., 2007; Remsen and DeSalle, 1998; Remsen and O'Grady, 2002; Russo et al., 1995; Yotoko et al., 2003).

Concerning the subgenus *Drosophila* species subdivision, although Robe et al. (2005) provided additional support for Throckmorton's (1975) *immigrans-Hirtodrosophila* and *virilis-repleta* radiations, the data presented by Kwiatowski and Ayala (1999), Remsen and DeSalle (1998), Remsen and O'Grady (2002) and Tatarenkov et al. (2001) better conform our three radiation hypothesis. In this case, although *Ddc* individual Bayesian analysis recovered a clade that, except for the inclusion of *Liodrosophila* and *D. funebris*, and for the exclusion of *Dorsilopa* and *D. immigrans*, intimately resembled Throckmorton's (1975) *immigrans-Hirtodrosophila* radiation, this result was not further supported in our TET. Instead, our combined results agree with Remsen and O'Grady's (2002) *immigrans-tripunctata* lineage, since besides including *D. funebris*, it also excludes the Hawaiian Drosophilidae and the *Dorsilopa*, *Hirtodrosophila* and *Samoaia* taxa.

As regards our TET *virilis-repleta* radiation, *D. polychaeta* is the early offshoot, in accordance with previous studies (DaLage et al., 2007; Tatarenkov and Ayala, 2001; Throckmorton, 1975; Wang et al., 2006). The clade joining *D. robusta* with *D. melanica* is the next to branch off, whereas *D. virilis* clusters with the clade comprising the

monophyletic *repleta* radiation, the *bromeliae/nannoptera* and the *annulimana/flavopilosa* sister groups lineages, a pattern largely congruent with Tatarenkov and Ayala (2001). In this case, although the clustering of *D. robusta* and *D. melanica* seem to be well established (Remsen and O'Grady, 2002; Throckmorton, 1975; Wang et al., 2006), the *D. virilis* branching position agrees with some previous studies (DaLage et al., 2007; Remsen and DeSalle, 1998; Remsen and O'Grady, 2002) but not with others (Kato et al., 2000; Robe et al., 2005; Tatarenkov et al., 2001; Throckmorton, 1975; Wang et al., 2006). The positioning of the *annulimana* group with the *flavopilosa* group conflicts with Throckmorton (1975), according to which it branches off earlier, whereas the clustering of the *nannoptera* and *bromeliae* groups entirely agrees with his work. The derived *repleta* radiation as suggested by Throckmorton (1975) is also recovered here, although the proposed *dreyfusi* and *canalinae* sister groups relationship was not encountered. At less inclusive levels, our data support the subdivision of the *mesophragmatica* group into three main lineages [(1) the *D. viracochi* lineage; (2) the *D. pavani/D. gaucha* lineage; (3) and the *D. gasici/D. brncici/D. mesophragmatica* lineage], in agreement with previous studies (Brncic et al., 1971; Brncic and Koref-Santibañez, 1957; Mota et al., in press; Nair et al., 1971). The *flavopilosa* group internal subdivision is also recovered, at least concerning the species here analyzed, since *D. incompta* (*nesiota* subgroup) is the furthest, whereas *D. cestri*, *D. cordeiroi* and *D. flavopilosa* (*flavopilosa* subgroup) are closely related.

In contrast with previous molecular evidences (DaLage et al., 2007; Remsen and O'Grady, 2002; Robe et al., 2005; Yotoko et al., 2003) the *immigrans-tripunctata* radiation relationships are well resolved in this study. In this sense, the basal divergence of *D. immigrans* is here recognized, a result highly congruent with previous analyses (Remsen and O'Grady, 2002; Robe et al., 2005; Throckmorton, 1975; Yotoko et al., 2003). The *quinaria*, *funebri* and probably *testacea* groups are the next to branch off, while the remaining clade defines a monophyletic *tripunctata* radiation, in agreement with DaLage et al. (2007), Robe et al. (2005) and Throckmorton (1975). The particular relationships presented by the *quinaria*, *funebri*, and *testacea* groups are, nevertheless, a little conflicting. These groups are presented in a nested succession in the *Amd* and *Amd + Ddc* phylogenies, whereas *Ddc* and the *immigrans-tripunctata* TET favor *quinaria* and *funebri* as sister groups. Notwithstanding, in DaLage et al. (2007) the *quinaria*, *funebri* and *testacea* groups are clustered together, with *quinaria* as the early offshoot, while in Robe et al. (2005) the *funebri* and *testacea* groups are sister taxa, more closely related to the *tripunctata* radiation clade than to the *quinaria* group. Remsen and O'Grady (2002), on the other hand, present the *quinaria*, *funebri* and

*testacea* lineages independently embedded within the *tripunctata* radiation, making it a paraphyletic taxon. The *funebri* and *testacea* groups also define a paraphyletic *tripunctata* radiation in Yotoko et al. (2003).

Within the *tripunctata* radiation clade recovered in our TET's three main lineages are defined, all of which occurred at about 18.5 Mya: (1) the *tripunctata* lineage; (2) the *cardini* lineage; (3) the *mediostriata* lineage. Contrary to previous studies (Robe et al., 2005; Yotoko et al., 2003), relationships are well defined across and within these lineages, although the short length of their internal basal branches, and their superimposed estimated divergence times agree with the rapid radiation hypothesis first presented by Throckmorton (1962). Also congruent with previous morphological and molecular evidences (DaLage et al., 2007; Frota-Pessoa, 1954; Kastriasis, 1969; Robe et al., 2005; Yotoko et al., 2003) is the absence of support to the *tripunctata* group monophyly in all the recovered trees, where members of this group appear intermingled among members of the *pallidipennis*, *guarani*, *cardini*, *calloptera* and *guaramunu* groups. Moreover, the position of the *tripunctata* group species along the tree do not agree with Frota-Pessoa's (1954) subdivision, since members of different subgroups are scattered across the *tripunctata* radiation clade. On the other hand, the monophyly of the *cardini*, *guarani* and *guaramunu* groups is largely recovered, at least with regard to the species here sampled. Concerning the *cardini* group relationships, the *cardini* subgroup was always paraphyletically defined with respect to the *dunni* subgroup, a result also presented by Brisson et al. (2006). The *cardini* group internal relationships were, however, highly unstable, mainly because of the *D. neocardini* floating positioning. In general, after the splitting of *D. cardini*, two main clades arise: (1) the *dunni* subgroup; (2) the clade joining *D. polymorpha*, *D. neocardini* and *D. cardinoides*. As this kind of clustering had not been previously reported, which is surprising taking into account the large number of alternative evolutionary hypotheses presented for this group (Brisson et al., 2006; Heed, 1962; Heed and Russell, 1971; Napp and Cordeiro, 1971), the *cardini* group may represent a seed to the study of complex evolution, where hybridization, ancestral polymorphisms and explosive radiations appear to have acted together to produce its present day diversity.

Another striking result presented by our analyses is in agreement with Tamura et al. (2004) in that 68% of the recovered divergence times are clustered close to periods of major climatic change, specially in paleoclimatic cooling periods that occurred throughout the Cenozoic. Hence, only 22 of the 75 nodes presented in our subgenus *Drosophila* TET are dated to near the Middle Oligocene and to the Early Miocene (35 to 15 Mya), known to be stable climatic epochs. Nevertheless, 21 of these

22 nodes refer exactly to the basal branching points within the three recovered radiations (see Figure 4). So, it is likely that a pattern different from that reported by Tamura et al. (2004) acted in the subgenus *Drosophila* basal evolution. Maybe some of these diversification events were allowed by the expansion to the New World, correlated to the exploration of new and different niches, which led to some kind of complex evolution where polymorphic ancestral populations rapidly and successively subdivided (Throckmorton, 1962). It is particularly interesting that both the *virilis-repleta* and the *immigrans-tripunctata* radiations share parallel evolutionary patterns, in which mainly Asian groups comprise the early offshoots (represented by the *polychaeta* and the *immigrans* groups, respectively), followed by the intermediate branchings of primarily Holarctic groups (represented by the *robusta*, *melanica* and *virilis* groups in the *virilis-repleta* radiation, and by the *quinaria*, *funebri* and *testacea* groups in the *immigrans-tripunctata* radiation), whereas the derived clades comprise Neotropical exclusive species. Thus, it is likely that similar historical events undertook both scenarios. Nevertheless, only the accrue of additional studies including new markers and new species will approach the real picture, proving this hypothesis true or false, and improving its details.

### **Acknowledgments**

We are grateful to Dr. Fabiano Pimentel Torres, MSc. Larissa Heinzemann, M.Sc. Newton Medeiros Vidal and M.Sc. Paulo Marcos Pinto for their help in cloning strategies; to Dr. Fabiano Pimentel Torres and MSc. Paloma Rubin for their assistance in the sequencing reactions; to Dr. Jean David, M.Sc. Adriana Ludwig, Dra. Daniela Cristina De Toni, M.Sc. Jonas da Silva Döge, Dr. Luciano Basso da Silva, M.Sc. Marco Silva Gottschalk, M.Sc. Marília Fantinel D'Avila and Dra. Myriam Budnik for the *Drosophila* species supply; to M.Sc. Adriana Ludwig, M.Sc. Hermes Schmitz and Nina Roth Mota for kindly reading the paper; and to Mr. Félix Nonnenmacher for the English language revision. We also thank the Brazilian agencies CNPq, FAPERGS, CAPES and PROPESQ-UFRGS for providing grants, fellowships, and research funding.

### **References**

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

- Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Contr. 19, 716-723.
- Bächli, G., 2008. TaxoDros: The Database on Taxonomy of Drosophilidae, v.1.03, Database 2008/01. <http://taxodros.unizh.ch/>. Last accessed on 08/01/2008.
- Baker, R.H., DeSalle, R., 1997. Multiple sources of character information and the phylogeny of Hawaiian drosophilids. Syst. Biol. 46, 654-673.
- Bremer, K., 1988. The limits of amino acid sequence data in angiosperm phylogenetics reconstruction. Evolution 42, 795-803.
- Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295-304.
- Brisson, J.A., Wilder, J., Hollocher, H., 2006. Phylogenetic analysis of the *cardini* group of *Drosophila* with respect to changes in pigmentation. Evolution 60, 1228-1247.
- Brncic, D., Koref-Santibañez, S., 1957. The *mesophragmatica* species group of *Drosophila*. Evolution 11, 300-310.
- Brncic, D., Nair, P.S., Wheeler, M.R., 1971. Cytotaxonomic relationships within the *mesophragmatica* species group of *Drosophila*. Univ. Texas Publ. 7103, 1-16.
- Clark, A.G., et al., 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. Nature 450, 203-218.
- DaLage, J.L., Dergoat, G.J., Maczkowiak, F., Silvain, J.F., Cariou, M.L., Lachaise, D., 2007. A phylogeny of Drosophilidae using the *Amyrel* gene: questioning the *Drosophila melanogaster* species group boundaries. J. Zool. Syst. Evol. Res. 45, 47-63.
- DeSalle, R. 1992a., The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. Mol. Phylogenet. Evol. 1, 31-40.
- DeSalle, R. 1992b., The origin and possible time of divergence of the Hawaiian Drosophilidae: Evidence from DNA sequences. Mol. Biol. Evol. 9, 905-916.

- Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113.
- Eveleth, D.D., Marsh, J.L., 1986. Evidence for evolutionary duplication of genes in the dopa decarboxylase region of *Drosophila*. *Genetics* 114, 469-483.
- Farris, J.S., Källersjö, M., Kluge, A., Bult, C., 1994. Testing significance of congruence. *Cladistics* 10, 315-320.
- Farris, J.S., Källersjö, M., Kluge, A., Bult, C., 1995. Constructing a significance test for incongruence. *Syst. Biol.* 44, 570-572.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using bootstrap. *Evolution* 39, 783-791.
- Frota-Pessoa, O. 1954. Revision of the *tripunctata* group of *Drosophila* with description of fifteen new species (Drosophilidae, Diptera). *Arq. Mus. Paranaense* 10, 253-304.
- Galtier, N., Gouy, M., 1995. Inferring phylogenies from DNA sequences of unequal base compositions. *Proc. Natl. Acad. Sci.* 92, 11317-11321.
- Galtier, N., Gouy, M., 1998. Inferring pattern and process: maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. *Mol. Biol. Evol.* 15, 871-879.
- Grimaldi, D.A., 1987. Amber fossil Drosophilidae (Diptera), with particular reference to the Hispaniolan taxa. *Am. Mus. Novitates* 2880, 1-23.
- Grimaldi, D.A., 1990. A phylogenetic, revised classification of the genera in the Drosophilidae (Diptera). *Bull. Am. Mus. Natl. Hist.* 197, 1-139.
- Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696-704.
- Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160-174.

- Heed, W.B., 1962. Genetic characteristics of island populations. Univ. Texas Publ. 6205, 173-206.
- Heed, W.B., Russell, J.S., 1971. Phylogeny and population structure in island and continental species on the *cardini* group of *Drosophila* studied by inversion analysis. Univ. Texas Publ. 7103, 91-130.
- Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754-755.
- Kastritsis, C.D., 1969. The chromosomes of some species of the *guarani* group of *Drosophila*. J. Heredity 60, 51-57.
- Katoh, T., Tamura, K., Aotsuka, T., 2000. Phylogenetic position of the subgenus *Lordiphosa* of the genus *Drosophila* (Diptera: Drosophilidae) inferred from Alcohol Dehydrogenase (*Adh*) gene sequences. J. Mol. Evol. 51, 122-130.
- Kumar, S., Gadagkar, S.R., 2001. Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. Genetics 158, 1321-1327.
- Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief. Bioinform. 5, 150-163.
- Kwiatowski, J., Ayala, F.J., 1999. Phylogeny of *Drosophila* and related genera: Conflict between molecular and anatomical analyses. Mol. Phylogenet. Evol. 13, 319-328.
- Kwiatowski, J., Krawczyk, M., Jaworski, M., Skarecky, D., Ayala, F.J., 1997. Erratic Evolution of Glycerol-3-Phosphate Dehydrogenase in *Drosophila*, *Chymomyza*, and *Ceratitis*. J. Mol. Evol. 44, 9-22.
- Kwiatowski, J., Skarecky, D., Bailey, K., Ayala, F.J., 1994. Phylogeny of *Drosophila* and related genera inferred from the nucleotide sequence of the Cu, Zn *Sod* gene. J. Mol. Evol. 38, 443-454.



- Lockhart, P.J., Stell, M.A., Hendy, M.D., Penny, D., 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* 11, 605-612.
- Markow, T.A., O'Grady, P.M., 2006. Phylogenetic relationships of Drosophilidae. In: Markow, T.A., O'Grady, P.M. *Drosophila: A guide to species identification and use*. Academic Press, London, pp. 3-64.
- Mota, N.R., Robe, L.J., Valente, V.L.S., Budnik, M., Loreto, E.L.S., (in press) Phylogeny of the *Drosophila mesophragmatica* group (Diptera, Drosophilidae): an example of Andean evolution. *Zool. Sci.*
- Nair P.S., Brncic D., Kojima K.I., 1971. II. Isozyme variations and evolutionary relationships in the *mesophragmatica* species group of *Drosophila*. *Stud. Genet.* VI, 17-28.
- Napp, M., Cordeiro, A.R., 1981. Inter-specific relationships in the *cardini* group of *Drosophila* studied by electrophoresis. *Rev. Brasil. Biol.* 4, 537-547.
- Nylander, J.A.A., 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Center, Uppsala University.
- Pagel, M., Meade, A., 2004. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Syst. Biol.* 53, 571-581.
- Pélandakis, M., Solignac, M., 1993. Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. *J. Mol. Evol.* 37, 525-543.
- Posada, C., Crandall, K.A., 1998. MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14, 817-818.
- Powell, J.R., DeSalle, R., 1995. *Drosophila* molecular phylogenies and their uses. *Evol. Biol.* 28, 87-138.
- Remsen, J., DeSalle, R., 1998. Character congruence and multiple data partitions and the origin of the Hawaiian Drosophilidae. *Mol. Phylogenet. Evol.* 9, 225-235.

- Remsen, J., O'Grady, P.O., 2002. Phylogeny of Drosophilinae (Diptera: Drosophilidae) with comments on combined analysis and character support. *Mol. Phylogenet. Evol.* 24, 249-264.
- Robe, L.J., Valente, V.L.S., Budnik, M., Loreto, E.L.S., 2005. Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: a nuclear versus mitochondrial gene approach. *Mol. Phylogenet. Evol.* 36, 623-640.
- Rodriguez, F., Oliver, J.L., Marin, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. *J. Theor. Biol.* 142, 485-501.
- Russo, C.A.M., Takezaki, N., Nei, M., 1995. Molecular phylogeny and divergence times of Drosophilid species. *Mol. Biol. Evol.* 12, 391-404.
- Rutschmann, F., 2005. Bayesian molecular dating using PAML/multidivtime. A step-by-step manual. University of Zurich, Switzerland. Available at <http://www.plant.ch/>.
- Sassi, A.K., Heredia, F., Loreto, E.L.S., Valente, V.L.S., Rohde, C., 2005. Transposable elements *P* and *gypsy* in natural populations of *Drosophila willistoni*. *Genet. Mol. Biol.* 28, 734-739.
- Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18, 502-504.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114-1116.
- Sorenson, M.D., Franzosa, E.A., 2007. TreeRot version 3. Boston University, Boston, MA.
- Staden, R., 1996. The Staden sequence analysis package. *Mol. Biotechnology* 5, 233-241.

Sturtevant, A.H., 1939. On the subdivision of the genus *Drosophila*. Proc. Natl. Acad. Sci. 25, 137-141.

Sturtevant, A.H., 1942. The classification of the genus *Drosophila* with the description of nine new species. Univ. Texas Publ. 4213, 5-51.

Swofford, D.L., 2003. PAUP: Phylogenetic Analysis using Parsimony (and other methods). Version 4. Sinauer Associates, Massachusetts.

Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer, Sunderland, pp. 407-514.

Tamura, K., Subramanian, S., Kumar, S., 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. Mol. Biol. Evol. 21, 36-44.

Tatarenkov, A., Ayala, F.J., 2001. Phylogenetic relationships among species groups of the *virilis-repleta* radiation of *Drosophila*. Mol. Phylogenet. Evol. 21, 327-331.

Tatarenkov, A., Kwiatowski, J., Sharecky, D., Barrío, E., Ayala, F.J., 1999a. On the evolution of *Dopa decarboxylase (Ddc)* and *Drosophila* systematics. J. Mol. Evol. 48, 445-462.

Tatarenkov, A., Sáez, A.G., Ayala, F.J., 1999b. A compact gene cluster in *Drosophila*: the unrelated *Cs* gene is compressed between duplicated *amd* and *Ddc*. Gene 231, 111-120.

Tatarenkov, A., Zurovcová, M., Ayala, F.J., 2001. *Ddc* and *Amd* sequences resolve phylogenetic relationships of *Drosophila*. Mol. Phylogenet. Evol. 20, 321-325.

Templeton, A.R., 1983. Convergent evolution and non-parametric inferences from restriction fragment and DNA sequence data. In: Weir, B. (ed.) Statistical Analysis of DNA Sequence Data. Marcel Dekker, New York, pp. 151-179.

Thorne, J.L., Kishino, H., 2002. Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* 51, 689–702.

Thorne, J.L., Kishino, H., Painter, I.S., 1998. Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* 15, 1647–1657.

Throckmorton, L.H., 1962. The problem of phylogeny in the genus *Drosophila*. *Stud. Genet. II. Univ. Texas Publ.* 6205, 207-343.

Throckmorton, L.H., 1975. The phylogeny, ecology and geography of *Drosophila*. In: King, R. C. (ed.), *Handbook of Genetics*. Plenum, New York, pp. 421-469.

Treier, M., Pfeifle, C., Tautz, D., 1989. Comparison of the gap segmentation gene *hunchback* between *Drosophila melanogaster* and *Drosophila virilis* reveals novel modes of evolutionary change. *EMBO J.* 8, 1517-1525.

Wang, B., Park, J., Watabe, H., Gao, J., Xiangyu, J., Aotsuka, T., Chen, H., Zhang, Y., 2006. Molecular phylogeny of the *Drosophila virilis* section (Diptera: Drosophilidae) based on mitochondrial and nuclear sequences. *Mol. Phylogenet. Evol.* 40, 484-500.

Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *BACIOS* 13, 555-556.

Yotoko, K.S.C., Medeiros, H.F., Solferini, V.N., Klaczko, L.B., 2003. A molecular study of the systematics of the *Drosophila tripunctata* group and the *tripunctata* radiation. *Mol. Phylogenet. Evol.* 28, 614-619.

Zilversmit, M., O'Grady, P., DeSalle, R., 2002. Shallow genomics, phylogenetics, and evolution in the family Drosophilidae. *Pac. Symp. Biocomput.*, 512-523.

## FIGURE LEGENDS

**Figure 1:** Structure of the *Amd*, *Ddc* and *Hb* genes, and strategy for amplification and sequencing. On the top of each figure, black arrows represent exons, while grey boxes

denote the introns. Just below these genomic maps, hatched boxes reflect the coding regions, and light grey arrows represent the different RNA precursors transcribed from different promoters. A) The *AmdEx4F* and the *Amd-bw* primers were used to amplify a region of 962 pb region from the *Amd* fourth exon. This gene encodes two different isoforms, A and B, derived from different transcripts (Adams et al., 2000). B) The *BPF* and *BPR* primers, both annealing in the *Ddc* fourth exon, generate a 1,184 pb amplicon. *Ddc* encodes three different isoforms, A, B and C, which originate through alternative transcription start sites and also through alternative splicing (Adams et al., 2000). C) *Hb106F* and *Hb903R* primers were used to amplify a region of variable length from the *hunchback* coding region. The deviations in size are due to the presence of a hypervariable region in the second half of the amplified fragments, which contrasts with the conserved first half that includes a Zinc-Finger region. Although two different precursor RNA's are transcribed, they present the same coding properties (Treier et al., 1989).

**Figure 2:** Majority-rule consensus trees obtained through the Bayesian inference analysis from the *Amd* data partition. Numbers at the nodes are the posterior probability estimates for each clade. Branch lengths are proportional to the scale given in substitutions per nucleotide. The taxonomic property of the most inclusive clades is indicated to the right, where quotation marks denote a paraphyletic status.

**Figure 3:** Majority-rule consensus trees obtained through the Bayesian inference analysis from the *Ddc* data partition. Numbers at the nodes are the posterior probability estimates for each clade. Branch lengths are proportional to the scale given in substitutions per nucleotide. The taxonomic property of the most inclusive clades is indicated to the right, where quotation marks denote a paraphyletic status.

**Figure 4:** Majority-rule consensus tree obtained through the Bayesian inference analysis from the *Amd* and *Ddc* combined data partition. Numbers to the left of the nodes are the posterior probability estimates for each clade, while numbers in small print to the right of each node are those referred to in Tables 4 and 5. Those numbers that are involved in a rectangle refer to nodes whose divergence is dated to the Oligocene and Early Miocene (35 to 15 Mya). Branch lengths are proportional to the scale given in substitutions per nucleotide. The taxonomic property of the most inclusive clades is indicated to the right, where quotation marks denote a paraphyletic status.

**Figure 5:** Majority-rule consensus tree obtained through the Bayesian inference analysis from the *Amd*, *Ddc*, *Hb* and *COII* combined data partition. Numbers to the left of the nodes are the posterior probability estimates for each clade, while numbers in small print to the right of each node are those referred to in Table 7. Branch lengths are proportional to the scale given in substitutions per nucleotide. The taxonomic property of the OTU's are indicated in the middle right of the figure, while a new proposal of internal subdivision for the *tripunctata* lineage is presented at the far right.

**Figure 1S:** Majority-rule consensus trees obtained through the Bayesian inference analysis from the *Hb* data partition. Numbers at the nodes are the posterior probability estimates for each clade. Branch lengths are proportional to the scale given in substitutions per nucleotide. The taxonomic property of the OTU's is indicated to the right.

Table 1

Taxonomic list of *Drosophila* taxa used in this study, with their source and GenBank accession numbers

Genus	Subgenus	Section	Group	Species	Origin	Accession numbers			
						<i>Amd</i>	<i>Ddc</i>	<i>Hb</i>	<i>COII</i>
<i>Drosophila</i>	<i>Drosophila</i>	<i>quinaria-tripunctata</i>	<i>guarani</i>	<i>D. ornatifrons_1</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444583</a>	<a href="#">EU446091</a>	<a href="#">EU447330</a>	AY162977
				<i>D. ornatifrons_2</i> <sup>d</sup>	Joinville, SC, Brazil	<a href="#">EU444584</a>	<a href="#">EU446090</a>	<a href="#">EU450643</a>	AY162978
				<i>D. subbadia</i> <sup>l</sup>	El Narango, México	<a href="#">EU444594</a>	<a href="#">EU446099</a>	<a href="#">EU447341</a>	AY847772
				<i>D. guaru</i> <sup>l</sup>	Joinville, SC, Brazil	<a href="#">EU444566</a>	<a href="#">EU446072</a>	<a href="#">EU447315</a>	AY847763
			<i>guaramunu</i>	<i>D. maculifrons_1</i> <sup>e</sup>	Santa Maria, RS, Brazil	<a href="#">EU444570</a>	<a href="#">EU446076</a>	<a href="#">EU447318</a>	AY847766
				<i>D. maculifrons_2</i> <sup>c</sup>	Florianópolis, SC, Brazil	<a href="#">EU444571</a>	<a href="#">EU446077</a>	<a href="#">EU447319</a>	AY162979
				<i>D. griseolineata_1</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444565</a>	<a href="#">EU446071</a>	<a href="#">EU447314</a>	AF478426
				<i>D. griseolineata_2</i> <sup>c</sup>	Florianópolis, SC, Brazil	-	<a href="#">EU446070</a>	<a href="#">EU447313</a>	-
			<i>tripunctata</i>	<i>D. tripunctata</i> <sup>l</sup>	Iowa River, Iowa, EUA	AF293728	AF324964	<a href="#">EU447342</a>	AF478432
				<i>D. nappae_1</i> <sup>a</sup>	Itapuã, RS, Brazil	<a href="#">EU444580</a>	<a href="#">EU446088</a>	<a href="#">EU447327</a>	AY162983
				<i>D. nappae_2</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444579</a>	<a href="#">EU446086</a>	<a href="#">EU447326</a>	-
				<i>D. nappae_3</i> <sup>b</sup>	Florianópolis, SC, Brazil	-	<a href="#">EU446087</a>	-	-
				<i>D. paraguayensis_1</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444588</a>	<a href="#">EU446094</a>	<a href="#">EU447334</a>	AY162986
				<i>D. paraguayensis_2</i> <sup>b</sup>	Florianópolis, SC, Brazil	<a href="#">EU444587</a>	<a href="#">EU446093</a>	<a href="#">EU447333</a>	AY162987
				<i>D. cuaso</i> <sup>b</sup>	Florianópolis, SC, Brazil	<a href="#">EU444563</a>	<a href="#">EU446068</a>	<a href="#">EU447311</a>	AY162984
				<i>D. roehrae</i> <sup>b</sup>	Florianópolis, SC, Brazil	<a href="#">EU444593</a>	-	<a href="#">EU447340</a>	-
				<i>D. mediopunctata_1</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444575</a>	<a href="#">EU446083</a>	<a href="#">EU447322</a>	AY162989
				<i>D. mediopunctata_2</i> <sup>a</sup>	Itatiaia, RJ, Brazil	<a href="#">EU444574</a>	<a href="#">EU446081</a>	<a href="#">EU447321</a>	AY162988
				<i>D. bandeirantorun_1</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444556</a>	<a href="#">EU446061</a>	<a href="#">EU447306</a>	AY162991
				<i>D. bandeirantorun_2</i> <sup>a</sup>	Derrubadas, RS, Brazil	<a href="#">EU444557</a>	<a href="#">EU446062</a>	<a href="#">EU447307</a>	AY162990
				<i>D. mediotriata_1</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444577</a>	<a href="#">EU446085</a>	<a href="#">EU447324</a>	AY847767
			<i>D. mediotriata_2</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444578</a>	<a href="#">EU446084</a>	<a href="#">EU447325</a>	AY847759	
			<i>D. paramediotriata_1</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444589</a>	<a href="#">EU446096</a>	<a href="#">EU447336</a>	AY162996	
			<i>D. paramediotriata_2</i> <sup>b</sup>	Florianópolis, SC, Brazil	<a href="#">EU444590</a>	<a href="#">EU446095</a>	<a href="#">EU447335</a>	AY162995	
			<i>D. crocina</i> <sup>l</sup>	Marko, Brazil	<a href="#">EU444562</a>	<a href="#">EU446067</a>	<a href="#">EU447310</a>	-	
			<i>D. unipunctata</i> <sup>l</sup>	Medellin, Colombia	<a href="#">EU444595</a>	<a href="#">EU446100</a>	<a href="#">EU447343</a>	AF246488	
			<i>D. mediodiffusa</i> <sup>l</sup>	Maricao, Porto Rico	<a href="#">EU444572</a>	<a href="#">EU446078</a>	<a href="#">EU447320</a>	AF246465	
			<i>D. mediopictoides</i> <sup>l</sup>	Boquete, Panama	<a href="#">EU444576</a>	<a href="#">EU446080</a>	<a href="#">EU447323</a>	-	
			<i>cardini</i>	<i>D. mediopicta</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444573</a>	<a href="#">EU446079</a>	-	AY847768
				<i>D. cardinoides</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444559</a>	<a href="#">EU446064</a>	<a href="#">EU447309</a>	AY162975
				<i>D. neocardini</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444581</a>	<a href="#">EU446089</a>	<a href="#">EU447328</a>	AY847770
				<i>D. polymorpha_1</i> <sup>a</sup>	Porto Alegre, RS, Brazil	<a href="#">EU444592</a>	<a href="#">EU446098</a>	<a href="#">EU447339</a>	AY162976
				<i>D. polymorpha_2</i> <sup>b</sup>	Florianópolis, SC, Brazil	<a href="#">EU444591</a>	<a href="#">EU446097</a>	<a href="#">EU447338</a>	-
				<i>D. nigrodunni</i>	Barbados, Caribe	<a href="#">EU444582</a>	<a href="#">EU446101</a>	<a href="#">EU447329</a>	AY173201
				<i>D. arawakana</i> <sup>l</sup>	St Kitts, Caribe	<a href="#">EU444555</a>	<a href="#">EU446060</a>	<a href="#">EU447305</a>	AY173165
				<i>D. cardini</i> <sup>f</sup>	Itaqui, RS, Brazil	<a href="#">EU444558</a>	<a href="#">EU446063</a>	<a href="#">EU447308</a>	AY162974
			<i>calloptera pallidipennis testaceae</i>	<i>D. ornatipennis</i> <sup>i</sup>	Guadalupe Island, Caribe	<a href="#">EU444585</a>	-	<a href="#">EU447331</a>	AF478419
				<i>D. pallidipennis</i> <sup>d</sup>	Joinville, SC, Brazil	<a href="#">EU444586</a>	<a href="#">EU446092</a>	<a href="#">EU447332</a>	AY162981
				<i>D. putrida</i>	-	AF293723	-	-	AF183992

		<i>quinaria</i>	<i>D. phalerata</i> <sup>l</sup>	Beirut, Lebanon	AF293721	AF293745	<u>EU447337</u>	AF147115
			<i>D. immigrans</i> <sub>1</sub>	-	AF293713	AF293738	-	AF478424
		<i>immigrans</i>	<i>D. immigrans</i> <sub>2</sub> <sup>a</sup>	Porto Alegre, RS, Brazil	<u>EU444568</u>	<u>EU446074</u>	<u>EU447317</u>	AY162993
			<i>D. immigrans</i> <sub>3</sub> <sup>d</sup>	Joinville, SC, Brazil	<u>EU444567</u>	<u>EU446073</u>	<u>EU447316</u>	AF519324
		<i>funebis</i>	<i>D. funebris</i> <sup>i</sup>	Mexico City, Mexico	AF293709	AF293734	<u>EU447312</u>	AF478422
			<i>D. gasici</i> <sub>1</sub> <sup>g</sup>	Arica, Chile	EF560567	EF559366	-	-
	<i>virilis-repleta</i>	<i>mesophragmatica</i>	<i>D. gasici</i> <sub>2</sub> <sup>g</sup>	Cochabamba, Bolivia	EF560568	EF559367	-	-
			<i>D. gasici</i> <sub>3</sub> <sup>g</sup>	Bogota, Colombia	EF560569	EF559368	-	-
			<i>D. brncic</i> <sup>β</sup>	Bogotá, Colombia	EF560566	EF559365	-	-
			<i>D. mesophragmatica</i> <sup>α</sup>	La Paz, Bolivia	EF560572	EF559369	-	-
			<i>D. gaucha</i> <sup>l</sup>	Campos do Jordão, SP, Brazil	AF324955	AF324971	-	-
			<i>D. pavan</i> <sup>β</sup>	La Florida, Chile	EF560570	EF559370	-	-
			<i>D. viracoch</i> <sup>β</sup>	Bogotá, Colombia	EF560571	EF559371	-	-
		<i>flavopilosa</i>	<i>D. flavopilosa</i> <sup>h</sup>	Santa Maria, RS, Brazil	<u>EU444564</u>	<u>EU446069</u>	-	-
			<i>D. cordeiro</i> <sup>h</sup>	Santa Maria, RS, Brazil	<u>EU444561</u>	<u>EU446066</u>	-	-
			<i>D. cestri</i> <sup>h</sup>	Santa Maria, RS, Brazil	<u>EU444560</u>	<u>EU446065</u>	-	-
			<i>D. incompta</i> <sup>h</sup>	Santa Maria, RS, Brazil	<u>EU444569</u>	<u>EU446075</u>	-	-
		<i>annulimana</i>	<i>D. annulimana</i> <sup>b</sup>	Florianópolis, SC, Brazil	<u>EU444554</u>	<u>EU446059</u>	-	-
			<i>D. aracataca</i>	-	AF324949	AF324965	-	-
		<i>repleta</i>	<i>D. hydei</i> <sup>β</sup>	Florianópolis, SC, Brazil	AF293712	AF293737	-	-
			<i>D. mulleri</i>	-	AF324958	AF324974	-	-
			<i>D. buzzatii</i>	-	AF324947	AF324980	-	-
			<i>D. repleta</i>	-	AF324961	AF324977	-	-
			<i>D. mercatorum</i> <sup>b</sup>	Florianópolis, SC, Brazil	AF324957	AF324973	-	-
			<i>D. ellisoni</i>	-	AF324953	AF324969	-	-
		<i>dreyfusi</i>	<i>D. camargo</i> <sup>i</sup>	Turrialba, Costa Rica	AF324951	AF324967	-	-
		<i>canalineae</i>	<i>D. canalinea</i>	-	AF324952	AF324968	-	-
		<i>bromeliae</i>	<i>D. bromeliae</i>	-	AF324950	AF324966	-	-
		<i>nannoptera</i>	<i>D. nannoptera</i>	-	AF324959	AF324975	-	-
		<i>virilis</i>	<i>D. virilis</i>	-	AF293729	AF293749	-	-
		<i>robusta</i>	<i>D. robusta</i>	-	AF293724	AF293747	-	-
		<i>melanica</i>	<i>D. melanica</i>	-	AF324956	AF324972	-	-
		<i>polychaeta</i>	<i>D. polychaeta</i>	-	AF324960	AF324976	-	-
		HPW*	<i>D. gymnobasis</i>	-	AF293710	AF293735	-	-
		MMP*	<i>D. mimica</i>	-	AF293716	AF293741	-	-
	<i>Sophophora</i>	<i>melanogaster</i>	<i>D. simulans</i>	-	AF293726	AY197770	-	-
			<i>D. erecta</i>	-	AF293708	-	-	-
		<i>obscura</i>	<i>D. pseudoobscura</i>	-	AF293722	AF293746	-	-
	<i>Dorsilopha</i>		<i>D. busckii</i> <sup>β</sup>	Porto Alegre, RS, Brazil	AF293707	AF293733	-	-
	<i>Scaptomyza</i>		<i>S. palmae</i>	-	AF293718	AF293743	-	-
			<i>S. adusta</i>	-	AF293704	AF293732	-	-
			<i>H. pictiventris</i>	-	AF293711	AF293736	-	-
	<i>Hirtodrosophila</i>		<i>Z. indianus</i> <sup>b</sup>	Porto Alegre, RS, Brazil	<u>EU444597</u>	<u>EU446103</u>	<u>EU447344</u>	AY847774
	<i>Zaprionus</i>		<i>Z. tuberculatus</i> <sup>i</sup>	-	AF293731	AF293751	<u>EU447345</u>	AF478440
			<i>L. aerea</i>	-	AF293715	AF293740	-	-
	<i>Liodrosophila</i>		<i>S. leonensis</i>	-	AF293725	AF293748	-	-
	<i>Samoaia</i>							



<i>Scaptodrosophila</i>			<i>S. latifasciaeformis</i> <sup>d</sup> <i>S. lebanonensis</i> <sup>j</sup>	Canal Zone, Panama Veyo, Utah	<u>EU444596</u> AF293714	<u>EU446102</u> AF293739	- -	- -
-------------------------	--	--	---	----------------------------------	-----------------------------	-----------------------------	--------	--------

(1) The classification presented by Bächli (2008) was adopted in this paper, except for the subdivision of the former *guarani* group into the *guarani* and *guaramunu* groups, for the desynonymization between *D. mediostriata* and *D. crocina*, and for the inclusion of the Hawaiian HPW and MMP drosophilids groups within the genus *Drosophila*. (2) When more than one population from a same species was sampled, different terminal numbers were added to the species name. (3) Population origin refers only to samples which were the source of new sequences. (4) Newly obtained sequences are underlined. (5) For the *COII* data set, the Acc. No. referred to *D. ornatipennis* is from *D. calloptera*, which was utilized as a placeholder.

\* Hawaiian picture-winged (HPW) and modified mouth parts (MMP) are groups of Hawaiian drosophilids.

Capital letters refer to the fly collectors and/or suppliers: <sup>a</sup>Dr. Luciano Basso da Silva, <sup>b</sup>M.Sc. Marco Silva Gottschalk, <sup>c</sup>Dra. Daniela Cristina De Toni, <sup>d</sup>MSc. Jonas da Silva Döge, <sup>e</sup>Dr. Elgion Lúcio Loreto, <sup>f</sup>M.Sc. Marícia Fantinel D'Avila e M.Sc. Marco Silva Gottschalk, <sup>g</sup>Dra. Myriam Budnik, <sup>h</sup>M.Sc. Adriana Ludwig, <sup>i</sup>Dr. Jean David, <sup>j</sup>Species obtained at Tucson *Drosophila* Stock Center.

Table 2

Descriptive characteristics of the data partitions used in this study

Partition	Absolute numbers <sup>1</sup>				Ts/Tv <sup>2</sup>	Base Frequencies <sup>2</sup>				$\chi^2$ Test
	OTU's	Sites	V Sites	PI Sites		A	C	G	T	
<i>Amd</i>	85	910	540	465	3.747	0.226	0.301	0.234	0.239	455.066 (df=252)**
<i>Ddc</i>	84	1,105	515	454	4.036	0.187	0.356	0.226	0.231	438.105 (df=249)**
<i>Amd+Ddc</i>	81	2,015	1,050	912	3.953	0.202	0.332	0.235	0.231	762.116 (df=240)**
<i>Hb</i>	42	507	209	191	4.237	0.227	0.367	0.198	0.208	160.339 (df=123)*
All data	39	3,194	1,344	1,086	3.720	0.205	0.291	0.259	0.245	706.870 (df=114)**

1- As obtained empirically by Mega 3.1 software.

2- As obtained by BaseML estimates under the HKY85 model (Hasegawa et al., 1985).

V = Variable; PI = Parsimony Informative; Ts/Tv = Transition/Transversion ratio.

\* Significant at the 0.05 level; \*\* Significant at the 0.01 level.

Table 3

Optimal substitution models and parameters for the different data sets, as indicated by the AIC test implemented by ModelTest 3.7

Partition	Model	Base frequencies	Rate Matrix	I	$\alpha$
<i>Amd</i>	GTR + I + G	A = 0.1911 C = 0.2906 G = 0.2872 T = 0.2311	A-C = 2.1591 A-G = 4.8375 A-T = 2.1889 C-G = 0.8887 C-T = 6.0419 G-T = 1.0000	0.3761	1.0021
<i>Ddc</i>	GTR + I + G	A = 0.1557 C = 0.3425 G = 0.2882 T = 0.2137	A-C = 3.0222 A-G = 5.9398 A-T = 2.6954 C-G = 1.0808 C-T = 8.0790 G-T = 1.0000	0.5043	0.9308
<i>Amd + Ddc</i>	GTR + I + G	A = 0.1658 C = 0.3213 G = 0.2892 T = 0.2238	A-C = 2.6366 A-G = 5.6765 A-T = 2.5885 C-G = 0.9353 C-T = 6.9707 G-T = 1.0000	0.4469	0.9445
<i>Hb</i>	TVM + I + G	A = 0.1988 C = 0.3614 G = 0.2321 T = 0.2078	A-C = 2.2718 A-G = 8.2577 A-T = 3.1541 C-G = 1.9343 C-T = 8.2577 G-T = 1.0000	0.5571	2.4703
All data	GTR + I + G	A = 0.2288 C = 0.2566 G = 0.2470 T = 0.2676	A-C = 2.3564 A-G = 5.4039 A-T = 3.4535 C-G = 2.2712 C-T = 9.9264 G-T = 1.0000	0.4592	0.9049

I = Proportion of invariable sites;  $\alpha$  = gamma distribution shape parameter.

GTR = General Time Reversible (Rodriguez et al., 1990); TVM = Transversal Model.

Table 4

Average Partitioned Bremer Support (PBS) scores for each gene partition across the subgenus *Drosophila* TET (Fig. 4)

Clade	Partitions	
	<i>Amd</i>	<i>Ddc</i>
1	5.4	5.6
2	4.2	0.8
3	1.2	3.8
4	2.7	-0.7
5	1.8	-0.8
6	9.2	9.8
7	14.0	-5.0
8	7.2	10.8
9	25.2	16.8
10	-0.8	3.8
11	28.2	30.8
12	0.2	-0.2
13	10.7	22.3
14	-1.3	1.3
15	0.2	-0.2
16	12.7	2.3
17	3.7	5.3
18	25.7	12.3
19	-0.8	4.8
20	6.0	5.0
21	46.2	28.8
22	0.2	-0.2
23	-3.5	7.5
24	1.2	12.8
25	38.7	12.3
26	0.2	-0.2
27	6.2	0.8
28	56.4	28.6
29	0.2	-0.2
30	29.4	16.6
31	8.7	-3.7
32	0.2	-0.2
33	8.5	-3.4
34	0.0	0.0
35	19.2	-9.2
36	47.2	28.8
37	10.5	-5.5
38	29.2	12.8
39	0.2	-0.2
40	0.0	0.0
41	7.2	11.8
42	12.2	10.8
43	13.7	10.3
44	9.9	5.1
45	19.2	10.8
46	10.2	0.8
47	16.2	-8.2
48	0.0	0.0
49	5.7	0.3
50	14.2	-7.2
51	7.2	-2.2
52	10.2	-4.2
53	2.2	1.8
54	21.2	16.8
55	4.2	1.8
56	2.2	-1.2
57	6.2	2.8
58	31.2	6.8
59	32.7	30.3

60	0.2	5.8
61	0.2	-0.2
62	0.2	-0.2
63	9.2	4.8
64	0.2	-0.2
65	0.2	-0.2
66	19.2	17.8
67	35.0	4.0
68	11.2	5.8
69	0.0	0.0
70	0.2	-0.2
71	4.2	0.8
72	0.2	-0.2
73	10.2	-6.2
74	29.2	-3.2
75	10.2	-6.2
$\Sigma$ PBS	809.9	364.2
Min steps	1,034	925
* $\Sigma$ PBS/ Min steps	0.78	0.39

\* PBS values summed across the tree and standardized by the minimum number of steps for each partition.

Table 5

Estimated divergence times of the subgenus *Drosophila* TET internal nodes, as presented in Figure 4

Node	Mean time	Standard deviation	Lower limit	Upper limit
1	2.09	0.67	1.04	3.63
2	4.15	1.15	2.29	6.76
3	5.66	1.45	3.24	8.92
4	0.27	0.19	0.014	0.74
5	0.65	0.32	0.17	1.44
6	1.25	0.49	0.51	2.42
7	6.26	1.56	3.65	9.75
8	0.67	0.42	0.07	1.71
9	4.64	1.39	2.45	7.86
10	14.14	2.78	9.37	20.13
11	1.59	0.72	0.51	3.36
12	18.48	3.30	12.71	25.56
13	1.57	5.26	7.48	2.78
14	6.93	1.61	4.26	10.53
15	8.03	1.79	5.04	11.99
16	4.08	1.14	2.25	6.67
17	8.82	1.91	5.61	12.97
18	10.75	2.24	6.93	15.56
19	2.32	0.70	1.22	3.91
20	1.18	0.41	0.55	2.14
21	3.04	0.85	1.70	4.99
22	18.32	3.31	12.66	25.54
23	19.74	3.46	13.71	27.14
24	2.98	0.94	1.52	5.22
25	4.98	1.36	2.83	8.17
26	15.70	2.98	10.58	22.15
27	0.73	0.36	0.17	1.60
28	1.62	0.55	0.77	2.95
29	16.94	3.10	11.55	23.50
30	1.44	0.71	0.35	3.14
31	15.54	2.96	10.41	21.86
32	18.44	3.29	12.67	25.34
33	22.13	3.73	15.60	29.87
34	29.07	4.48	20.97	38.32
35	31.08	4.62	22.64	40.42
36	0.97	0.74	0.065	2.82
37	41.74	5.55	31.52	52.64
38	20.30	3.82	13.52	28.41
39	37.00	5.20	27.43	47.33
40	43.59	5.65	33.18	54.69
41	0.18	0.14	0.007	0.52
42	2.60	0.80	1.34	4.46
43	5.32	1.32	3.18	8.23
44	1.12	0.50	0.30	2.30
45	7.09	1.60	4.42	10.64
46	5.01	1.17	3.08	7.65
47	6.62	1.45	4.16	9.81
48	6.64	1.47	4.17	9.95
49	8.15	1.70	5.27	11.89
50	10.37	2.05	6.84	14.83
51	14.05	2.53	9.61	19.43
52	16.23	2.85	11.20	22.25
53	17.91	3.03	12.51	24.22
54	11.79	2.38	7.67	16.92
55	19.66	3.27	13.81	26.56
56	6.81	1.68	4.04	10.54
57	8.04	1.84	4.95	12.00
58	11.45	2.33	7.46	16.48
59	4.75	1.33	2.63	7.81
60	20.99	3.35	14.90	27.86
61	24.82	3.71	17.96	32.29
62	27.67	3.96	20.33	35.58

63	21.98	3.56	15.55	29.24
64	31.82	4.31	23.82	40.14
65	34.61	4.66	26.01	43.68
66	15.73	2.93	10.48	21.77
67	17.93	3.20	12.23	24.43
68	30.75	3.72	23.72	36.68
69	39.52	4.98	30.15	48.92
70	42.17	5.40	32.14	52.53
71	44.78	5.76	34.11	56.07
72	48.99	6.48	37.12	61.92
73	52.85	7.01	40.07	67.02
74	39.98	6.53	28.43	53.49
75	55.08	7.33	41.77	69.98

Table 6

p values resulting from PHT between *Amd*, *Ddc*, *Hb* and *COII* data partitions as concerns the *immigrans-tripunctata* radiation species

	<i>Amd</i>	<i>Ddc</i>	<i>Hb</i>	<i>COII</i>
<i>Amd</i>	-	0.057	0.021*	0.001**
<i>Ddc</i>		-	0.003**	0.001**
<i>Hb</i>			-	0.001**
<i>COII</i>				-

\* Significant at the 0.05 level; \*\* Significant at the 0.01 level.



Table 7

Average Partitioned Bremer Support (PBS) scores for the *Hb* and *COII* gene partitions across the *immigrans-tripunctata* radiation TET (Fig. 5)

Clade	Partition	
	<i>Hb</i>	<i>COII</i>
1	6.0	4.0
2	7.0	10.0
3	0.0	0.0
4	2.0	0.0
5	2.0	0.0
6	10.0	-3.0
7	5.0	-1.0
8	8.5	1.5
9	9.5	5.0
10	0.0	0.0
11	4.0	1.0
12	0.0	0.0
13	-2.0	9.0
14	-2.0	9.0
15	10.0	7.0
16	2.0	1.0
17	17.0	0.0
18	11.0	11.0
19	14.0	17.0
20	40.0	4.0
21	0.0	0.0
22	0.0	0.0
23	17.7	21.7
24	0.0	0.0
25	2.5	25.0
26	34.0	4.0
27	0.0	0.0
28	12.5	8.5
29	21.0	1.5
30	0.0	0.0
31	0.0	0.0
32	3.5	3.5
33	4.0	-0.5
34	4.5	2.0
35	29.0	31.0
36	14.0	6.0
$\Sigma$ PBS	286.7	178.2
Min steps	303	380
* $\Sigma$ PBS/ Min steps	0.95	0.47

\* PBS values summed across the tree and standardized by the minimum number of steps for each partition.

Figure 1

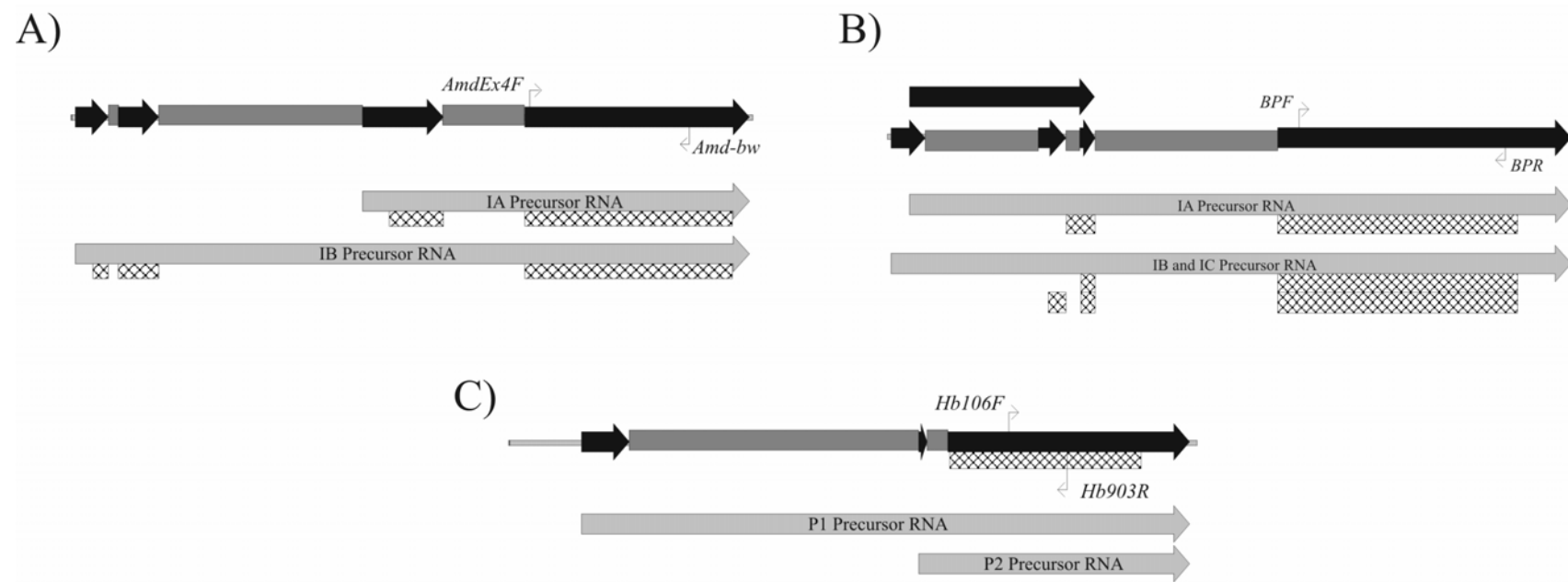


Figure 2

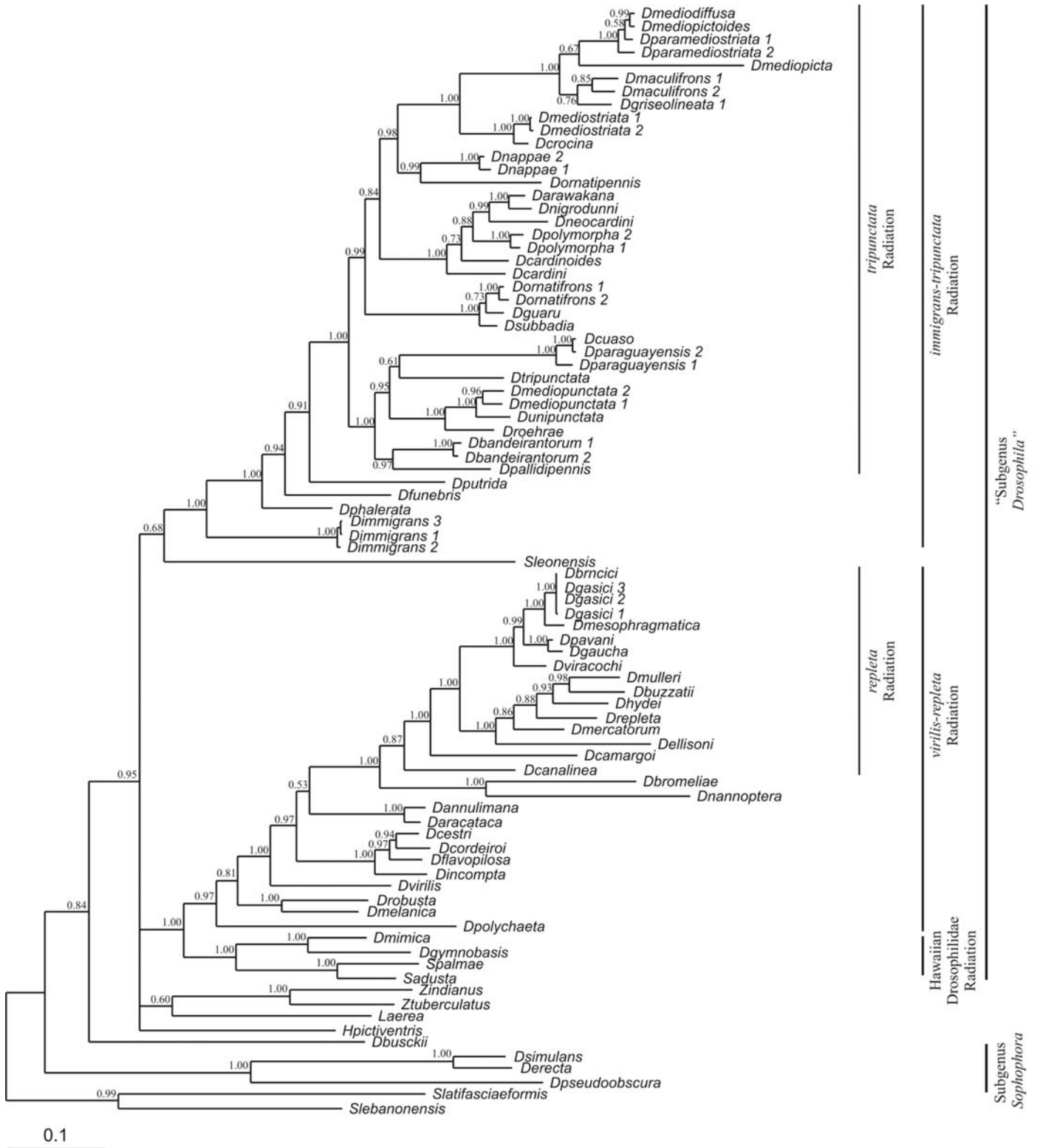


Figure 3

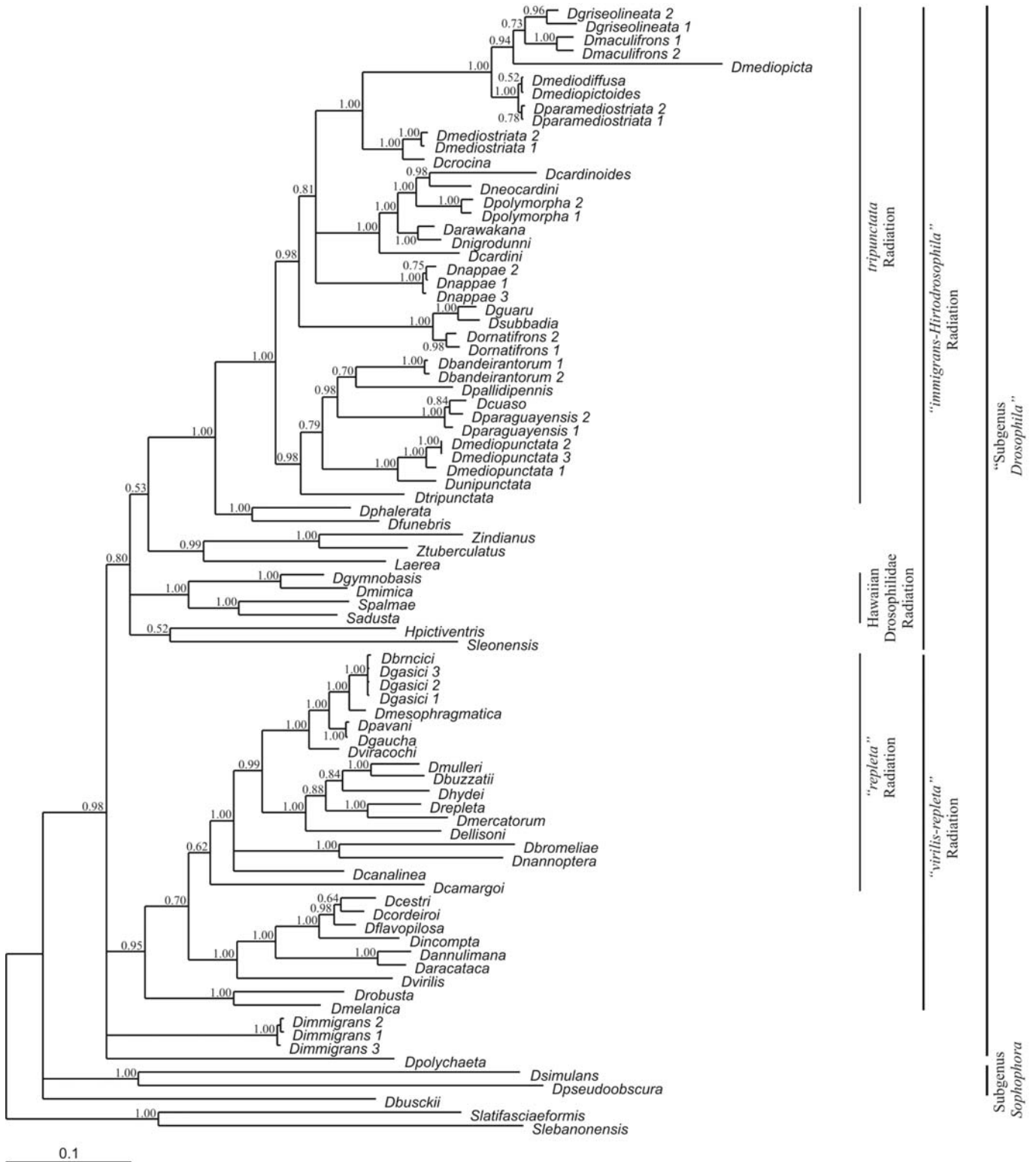


Figure 4

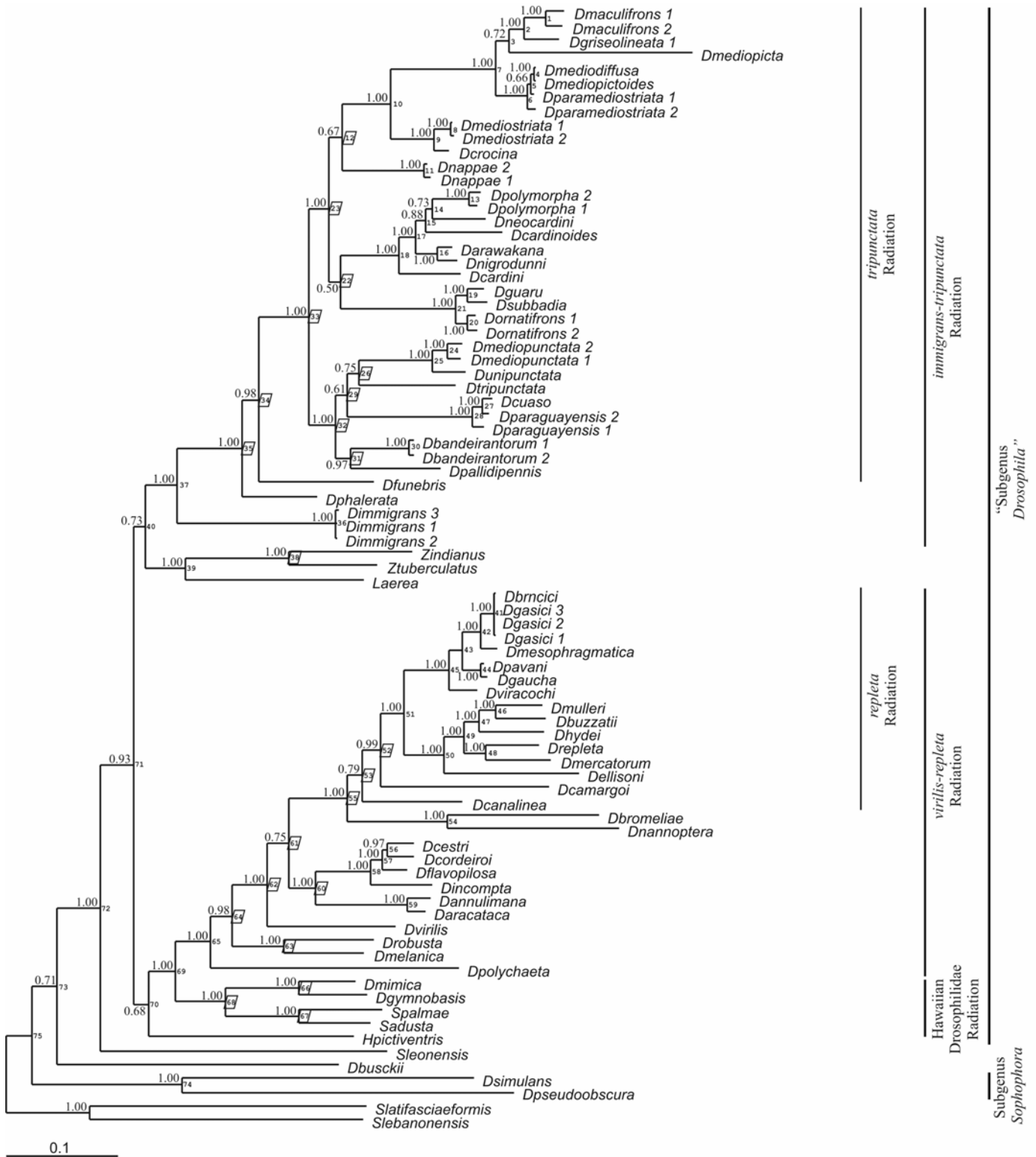
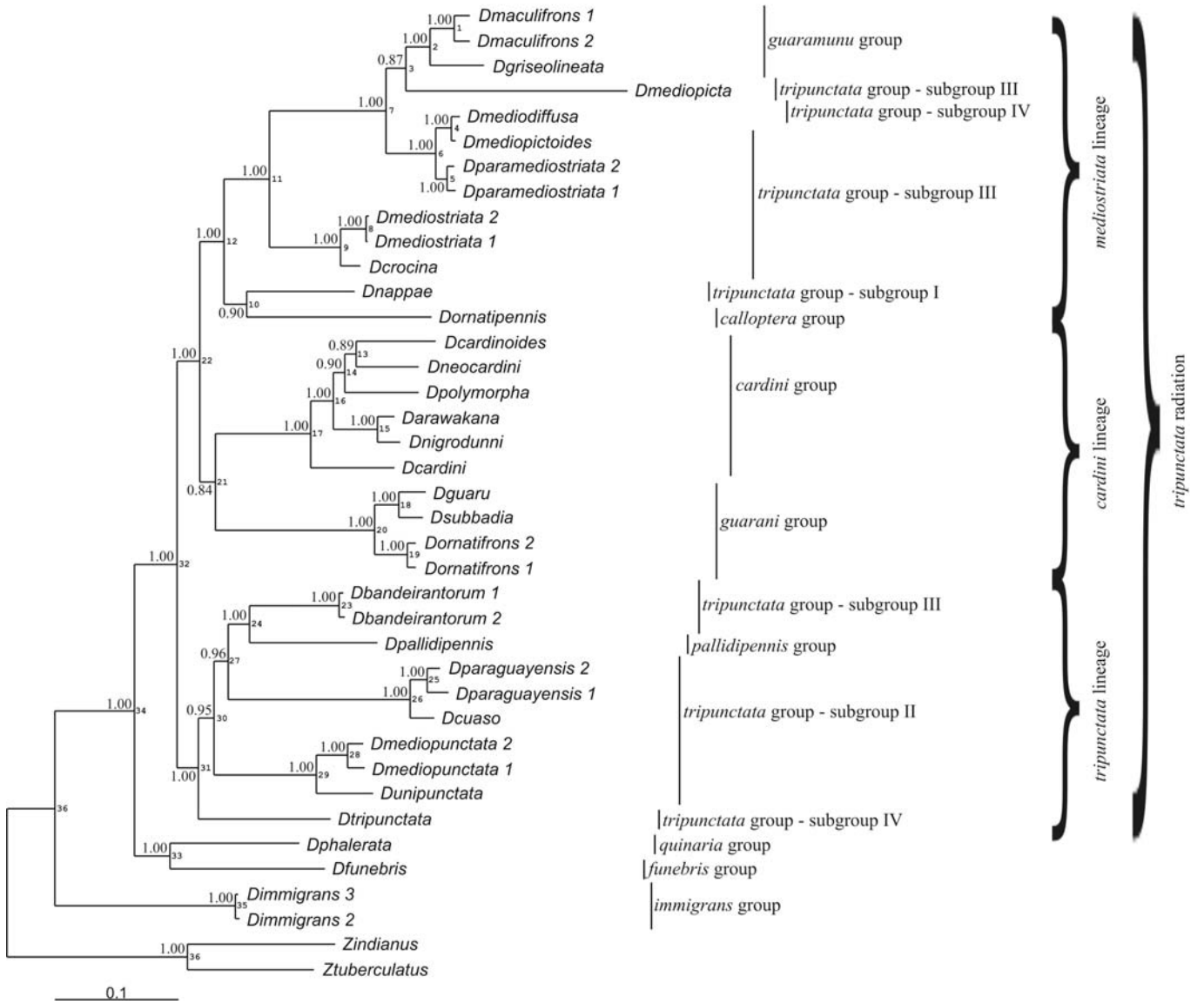
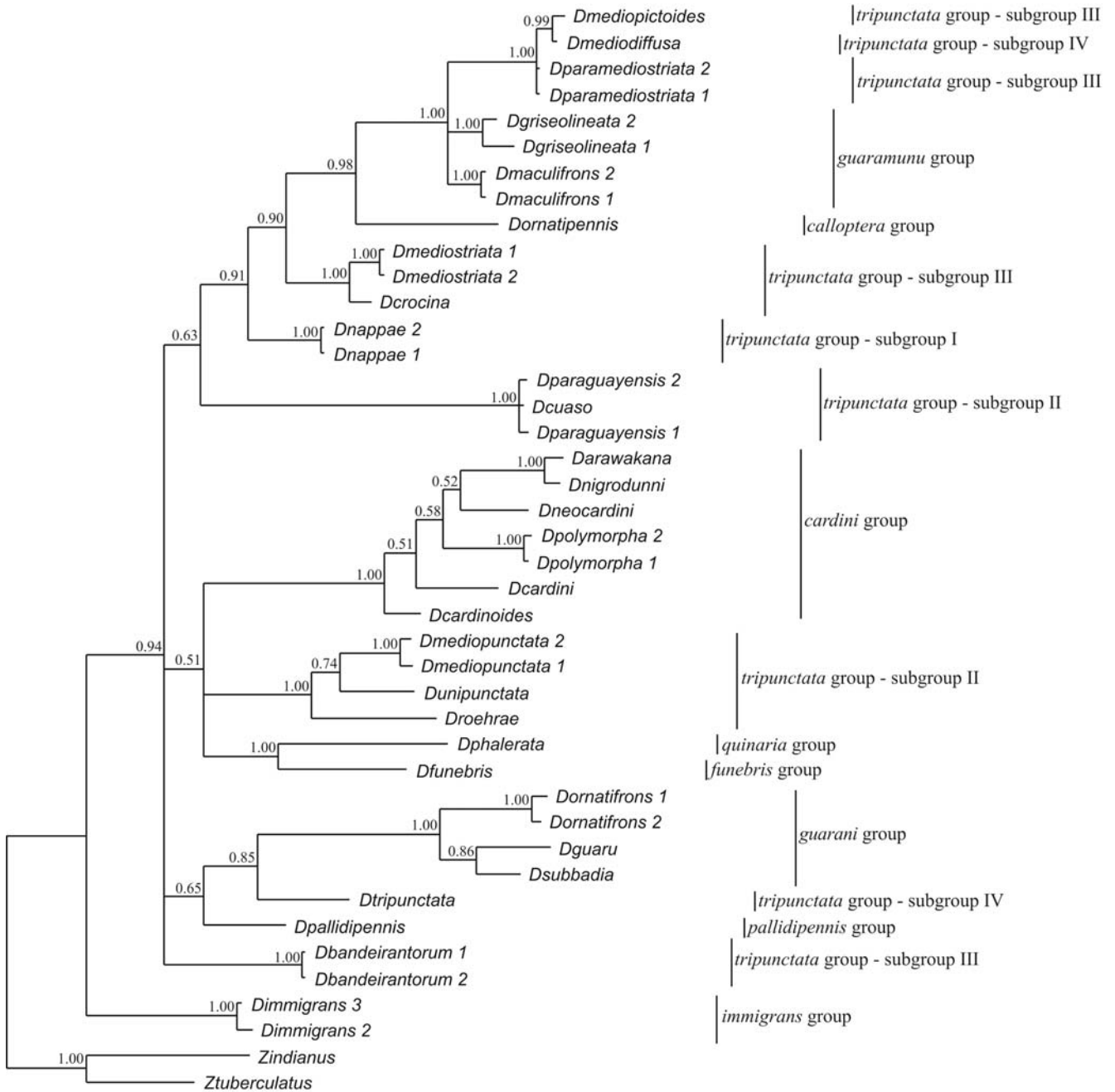


Figure 5



Supplementary Material – Figure 1S



## CAPÍTULO IV

---

Trabalho aceito para publicação na revista Zoological Science

### **Phylogeny of the *Drosophila mesophragmatica* group (Diptera, Drosophilidae): an example of Andean evolution<sup>1</sup>**

Nina R. Mota\*<sup>1,2</sup>, Lizandra J. Robe\*<sup>1</sup>, Vera L. S. Valente<sup>1</sup>, Myriam Budnik<sup>3</sup>, Elgion L. S. Loreto<sup>1,2</sup>

\* These authors contributed equally for this study.

1 – Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul (UFRGS), CP 15053, Porto Alegre, Rio Grande do Sul, Brazil, 91501-970

2 – Departamento de Biologia, Universidade Federal de Santa Maria (UFSM), Camobi Km 09, CP 5050, Santa Maria, Rio Grande do Sul, Brazil, 97105-900

3 – Universidad de Chile, Facultad de Medicina, Santiago, Chile

Corresponding author: E. L. S. Loreto, E-mail: [elgion@base.ufsm.br](mailto:elgion@base.ufsm.br), phone: 55-55-3220-8912, fax: 55-55-3220-8020.

---

<sup>1</sup> Este trabalho fez parte do Trabalho de Conclusão de Curso da aluna Nina Roth Mota, co-orientada pela doutoranda Lizandra Jaqueline Robe junto ao Curso de Bacharelado em Ciências Biológicas da UFSM.



## ABSTRACT

The *mesophragmatica* group of *Drosophila* belongs to the *virilis-repleta* radiation of the *Drosophila* subgenus. This group comprises thirteen Neotropical species endemic to south-American continent, which seem to be fundamentally Andean in their distribution. The *mesophragmatica* group phylogeny has been inferred previously by other authors based on morphological, cytological and isozyme analyses. However, the relationships within the group have not yet been completely resolved, although its monophyletic origin has already been confirmed by molecular data. This work attempts to enhance the molecular approach of the relationships among the species of the *mesophragmatica* group, using both nuclear and mitochondrial markers. The phylogenetic analyses were performed using fragments of the nuclear *alcohol dehydrogenase* – *Adh* (631 bps), *alpha-methyl dopa* – *Amd* (1,211 bps), *dopa-decarboxylase* – *Ddc* (1,105 bps) and *hunchback* – *Hb* (687 bps) genes and the mitochondrial *cytochrome oxidase subunit II* – *COII* (672 bps) gene, performing a total of 4,306 bps. The sequences obtained for eight representatives of the *mesophragmatica* group were analyzed both individually and in combination, based on distance methods, maximum parsimony and maximum likelihood. Our results support the subdivision of the *mesophragmatica* group into three main lineages: the first composed by *D. viracochi*; the second presenting a clade grouping the sibling *D. pavani* and *D. gaucha*; and the third one encompassing *D. gasici*, *D. brncici* and *D. mesophragmatica*. The recovered scenario also suggests that *D. viracochi* is the *mesophragmatica* group early offshoot, with this and other early branchings occurring in the Pliocene/Pleistocene epochs, possible associated with glacial Andean refuges. Also based on the obtained phylogenies, we present a genealogical view of the evolution of previously described characters within the group.

**Key words:** *Drosophila*, *mesophragmatica* group, molecular phylogeny, Andean System of Mountains, *Adh*, *Amd*, *COII*, *Ddc*, *hunchback*.

## INTRODUCTION

The *mesophragmatica* group, described by Brncic and Koref-Santibañez (1957), belongs to the *virilis-repleta* radiation of the *Drosophila* subgenus (Robe et al., 2005; Throckmorton, 1975). According to Vela and Rafael (2004), the group includes thirteen Neotropical species endemic to the south-American continent: *D. altiplanica* Brncic and Koref-Santibañez, 1957; *D. brncici* Hunter and Hunter, 1964; *D. camaronensis* Brncic, 1957; *D. canescens* Duda, 1927; *D. gasici* Brncic, 1957; *D. gaucha* Jaeger and Salzano, 1953; *D. mesophragmatica* Duda, 1927; *D. orkui* Brncic and Koref-Santibañez, 1957; *D. pavani* Brncic, 1927; *D. viracochi* Brncic and Koref-Santibañez, 1957; *D. amaguana* Vela and Rafael, 2004; *D. ruminahuii* Vela and Rafael, 2004; and *D. shyri* Vela and Rafael, 2004. All members of the *mesophragmatica* group can be distinguished by their external characteristics, with the exception of *D. gaucha* and *D. pavani*, which are sibling species (Brncic and Koref-Santibañez, 1957).

With the exception of *D. gaucha* which can be found even in the south of Brazil, all the species of this group are fundamentally Andean in their geographic distribution (Brncic et al., 1971). The configuration of this System of Mountains facilitates the occurrence of rapid speciation, since it leads to the formation of local isolated populations. According to Brncic (1957), the *mesophragmatica* group can be considered as derived from such evolutionary scenario, since its species are in general allopatric. In fact, it is probable that the actual sympatric regions are the product of recent secondary contact (Nair et al., 1971).

Brncic and Koref-Santibañez (1957) presented a key based on external morphological characteristics which permitted to identify three subgroups within the six species then described for the *mesophragmatica* group: the first presenting species with convergent anterior scutellars bristles (*D. viracochi*); the second with species dark brown which have divergent anterior scutellars (*D. mesophragmatica*, *D. orkui*, *D. altiplanica*); and the third with species light brown which have divergent basal scutellars (*D. pavani* and *D. gaucha*). Nacrur (1958), based on the direction of the basal scutellars described by Brncic and Koref-Santibañez (1957), formally proposed to divide the six *D. mesophragmatica* group species into two taxonomic lineages. This proposal was recently recognized and updated by Vela and Rafael (2004), which also renamed those two subgroups: the *viracochi* subgroup, comprising *D. viracochi* and *D. ruminahuii*; and the *mesophragmatica* subgroup, containing the other eleven species of the group.

Phylogenetic relationships of the *mesophragmatica* group species have been inferred by other authors based on cytological (Brncic et al., 1971) and isozyme (Nair et al., 1971) analyses. These studies subdivided the *mesophragmatica* group in three lineages: the first one formed by *D. mesophragmatica*, *D. brncici* and *D. gasici*; the second by *D. viracochi*; and the third by *D. gaucha* and *D. pavani*. Although both of these studies were in general agreement with the previously mentioned morphological data, the number of subgroups and their external and internal relationships remained until now unsolved. Besides that, the *mesophragmatica* group phylogeny has not yet been widely studied based on molecular data. The group monophyly and its origin from the *virilis-repleta* radiation were confirmed by Carrasco et al. (2003) and by Robe et al. (2005), but these works presented contrasting results concerning the species relationships.

Due to this controversy, the aim of the present study is to shed more light on the evolutionary patterns involved on the *mesophragmatica* group speciation through a molecular framework. Five independent loci were chosen to approach this issue: four nuclear (*Adh* – alcohol dehydrogenase, *Amd* –  $\alpha$ -methyldopa, *Ddc* – dopa-decarboxylase and *Hb* - hunchback) and one mitochondrial gene (*COII* – cytochrome oxidase subunit II). These partitions were sequenced for the six species of the group for which stocks are now available: *D. brncici*, *D. gasici*, *D. mesophragmatica*, *D. gaucha*, *D. pavani* and *D. viracochi*. Phylogenetic analyses were performed for each gene individually and in combination, in order to finally obtain a total evidence tree, which further refine the relationships within the group.

## **MATERIALS AND METHODS**

Table 1 lists all the species used in the present study, as well as their collection site and the Genbank accession numbers at National Center for Biotechnology Information (NCBI) for each of the studied markers. We chose *D. hydei* as the outgroup, because there is ample evidence that the *repleta* group is closely related to the *mesophragmatica* group (Durando et al., 2000; Tatarenkov and Ayala, 2001; Remsen and O'Grady, 2002; Robe et al., 2005). Tatarenkov and Ayala (2001) and Robe et al. (2005) even presented some evidences that the *repleta* group is the sister clade of the *mesophragmatica* group.

The total DNA of each species or population established isolineage was extracted according to the protocol described by Sassi et al. (2005). The five target loci were amplified from each taxon under standard PCR conditions. PCR primers for the *Amd*, *Adh*, *Ddc* and *COII* genes were synthesized based on previous studies, while de *Hb* primers were

newly designed for the present study: 5' TCG CAY GGM AAG ATG AAG AAC TAC AA 3' (Hb106F – forward primer) and 5' TCC AGT ASG GGA ARA AGG CCA TG 3' (Hb903R – reverse primer). All these set of primers have their properties referred to in Table 2.

The obtained PCR products were checked by horizontal electrophoresis in 0.8% agarose gel stained with ethidium bromide, and once a unique band of the expected size was obtained, it was purified with a solution of PEG 13% and NaCl 1.6M. Otherwise, when the eletrophoregram presented unspecific products the expected amplicon was purified with QIAquick Gel Extraction Kit (Qiagen, São Paulo, Brazil) according to the manufacturer's protocol. DNA sequencing was performed directly using the purified amplicons in a MegaBACE 500 automatic sequencer. The dideoxy chain-termination reaction was implemented using the DYEnamic ET® kit (Amersham, São Paulo, Brazil). Both strands were sequenced through the use of the same primers implemented on the PCR reaction.

The nucleotide sequences had their identity initially confirmed using the BlastN program available at the NCBI site. Inspection, edition and assembly of automated sequences were subsequently accomplished using the Staden Package Gap 4 program (Staden, 1996), where from a confidence consensus sequence was obtained for each taxon. These sequences were aligned using the ClustalX 1.81 program (Jeanmougin et al., 1998) according to the system default parameters.

The MEGA 3.1 (Kumar et al., 2004) was used to calculate distances, to construct evolutionary trees with the neighbor joining algorithm (NJ) (Saitou and Nei, 1987), and for calculating several descriptive parameters. The individual and combined phylogenetic trees obtained were also compared to those obtained through Maximum Parsimony (MP) and

Maximum Likelihood (ML) by PAUP 4.0b10 (Swofford, 2003). The MP analysis was performed through the equal weight criterion, while the ML analysis followed the model selected by the Akaike information criterion (AIC) test (Akaike, 1974) executed using the ModelTest 3.7 program (Posada and Crandall, 1998). In all cases the confidence values of each clade were measured by the bootstrap test (Felsenstein, 1985) with 1,000 replications. For MP the bootstrap search was implemented through branch-and-bound algorithm, and for ML the confidence values were obtained through a heuristic search, with tree-bisection-reconnection (TBR) swaps executed on initial NJ trees. In relation to gaps, on the NJ analysis a pairwise deletion decision was taken, while on the MP analysis all sites were used, with gaps being treated as a new state.

## RESULTS

The five partial gene sequences were analyzed individually and in combination for all the nine taxa. The properties of each of these datasets are presented in Table 3. All individual partitions presented a level of transition/transversion bias, and all, except *Adh*, showed a significant nucleotide frequency bias. Concerning the ModelTest selection, only for *Ddc* and *Hb* the null hypothesis of a single transversion rate was accepted, leading to the adoption of the Tamura-Nei model (TrN). For all the other datasets more complex evolutionary models were selected. The heterogeneity of substitution rates among sites was corrected for through the adoption of a gamma distribution (G) - in the cases of *Adh* and the combined dataset - or a proportion of invariable sites (I) - for *Amd*, *Ddc*, *Hb* and *COII*.

The NJ inferred phylogenies for each of these six datasets are illustrated in Fig. 1, where it can also be seen the *bootstrap* values presented for each of the reconstructed nodes

under other phylogenetic methods. In general, the *mesophragmatica* group could be subdivided into three main clades: one of them composed by *D. viracochi*; another one presenting a clade grouping the sibling *D. pavani* and *D. gaucha*; and the third one mixing the representatives of *D. gasici*, *D. brncici* and *D. mesophragmatica*. This scenario, however, was not recovered by the *Hb* dataset analysis, where *D. mesophragmatica* appears displaced from the clade grouping *D. gasici* and *D. brncici*, acquiring a different position according to the adopted phylogenetic method.

According to *Amd*, *Ddc* and the combined analysis, through NJ, MP and ML, the first *mesophragmatica* group species to branch off is *D. viracochi*. This early offshoot of *D. viracochi* is also supported by the *Adh* and *Hb* NJ and MP trees, and by the *COII* ML topology. In all these cases, with the exception of the *Hb* NJ tree (Fig. 1D), the next clade to branch off is the one composed by *D. gaucha* and *D. pavani*. The derived clade, on the other hand, intermingles representatives of *D. gasici*, *D. brncici* and *D. mesophragmatica*, although the most frequently recovered scenario places this last species as this clade's early offshoot (relationship presented by *Adh*, *Amd*, *Ddc* and combined analysis for all three methods and by *Hb* for MP).

## DISCUSSION

### *D. mesophragmatica* group taxonomy

In general, our data concerning the *D. mesophragmatica* group internal subdivision support the results presented by Brncic and Koref-Santibañez (1957), Brncic et al. (1971) and Nair et al. (1971), that is, that these six species could be divided into three subgroups. We propose here to rename these subgroups, based on the first species described for each of

them, in this way: subgroup *viracochi*, composed by *D. viracochi*; subgroup *gaucha*, comprising the siblings *D. gaucha* and *D. pavani*; and subgroup *mesophragmatica*, containing at least *D. mesophragmatica*, *D. brncici* and *D. gasici*.

This proposal is not in complete disagreement with that of Nacur (1958) and Vela and Rafael (2004), which recognized only two subgroups within the *D. mesophragmatica* group, one presenting *D. viracochi* and the other grouping all the remaining species here studied. But given the deep divergence presented by the three main clades recovered by our phylogenies (Figs. 1A, 1B, 1C, 1E and 1F), we prefer to adopt a subdivision which reflect these branch lengths results.

#### *D. mesophragmatica* group evolution

Although some disagreements can be detected between our individual analysis phylogenies, specially in relation to the early *D. mesophragmatica* group offshoot, our total evidence tree (TET), as also our individual analysis of the four nuclear genes, point out that *D. viracochi* is the first species to branch off. This result is in perfect agreement with the data presented by Nacur (1958), but contrasts with the phylogeny depicted by Nair et al. (1971), which presents the clade grouping *D. gaucha* and *D. pavani* in the *D. mesophragmatica* group basal position. Brncic et al. (1971) made an attempt to establish the cytotaxonomic relationships within the *D. mesophragmatica* species group of *Drosophila*, but they did not try to polarize the topology obtained.

Another point that is a little confusing in our data concerns the relationships between the three *D. gasici* populations and *D. brncici*. In all the obtained phylogenies, these two species appear intermingled, and in none case did the *D. gasici* representatives



constitute a monophyletic group. According to Brncic et al. (1971) these two species do not cross with any other species of the group, that is, they are entirely isolated entities and so, good biological species. Nevertheless, these same authors obtained some results similar to ours, pointing out that from the cytological perspective, the Bolivian population of *D. gasici* appears to be closer to *D. brncici* than to the other populations of its same species. From the molecular point of view, on the other hand, the Chilean and Colombian populations of *D. gasici* appear to be closer to *D. brncici*, than to the Bolivian population. If the cytological and molecular data agreed in the established relationships, the scenario of a recent peripatric speciation event of *D. brncici* could be inferred with accuracy. As this is not the case, it is more probable that these findings reflect a state of high ancestral polymorphism, followed by insufficient lineage sorting, given by a recent and rapid speciation event. More than this, as neither the mitochondrial gene did recover a monophyletic *D. gasici*, we can conclude that this speciation phenomenon is almost incipient, having occurred more recently than the ultimate coalescence of the different mitochondrial alleles currently present in these taxa. So, in general, relationships among the members of the *mesophragmatica* subgroup are confounded by shared ancestral polymorphism, as observed for sequences of mtDNA and nuclear DNA.

For the *COII* gene, even *D. mesophragmatica* appears to be affected by ancestral polymorphism shared with the other members of its subgroup (Fig. 1E). Nevertheless, as mtDNA sequences are known to have faster rates of coalescence than nuclear genes (Moore, 1995) and faster accumulation of phylogenetic information (Caletka and McAllister, 2004), it is not expected, a priori, that only mitochondrial genes reveal such a pattern, specially when other four different nuclear genes are simultaneously considered. In

this case however, as the confidence value of its cluster is very low its significance can be taken as negligible, although limited uniparental introgression could give a compatible explanation.

The geographical distribution of the analyzed species (Fig. 2), summed by the topology and branch lengths of our trees (Fig. 1) can led us to infer two most parsimonious evolutionary scenarios involved on the *D. mesophragmatica* group diversification: the first one is that of a north-central Andean ancestral distribution, followed by speciation and secondary dispersion, given that most of the analyzed species were collected at least once in the area encompassing Bolivia, Colombia, Ecuador and/or Peru, and given that this is exactly the geographical area occupied by *D. viracochi*. In this scenario, the *pavani-gaucha* clade ancestral would have dispersed southwards, while the *mesophragmatica-brncici-gasici* clade ancestral would have dispersed northwards. An alternative, but also possible scenario, would be that of a high dispersed *mesophragmatica* group ancestral, followed by more localized speciation and more recent secondary dispersion events. Also in this scenario, the *viracochi* subgroup ancestral would have occupied a north-central Andean distribution, the *pavani-gaucha* and the *mesophragmatica-brncici-gasici* ancestrals a southwards and northwards distribution area, respectively.

In any case, it is possible to infer the action of the Andean System of Mountains on the diversification of the group. After confirming by the hLRT (hierarchical Likelihood Ratio Test) that the molecular clock applies to the set of *Adh* sequences ( $-2 \log \Delta = 7.05$ , 7 df,  $p > 0.1$ ), the evolutionary rate established by Russo *et al.* (1995) for the third codon positions of this gene as  $1.0 \times 10^{-8}$  per site per year per lineage was used to estimate the divergences datings. Hence, we could infer that the group originated at approximately 3.7

million of years ago, while the divergence time between *D. mesophragmatica* and the clade *gasici-brncici*, as well as that between *D. gaucha* and *D. pavani*, is estimated at 1.27 mya. This situates the *D. mesophragmatica* group evolution around the Pliocene epoch of the Tertiary Period and the boundaries of the Pleistocene epoch of the Quaternary Period, which are characterized by great climatic oscillations, including glaciation events. According to Salgado-Labouriau (1994) the Andes as a whole were affected by glacial expansions, in a way that it is possible to infer the action of distinct refuge areas in the first diversifications of the group. This makes the second scenario depicted above especially attractive, and allows the speculation of an intimate association between the supposed localized first speciation events and the possible refuges.

#### Genealogical view of character evolution within the *D. mesophragmatica* group

Nacur (1958) proposed to divide the *mesophragmatica* group into two subgroups based on the direction of the basal scutellar setae, considering that *D. viracochi* has convergent basal scutellars whereas the remaining members of the group have divergent basal scutellars. Applying this character distribution to our TET, it can be easily inferred that convergent basal scutellars is a plesiomorphic state within the group, given that most of the *repleta* group species have convergent basal scutellars (Dobzhansky and Pavan, 1943; Patterson, 1943; Vilela, 1983).

In relation to the *mesophragmatica* group metaphase chromosomes, Brncic (1957) have pointed out that *D. gaucha*, *D. pavani* and *D. viracochi* all have one pair of V-shaped chromosomes, three pairs of rods and a pair of dots. Nevertheless, Brncic et al. (1971) described that in *D. gasici* and *D. brncici* the dots are more enlarged, having the appearance

of a small pair of rods. According to these authors, the dots presented a tendency of increasing the size due to the addition of extra heterochromatin, which is even more evident in *D. mesophragmatica*. Applying this distribution again in our TET, we can corroborate the hypothesis developed by Brncic et al. (1971) of a tendency towards increasing the size of the dots as more derived species are considered. So, a small pair of dots as presented by *D. gaucha*, *D. pavani* and *D. viracochi* can be considered the plesiomorphic state, while the dots resembling a pair of rods is one sinapomorphic state shared by the *mesophragmatica* subgroup species.

Also according to Brncic et al. (1971) *D. pavani* and *D. gaucha* differ by at least 11 inversions from the other members of the group. So, besides this state can be considered a sinapomorphy shared by the members of the *gaucha* subgroup, it can also be deduced that this lineage was submitted to high levels of rearrangements, only a small subset of which were selected for, since no intermediates of these paracentric inversions were encountered (Brncic et al., 1971). Maybe this could be the effect of selection acting under individuals exploring new habitats - as represented by the Andes southwards - and even leaving the Andean Mountain System, as is the case of some descendants of this lineage, given that *D. gaucha* can be found in Argentina, Uruguay and also in Brazil.

### **Acknowledgments**

We would like to thank Paloma Rubin for help with the sequencing reactions and Dr. Paulo Hoffman for contributions and suggestions concerning the text. This study was supported by research grants and fellowships from CNPq.

**This study is dedicated to the memory of Dr. Danko Brncic, in honor to his pioneer work on the biology, genetics and evolution of the *Drosophila mesophragmatica* group.**

## **REFERENCES**

- Akaike H (1974) A new look at the statistical model identification. IEEE T Automat Contr 19: 716-723
- Bächli G (2007) Taxodros: The database on taxonomy of Drosophilidae, v.1.03, Database 2007/06. <http://taxodros.unizh.ch/>. Last accessed on 15/08/2007
- Brncic D (1957) A comparative study of chromosomal variation in species of the *mesophragmatica* group of *Drosophila*. Genetics 42: 789-805
- Brncic D, Koref-Santibañez S (1957) The *mesophragmatica* species group of *Drosophila*. Evolution 11: 300-310
- Brncic D, Nair PS, Wheeler MR (1971) Cytotaxonomic relationships within the *mesophragmatica* species group of *Drosophila*. Univ Texas Publ 7103: 1-16
- Caletka BC, McAllister BF (2004) A genealogical view of chromosomal evolution and species delimitation in the *Drosophila virilis* species subgroup. Mol Phylogenet Evol 33: 664-670
- Carrasco SF, Prado LF, Godoy-Herrera R (2003) Molecular phylogeny of the *mesophragmatica* species group inferred from Cytochrome Oxidase II sequence. D I S 86: 72-75

- Dobzhansky T, Pavan C (1943) Studies on Brazilian species of *Drosophila*. Bolm Fac Filos Cienc S Paulo 36: 7-72
- Duda O (1927) Die Sudamerikanischen Drosophiliden unter Beruchsichtigung auch der anderen Neotropischen sowie Nearktischen Arten. Archiv fur Naturgeschichte 91(A)11/12: 1-229
- Durando CM, Baker RH, Etges WJ, Heed WB, Wasserman M, DeSalle R (2000) Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. Mol Phylogenet Evol 16: 296- 307
- Felsenstein J (1985) Confidence limits on phylogenies: an approaching using bootstrap. Evolution 39: 783-791
- Hunter AS, Hunter RA (1964) The *mesophragmatica* species group of *Drosophila* in Colombia. Ann Entomol Soc Am 57: 732-736
- Jaeger CP, Salzano FM . (1953) *Drosophila gaucha*, a new species from Brazil. Rev Bras Biol 13: 205-208
- Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple Sequence Alignment with ClustalX. Trends Biochem Sci 23: 403-405
- Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120
- Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics 5: 150-163

- Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear gene-trees. *Evolution* 49: 718-726
- Nacur J (1958) Genitalia masculina de *Drosophila* do grupo *mesophragmatica* (Diptera). *Rev Bras Biol* 18: 243-249
- Nair PS, Brncic D, Kojima KI (1971) II. Isozyme variations and evolutionary relationships in the *mesophragmatica* species group of *Drosophila*. *Stud Genet* VI: 17-28
- O'Grady P (1998) Reevaluation of phylogeny in the *Drosophila obscura* species group based on combined analysis of nucleotide sequences. *Mol Phylogenet Evol* 12: 124-139
- Patterson JT (1943) The Drosophilidae of the Southwest. *Univ Texas Publ* 4313: 7-216
- Posada C, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818
- Remsen J, O'Grady P (2002) Phylogeny of Drosophilinae (Diptera: Drosophilidae) with comments on combined analysis and character support. *Mol Phylogenet Evol* 24: 249-264
- Robe LJ, Valente VLS, Budnik M, Loreto ELS (2005) Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: A nuclear versus mitochondrial gene approach. *Mol Phylogenet Evol* 36: 623-640
- Rodriguez F, Oliver JL, Marin A, Medina JR (1990) The general stochastic model of nucleotide substitution. *J Theor Biol* 142: 485-501

- Russo CAM, Takezaki N, Nei M (1995) Molecular phylogeny and divergence times of *Drosophilid* species. *Mol Biol Evol* 12: 391-404
- Salgado-Laboriau ML (1994) O Período Quaternário. In: “História Ecológica da Terra” Ed Salgado-Laboriau ML, Editora Edgard Blücher, São Paulo, pp 255-280
- Saitou N, Nei M (1987) The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 9: 406-425
- Sassi AK, Heredia F, Loreto ELS, Valente VLS, Rohde C (2005) Transposable elements *P* and *gypsy* in natural populations of *Drosophila willistoni*. *Genet Mol Biol* 28: 734-739
- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene- Sequences and a compilation of conserved polymerase chain reaction primers. *Ann Entomol Soc Am* 87: 651-701
- Staden R (1996) The Staden sequence analysis package. *Mol Biotechnol* 5: 233-241
- Swofford DL (2003) PAUP: Phylogenetic Analysis using Parsimony (and other methods). Version 4.0. Sinauer Associates, Massachusetts
- Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C biases. *Mol Biol Evol* 9: 678-687
- Tamura K, Nei M (1993) Estimation of the number of base nucleotide substitution in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10: 512-526
- Tatarenkov A, Ayala FJ (2001) Phylogenetic relationships among species groups of the *virilis-repleta* radiation of *Drosophila*. *Mol Phylogenet Evol* 21: 327-331



- Tatarenkov A, Zurovcová M, Ayala FJ (2001) *Ddc* and *Amd* sequences resolve phylogenetic relationships of *Drosophila*. *Mol Phylogenet Evol* 20: 321-325
- Throckmorton LH (1975) The phylogeny, ecology and geography of *Drosophila*. In: “Handbook of Genetics”, King RC, Plenum, New York, pp 421-469
- Vela D, Rafael V (2004) Three new andean species of *Drosophila* (Diptera, Drosophilidae) of the *mesophragmatica* group. *Iheringia* 94: 295-299
- Vilela CR (1983) A revision of the *Drosophila repleta* species group (Diptera, Drosophilidae). *Rev Bras Entomol* 27: 1-114
- Zharkikh A (1994) Estimation of evolutionary distances between nucleotide sequences. *J Mol Evol* 39: 315-329

### Figure Legends:

Fig. 1. Neighbor Joining trees resulting from the analysis of: A, the *Adh* data set; B, the *Amd* data set; C, the *Ddc* data set; D, the *Hb* data set; E, the *COII* data set; F, the combined data set. Bootstrap supports (NJ, ML and MP) are indicated above each node. When a given clade was not recovered or a polytomy was obtained, a – sign is depicted. The branch lengths are proportional to the scale given in substitutions per nucleotide. Abbreviations of *D. gasici* populations: AR – Arica, Chile; CA – Cochabamba, Bolivia; CO – Bogota, Colombia.

Fig. 2. Known geographic distribution of the six *D. mesophragmatica* group species utilized in this work, according to Bächli (2007).

Table 1. Taxonomic list of *Drosophila* species used in this study, with their collection locality and Genbank accession numbers

Group	Subgroup	Species	Collection Locality	Genbank Accession Numbers				
				<i>Adh</i>	<i>Amd</i>	<i>Ddc</i>	<i>Hb</i>	<i>COII</i>
<i>mesophragmatica</i>	<i>mesophragmatica</i>	<i>D. brncici</i>	Bogotá, Colombia	<u>EF560573</u>	<u>EF560566</u>	<u>EF559365</u>	<u>EF559356</u>	AY847757
		<i>D. gasici</i>	Cochabamba, Bolivia	<u>EF560575</u>	<u>EF560568</u>	<u>EF559367</u>	<u>EF559358</u>	AY847761
			Arica, Chile	<u>EF560574</u>	<u>EF560567</u>	<u>EF559366</u>	<u>EF559357</u>	AY847760
			Bogotá, Colombia	<u>EF560576</u>	<u>EF560569</u>	<u>EF559368</u>	<u>EF559359</u>	AY847762
		<i>D. mesophragmatica</i>	La Paz, Bolivia	<u>EF560577</u>	<u>EF560572</u>	<u>EF559369</u>	<u>EF559362</u>	AY847769
		<i>D. gaucha</i>	Campos do Jordão, Brazil	<u>EF560579</u>	AF324955	AF324971	<u>EF559360</u>	AF478423
	<i>D. pavani</i>	La Florida, Chile	<u>EF560580</u>	<u>EF560570</u>	<u>EF559370</u>	<u>EF559363</u>	AY847771	
	<i>viracochi</i>	<i>D. viracochi</i>	Bogotá, Colombia	<u>EF560578</u>	<u>EF560571</u>	<u>EF559371</u>	<u>EF559364</u>	AY847773
	<i>repleta</i>	<i>hydei</i>	<i>D. hydei</i>	X58694	AF293712	AF293737	<u>EF559361</u>	AF478429

Notes:

- (1) We followed here the classification proposed by Nacur (1958) and recently renamed by Vela and Rafael (2004).
- (2) Newly obtained sequences are underlined.

Table 2. Summary of the properties of the five set of primers used in this study

Locus	Primers	Amplicon Size	Amplicon content	Analysed Size	Amplicon content	Annealing Temperature	Reference
<i>Adh</i>	4682 and 4683	≈ 771bp	End of first exon, first intron, second exon, second intron and beginning of third exon	631bp	End of first exon, first intron, second exon, second intron and beginning of third exon	55°C	O'Grady, 1998
<i>Amd</i>	Amd-un2 and Amd-bw	1307bp	End of third exon, third intron and part of fourth exon	1211bp	End of third exon, third intron and part of fourth exon	58°C	Tatarenkov et al., 2001
<i>Ddc</i>	BPF and BPR	1183bp	Part of fourth exon	1105bp	Part of fourth exon	Touch down 60-55°C	Tatarenkov et al., 1999
<i>Hb</i>	Hb106F and Hb903R	≈ 820bp	Part of third exon	687bp	Part of third exon	Touch down 58-53°C	Newly designed
<i>COII</i>	TL2-J-3037 and TK-N-3785	786bp	Part of tLeu, 684bp of the entire COII gene, part of tLys	672bp	Part of COII gene	55°C	Simon et al., 1994

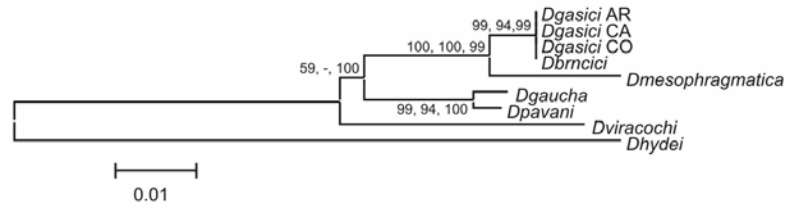
Table 3. Properties of the five molecular datasets

	<i>Adh</i>	<i>Amd</i>	<i>Ddc</i>	<i>Hb</i>	<i>COII</i>	Combined
Number of positions	631 bp	1233 bp	1105 bp	687 bp	672 bp	4306 bp
Number of variable sites	110 bp	198 bp	156 bp	96 bp	125 bp	685 bp
Number of parsimoniously informative sites	26 bp	54 bp	43 bp	27 bp	62 bp	212 bp
Nucleotide frequencies						
- A	26.5	21.8	20.6	26.2	31.7	24.4
- C	26.9	24.7	27.0	34.5	14.5	25.5
- G	23.1	29.7	29.5	23.2	13.5	25.1
- T	23.5	23.8	22.9	16.2	40.3	24.9
Transition/Transversion ratio	0.9	1.8	2.2	1.3	1.4	1.5
NJ model <sup>a</sup>	K2P	T3P	T3P	TrN	TrN	K2P
ModelTest best fit model <sup>b</sup>	SYM + G	TVM + I	TrN + I	TrN + I	GTR + I	SYM + G
Proportion of invariable sites	-	0.4798	0.5698	0.7209	0.7168	-
Gamma distribution shape parameter	0.4674	-	-	-	-	0.3019
Rate matrix						
- rAC	0.3713	3.3324	1.0000	1.0000	12.9211	1.8486
- rAG	1.2004	8.1065	3.6807	2.3602	40.9757	4.2673
- rAT	1.6801	3.0115	1.0000	1.0000	20.0944	2.6091
- rCG	0.8397	1.6551	1.0000	1.0000	0.0000	1.2585
- rCT	2.7682	8.1065	6.7409	3.9436	134.5600	7.5862

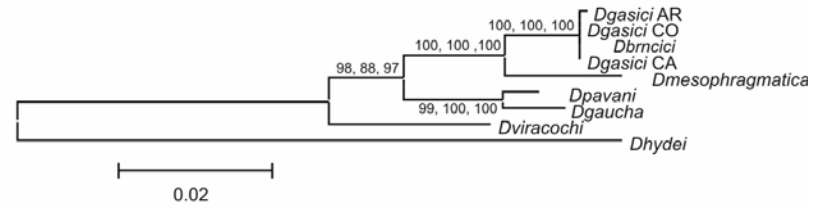
<sup>a</sup> NJ model was selected according to base frequencies and transition/transversion ratio bias, which were calculated with MEGA3.1; <sup>b</sup> Best fit model parameters were calculated with the Akaike Information Criterion (AIC) (Akaike, 1974) in ModelTest 3.7: K2P, Kimura 2-Parameter model (Kimura, 1980); T3P, Tamura 3-Parameter model (Tamura, 1992); TrN, Tamura-Nei model (Tamura and Nei, 1993); SYM, symmetrical model (Zharkikh, 1994); TVM, transversional model; GTR, General Time Reversible (Rodriguez et al., 1990); I, proportion of invariable sites; G, shape parameter of gamma distribution.

Figure 1

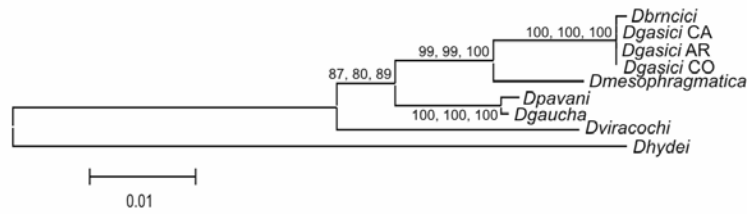
A)



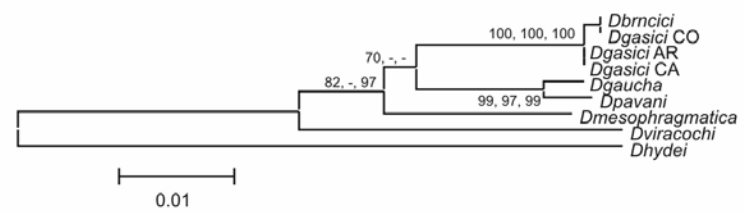
B)



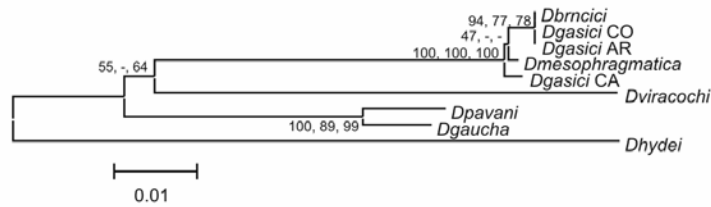
C)



D)



E)



F)

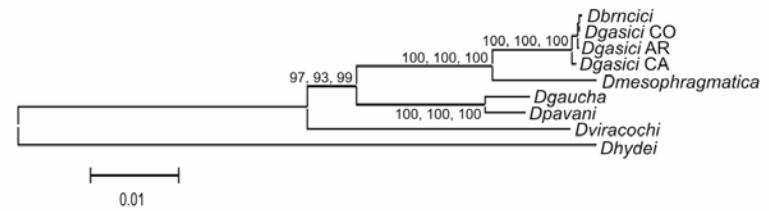


Figure 2



## CAPÍTULO V

---

Trabalho a ser submetido à revista Journal of Zoological Systematics and Evolutionary Research

### **Taxonomic boundaries and the phylogenetic relationships within the *Drosophila willistoni* subgroup**

Lizandra J. Robe<sup>1\*</sup>, Juliana Cordeiro<sup>1</sup>, Elgion L. S. Loreto<sup>1,2</sup>, Vera L. S. Valente<sup>1</sup>

1 – Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul (UFRGS), CP 15053, Porto Alegre, Rio Grande do Sul, Brazil, 91501-970

2 – Departamento de Biologia, Universidade Federal de Santa Maria (UFSM), Camobi Km 09, Santa Maria, Rio Grande do Sul, Brazil, 97105-900

\* Corresponding author: Phone (55)(51) 3308-6713; Fax (55)(51) 3308-7311

E-mail addresses: [lizbiogen@yahoo.com.br](mailto:lizbiogen@yahoo.com.br) (LJ Robe), [jlcd@yaho.com.br](mailto:jlcd@yaho.com.br) (J Cordeiro), [elgion@base.ufsm.br](mailto:elgion@base.ufsm.br) (ELS Loreto), [vera.gaiesky@ufrgs.br](mailto:vera.gaiesky@ufrgs.br) (VLS Valente).

## Abstract

The *Drosophila willistoni* group comprises 23 species, which are subdivided into three different subgroups: *alagitans*, *bocainensis* and *willistoni*. The *willistoni* subgroup, particularly, represents a complex with varying taxonomic levels. It encompasses *D. willistoni*, and its five sibling species: *D. equinoxialis*, *D. insularis*, *D. paulistorum*, *D. pavlovskiana* and *D. tropicalis*. Of these, *D. equinoxialis*, *D. tropicalis* and *D. willistoni* present differentiation at the subspecific level, whereas *D. paulistorum* represents a superspecies, compounded by six semispecies: Amazonian, Andean-Brazilian, Centro-American, Interior, Orinocan and Transitional. Despite the importance presented by this group in previous, present and likely in future biological studies, its evolutionary history still presents a series of opened questions. Aiming to contribute to the understanding of the evolution of this challenging group of species, we gathered nucleotide sequences from the mitochondrial *cytochrome oxidase II – COII* gene, and from the nuclear *alcohol dehydrogenase - Adh*, *hunchback - Hb* and *alpha methyl dopa – Amd* partitions. All these data sets, along with previously published ones from the *cytochrome oxidase I – COI* and *period – per* loci were analyzed both individually and simultaneously through different phylogenetic methods. Although high levels of incongruence were detected among some partitions, our data finally lead to a robust tree which further refines several aspects related to the *willistoni* subgroup phylogeny. In this respect, *D. insularis*, *D. tropicalis*, *D. willistoni* and *D. equinoxialis* successively branched off from the *willistoni* subgroup main stem, which recently subdivided to originate *D. paulistorum* and *D. pavlovskiana*. The taxonomic property of this last taxon is, nevertheless, questioned, since it seems to derive from the recent *D. paulistorum* radiation exactly in the same way as the other semispecies do. Furthermore, there are evidences that Pleistocene Amazonian refuges acted in the origin and evolution of these species still in *statu nascendi*.

**Key-words:** *Drosophila willistoni* subgroup; *D. paulistorum* superspecies; semispecies; phylogeny; divergence dating estimates.



## INTRODUCTION

The *willistoni* group of *Drosophila*, together with the *saltans* group, represents the New World radiation of the *Sophophora* subgenus (O'Grady and Kidwell 2002; Throckmorton 1975). This group currently encompasses 23 species, which are subdivided into three main subgroups (Bächli 2008): *alagitans* (five species), *bocainensis* (twelve species) and *willistoni* (six species). The *willistoni* subgroup comprises *D. willistoni* and its five sibling species: *D. equinoxialis*, *D. insularis*, *D. paulistorum*, *D. pavlovskiana* and *D. tropicalis*. Moreover, according to Spassky et al. (1971), there may be a seventh sibling species, the Carmody strains (Carmody 1965), which appears to be intermediate between *D. paulistorum* and *D. equinoxialis*, since it crosses with some but not with other strains of both these species.

As the *willistoni* subgroup species are very similar in their external morphology (Burla et al. 1949), several attempts have been conducted in order to improve their differentiation with basis on: external male genitalia (Spassky 1957), crossing tests (Spassky et al. 1971), inspection of polytenic chromosomes (Rohde et al. 2006) and allozyme variations of acid phosphatase-1 (*AcpH-1*) (Garcia et al. 2006). Despite this high similarity, these sibling species generally fail to cross hybridize, since they exhibit varying degrees of ethological isolation (Ehrman and Powell 1982). According to Spassky et al. (1971), the most striking exception to this rule seems to be the crossing between *D. paulistorum* and *D. pavlovskiana*, which produces sterile hybrid males and fertile hybrid females. Winge and Cordeiro (1995), nevertheless, report the occurrence of interespecific crosses with varying degrees of reproductive affinities between the other species of the sibling group. Actually, these authors detect that even fertile hybrid males and females

could be obtained in the crosses involving *D. paulistorum* Transitional females and *D. willistoni* males.

Another level of taxonomic complexity is added to the *willistoni* subgroup by the composition of the *D. paulistorum* superspecies, which is subdivided into six indistinguishable semispecies or species in *statu nascenti* (Dobzhansky and Spassky 1959; Pérez-Salas et al. 1970): Amazonian, Andean-Brazilian, Centroamerican, Interior, Orinocan and Transitional. These semispecies exhibit high levels of ethological isolation, but when they cross, unviable hybrids or sterile males and fertile females are produced (Ehrman and Powell 1982). Exceptions to this rule are provided by the crosses involving the Transitional semispecies and the Andean-Brazilian or the Centroamerican ones, which produce viable and fertile offspring in some cases (Dobzhansky et al. 1969).

Furthermore, the *willistoni* subgroup also presents genetic divergence at the subspecies level. In this way, *D. willistoni* differentiates into the *willistoni* and *quéchua* subspecies (Ayala and Tracey 1973); *D. equinoxialis* presents the *equinoxialis* and *caribbensis* subspecies (Ayala et al. 1974); and *D. tropicalis* subdivides into the *tropicalis* and *cubana* subspecies (Townsend 1954). As these are geographically isolated taxa, they present a lower degree of ethologic isolation, although also only sterile hybrid males are produced in most of their crossings.

Hence, the *willistoni* subgroup is a complex of various taxonomic levels, containing a sequential array of evolutionary stages (Ehrman and Powell 1982). It can thus be considered as an ideal material for all sorts of studies into the processes and products of evolution, providing uncommon opportunities for studies on the genetics of the speciation process. The *willistoni* group species also presents other remarkable properties, as their peculiar substitution patterns, characterized by a high A+T content (Tarrío et al. 2000;

Rodríguez-Trelles et al. 2000a and 2000b; Powell et al. 2003), and their singular methylation process, possibly involved in sex-specific dosage compensation (Garcia et al. 2007).

Despite this, however, the *willistoni* subgroup phylogeny is not completely understood. Spassky et al. (1971) summarized the biogeographic, genetic, cytological and biochemical evidences available for the group at that time in a schematic diagram which provides some insights into the species and semispecies relationships. Later, Ayala et al. (1974) developed a dendrogram for the *willistoni* subgroup species and semispecies based on genetic differentiation at the allozyme level. More recently, Gleason et al. (1998) and Tarrío et al. (2000) clarified several aspects related to the phylogenetic relationships within the subgroup by studying different molecular markers (the nuclear *period – per* and *alcohol dehydrogenase – Adh* and the mitochondrial *cytochrome oxidase I – COI* in the first; the nuclear *xanthine dehydrogenase – Xdh* in the second). Nevertheless, the conflicting results presented by some of these studies, added to the absence of an explicit molecular hypothesis related to the semispecies relationships, leaves several unanswered questions concerning the evolution of the *willistoni* subgroup.

Thus, in the present study we seek to contribute to the understanding of the evolutionary scenario that led to the origin and diversification of this outstanding group of species. We also aim to test the molecular divergence levels presented by several of its constituent taxa and its correlation with the degree of reproductive isolation between them. In this way, we have gathered additional molecular sequences from the mitochondrial *cytochrome oxidase II – COII* gene, and from the nuclear *Adh*, *hunchback - Hb* and *alpha methyl dopa – Amd* partitions. Examining these data, along with previously published ones from the *COI* and *Per* loci (Gleason et al. 1998), in both individual and combined analyses,

led to a strongly supported tree which further refines several aspects related to the *willistoni* subgroup phylogeny. Nevertheless, the absence of congruent levels of divergence between the *D. paulistorum* semispecies, added to the paraphyletic status presented by some of the focused taxa, and to the small internal branches evidenced specially in the derived part of the recovered trees further endorses the proposed complex and recent patterns of evolution presented by the *willistoni* subgroup, supporting the notion of introgression and/or ancestral polymorphisms incompletely sorted due to the occurrence of rapid radiations.

## **MATERIALS AND METHODS**

### **1. Species**

Wildtype stocks from five of the six *D. willistoni* sibling species, in addition to stocks of the six *D. paulistorum* semispecies, were included in this study (Table 1), and are maintained in a  $17^{\circ}\text{C} \pm 1^{\circ}\text{C}$  chamber in our laboratory. Unfortunately, we could not obtain stocks from *D. pavlovskiana* or from the Carmody strain, since these are rare endemic and no strain is now available at any Stock Center. *D. sturtevantii*, a member of the *saltans* group, which is considered the *willistoni* group sister lineage (O'Grady and Kidwell 2002; Throckmorton 1975), and/or species from the *bocainensis* subgroup, were utilized as outgroup taxa.

### **2. DNA manipulation**

Genomic DNA was isolated according to the protocol described by Sassi et al. (2005). The target loci *Amd*, *Adh*, *Hb* and *COII* were amplified from each taxon using standard PCR conditions, with the use of the oligonucleotides described in Mota et al. (in press). The obtained PCR products were checked by horizontal electrophoresis in 0.8%

agarose gel stained with ethidium bromide. The *Amd* visualized amplicons were excised and purified using the Qiaquick Gel Extraction System (Qiagen). *Adh*, *Hb* and *COII*, on the other hand, had their fragments purified with the use of a solution containing PEG 13% and NaCl 1.6M. The sequencing reactions were then directly performed on a MegaBACE 500 automatic sequencer, through the use of the DYEnamic ET kit (Amersham) and each of the amplification primers. All *Per* and *COI* sequences were obtained from GenBank and their accession numbers are listed in Table 1, together with the other gene sequences references.

### 3. Data analysis

First the electropherograms were assembled and visualized through the use of the Staden Package Gap4 program (Staden 1996), where the variable sites within a taxon were coded as polymorphic. The contigs obtained for each partition were then joined to previously published ones. The alignment of the *COI*, *COII*, *Adh* and *Amd* loci was trivial, and accomplished using the ClustalW algorithm as implemented in Mega 3.1 (Kumar et al. 2004). The more divergent sequences involving the *Hb* and *Per* coding regions were aligned through the use of the Muscle software (Edgar 2004). The alignments obtained for each of the six different partitions were individually tested for substitutional saturation in DAMBE 4.5.22 (Xia and Xie 2001).

The qui-square test implemented in PAUP version 4.0b10 (Swofford 2003) was performed in order to test for base frequency equilibrium across taxa. Besides, the Mega 3.1 program was used to calculate p distances and other empirical descriptive parameters. Additional parameters related to the best-fit substitution model for the data under a likelihood framework were obtained in the ModelTest 3.7 (Posada and Crandall 1998) out-file.

Phylogenetic analyses were first performed on each of the six different data sets taken individually. Combined analyses were also done in three different ways: the mitochondrial concatenated partitions; the nuclear concatenated partitions; all loci combined. PAUP 4.0b10 was employed for the Maximum Parsimony (MP) and the Neighbor Joining (NJ) analyses, whereas PhyML 2.4.4 (Guindon and Gascuel 2003) was used for the Maximum Likelihood inferences. The Bayesian analyses were performed on MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001) for the individual and combined data sets and on BayesPhylogenies (Pagel and Meade 2004) only for the combined ones. Anyway, MP analyses were performed through an equal weight criterion using either the branch-and-bound or the heuristic algorithm (only for the individual *Hb* and *Adh* data sets). These heuristic searches were carried out with 100 replicates of random stepwise addition, followed by tree-bisection-reconnection (TBR) branch swapping. ML and NJ searches followed the model selected by the Akaike Information Criterion (AIC) (Akaike 1974) executed in the ModelTest 3.7. The ML ones, specifically, were performed using as input a tree generated by a NJ search, with the model parameters estimated and optimized from the data. The MrBayes inferences, otherwise, followed the individual MrModeltest 2.2 (Nylander 2003) AIC output, and all parameters were unlinked before running the combined analysis. The BayesPhylogenies analyses, on the other hand, were always conducted through the General Time Reversible (GTR) (Rodriguez et al. 1990) model with rate heterogeneity corrected through a discrete gamma distribution, with the number of patterns and the number of different rates set to the number of independent matrixes. In both Bayesian analyses, Markov Chain Monte Carlo (MCMC) algorithm was implemented through at least 1,000,000 generations, with trees saved every 1,000 generations.

Support for each clade in the ML, NJ and MP trees was determined by the bootstrap proportions (Felsenstein 1985) with 1,000 replications, whereas for the Bayesian inferences the posterior probability (PP) of each clade on the 50% majority rule consensus tree was measured. Furthermore, the contribution of each individual partition to each node recovered by the total evidence tree (TET) was assessed through the Partitioned Bremer Support scores (PBS) (Bremer 1988 and 1994), as performed by TreeRot 3.0 (Sorenson and Franzosa 2007) and PAUP 4.0b10.

Before combining the individual data sets, pairwise Partition Homogeneity Tests (PHT) (Farris et al. 1994 and 1995), as implemented in PAUP 4.0b10, were conducted to test for incongruence between the different loci. Additionally, the alternative topologies obtained under the different data sets were compared using the Templeton test (Templeton 1983), performed under the parsimony criteria, and the Shimodaira-Hasegawa test (Shimodaira and Hasegawa 1999) performed under the likelihood criteria. The Shimodaira-Hasegawa topology test was implemented using RELL bootstraps with 10,000 replicates.

Divergence times were also estimated through the use of the *Adh* third codon positions divergences and the evolutionary rate established by Russo et al. (1995) for these as  $1.0 \times 10^{-8}$  per site per year per lineage. Nevertheless, these estimates were obtained only after evaluating constancy of evolutionary rates across the *willistoni* group *Adh* tree through a molecular clock hierarchical Likelihood Ratio Test (hLRT).

## RESULTS

### 1. The data sets

Partial sequences from six different loci, encompassing more than 4,600 bp, were analyzed for at least twelve OTU's (Operational Taxonomic Units) traditionally placed at the *willistoni* group. The properties of each of these six different data sets are presented in Table 2, together with the properties of the three concatenated data sets. In general, although the qui-square test did not detect any significant signal of base composition heterogeneity between taxa, all individual partitions showed a significant nucleotide frequency bias, and also a detectable deviation in the transition/transversion ratio. Besides that, all matrixes presented significant levels of heterogeneity in substitution rates among sites, which were corrected for through the implementation of a proportion of invariable sites (I option) and/or a continuous gamma distribution (G option). The nucleotide substitution patterns were some skewed in all data sets, although the nuclear *Adh* and *Amd* were, in one extreme, the most symmetrical, whereas the mitochondrial *COI* and *COII* detached as the most assymetrical partitions. All these factors appear to have acted together to produce some levels of homoplasy and saturation, even at lower taxonomic ranges (see bellow).

### 2. Phylogenetic inferences derived from the mitochondrial markers

Phylogenetic analyses individually conducted on each of the two mitochondrial partitions, as well as on a concatenated mitochondrial data set (Fig. 1), yielded extremely incongruent results, although in general the Bayesian, ML, NJ and MP methods recovered similar topologies within each of these data sets. Hence, the major source of agreement



between the three different topologies concerns the *bocainensis* subgroup polyphyly, which is contrasted to the *willistoni* subgroup monophyly. In this sense, within the *bocainensis* subgroup, the sister species relationship between *D. capricorni* and *D. sucinea*, to the exclusion of *D. nebulosa*, is recurrently recovered. Within the *willistoni* subgroup, nevertheless, the picture is more confusing. Both *COI* and *COII* data sets, for example, present the *D. paulistorum* superspecies entirely paraphyletic in respect to the other five sibling species, although the mitochondrial combined analysis recovered the *D. paulistorum* paraphyly only in respect to *D. equinoxialis* and *D. pavlovskiana*.

Concerning the *willistoni* subgroup individual OTU's relationships, Figure 1A shows the Bayesian phylogeny recovered under the *COI* data set. This partition provides unexpected results, presenting the *D. paulistorum* Amazonian semispecies as the *willistoni* subgroup oldest lineage (PP of 0.72), followed by the clade joining *D. insularis* and *D. tropicalis* (PP of 0.67). *D. willistoni* is the next to branch off (PP of 0.92), and its sister clade furthest divide into two main clades: (1) the first presenting the Carmody strain as closer to *D. paulistorum* Andean-Brazilian semispecies than to *D. equinoxialis* (PP of 0.99); (2) the second clustering *D. paulistorum* Orinocan and Transitional semispecies as sister to the clade joining *D. paulistorum* Centro-American semispecies and *D. pavlovskiana* with *D. paulistorum* Interior semispecies (PP of 0.97).

On the other hand, although *COII* agrees with *COI* in respect to the *D. paulistorum* paraphyly, its recovered phylogenetic relationships presented in Figure 1B are entirely different. In this case, a weakly supported clade (PP of 0.57) encompassing *D. equinoxialis* in addition to the homo-sequential *D. pavlovskiana* and *D. paulistorum* Centro-American and Transitional semispecies is the early offshoot (PP of 0.81). The remaining OTU's split off to divide into two sister lineages: (1) the first grouping the two *D. tropicalis* subspecies

with *D. insularis* and *D. willistoni* (PP of 1.00); (2) the second presenting the four strains of *D. paulistorum* Andean-Brazilian semispecies (three of which are also homo-sequential – AB, RP and JAI) and the two strains of *D. paulistorum* Amazonian semispecies together with the *D. paulistorum* Interior and Orinocan semispecies (PP of 0.99).

Although *COI* and *COII* presented these conflicting phylogenetic signals – also detected through a significant PHT result (Table 3) -, combined analyses were performed in order to overwhelm this heterogeneity (Rokas et al. 2003). Indeed, the mitochondrial total evidence tree (Fig. 1C) relaxes some of the above mentioned incongruities in recovering *D. tropicalis*, *D. insularis* and *D. willistoni* as an independent branch (PP of 0.99). However, *D. paulistorum* remains paraphyletic, and most of its internal relationships are not strongly supported, better resembling the *COII* topology as concerns its major subdivision.

Analyzing the saturation plots presented above each individual partition phylogeny (Fig. 1), it can be seen that there are significant evidences of transitional saturation both in the *COI* and in the *COII* data sets. So, the presented incongruent results, as also as the low statistical support observed in some of the recovered clades, could be attributed to high levels of homoplasy and multiple substitutions, characterizing a significant saturation detectable even in lower levels of divergence.

### 3. Phylogenetic inferences derived from the nuclear markers

The phylogenetic analyses performed on the nuclear partitions besides presenting no so strong evidences of substitutional saturation, provide a clearer and more robust picture of the *willistoni* subgroup evolution (Fig. 2). The lower levels of heterogeneity presented by the nuclear partitions were also evidenced through the PHT, which did not detect significant levels of incongruence in most pairwise comparisons (Table 3) – exception

made by the *per* versus *Amd* test. Moreover, as previous studies have demonstrated that combining data sets can improve phylogenetic accuracy (Baker and DeSalle 1997; Remsen and DeSalle 1998), combined approaches were also here used.

Nevertheless, as concerning the *bocainensis* subgroup, some conflicts are retained, and there is only a tenuous support to its monophyly. As for the mitochondrial data, the phylogenetic affinity between *D. sucinea* and *D. capricorni* is well supported, whereas the positioning of *D. fumipennis* and *D. nebulosa* are highly unstable.

The monophyly and the phylogenetic relationships within the *willistoni* subgroup are, nevertheless, better resolved. *D. insularis* is always defined as the *willistoni* subgroup early offshoot, whereas the other siblings are recovered in a nested succession. In this case, *Adh*, *Amd*, *per* and the nuclear simultaneous analysis all favor *D. tropicalis* as the next to branch off following *D. insularis*, although *D. willistoni* occupies this positioning in the *Hb* analyses. Thus, except for *Hb*, *D. willistoni* is generally presented as the sister species of a main clade that ensures the phylogenetic affinities between *D. equinoxialis*, *D. pavlovskiana* and *D. paulistorum*. Hence, whenever included in the analyses, *D. pavlovskiana* appears embedded among the *D. paulistorum* strains, so as to comprise a big clade closely related to *D. equinoxialis* (Figs. 2A and 2D). The Carmody strain nuclear sequences, on the other hand, appear to belong to the *D. equinoxialis* lineage (Fig. 2D).

As concerns the relationships between the *D. paulistorum* semispecies, a better resolved picture is also in general accomplished. In this sense, a paraphyletic Andean-Brazilian semispecies is recovered by both *Adh* and *Amd*, which together with *per* and the nuclear simultaneous analyses suggest a close relationship between this semispecies and the Orinocan one. *Amd*, *per* and the nuclear simultaneous analyses also agree in the clustering of the Centroamerican, the Interior and the Transitional semispecies. The Amazonian

semispecies, nevertheless, besides paraphyletic, presents a more floating positioning, being presented as the Centroamerican sister semispecies by *Adh*, but comprising the *D. paulistorum* superspecies early offshoot according to *Amd*, *per* and the nuclear combined analyses. The short internal branches evidenced in this part of the *willistoni* subgroup topology suggest that incomplete lineage sorting may be affecting the obtained results. Moreover, *Hb* completely failed to establish any relationship within the *D. paulistorum* complex.

#### 4. The total evidence phylogeny

As can be detected by the results above depicted and also by the inspection of Table 3, there is a great deal of incongruence between the mitochondrial and the nuclear data sets. However, through an analysis of the saturation plots recovered for the mitochondrial partitions, it can be inferred that the conflict arose mainly due to differences in internal homoplasies presented by the data, although the process of incomplete lineage sorting may also be causing some discrepancies. Since combining is a way of attempting to amplify signal over noise (Baker and DeSalle 1997; Remsen and DeSalle 1998), the total evidence phylogeny obtained through a combined approach is here considered as the best phylogenetic hypothesis until now available for the *willistoni* subgroup. In accordance to this, among the nine different data sets, only *COI* rejected the TET under both the Templeton and the Shimodaira-Hasegawa Tests (Supplementary Material – Table 1).

This total evidence hypothesis presents *D. insularis* branching off at the base of the *willistoni* subgroup tree (PP of 1.00), in which it is followed by *D. tropicalis* (PP of 1.00). *D. willistoni*, on the other hand, is shown to be the sister taxon of the remainder of the subgroup (PP of 1.00), that is, *D. equinoxialis* and *D. paulistorum*, which are monophyletic

with respect to each other. As regards to the *D. paulistorum* complex internal relationships, Centroamerican shares a most recent common ancestor with Transitional, to the exclusion of the other four semispecies (PP of 1.00 and 0.98), which comprise an independent lineage (PP of 0.99 and 1.00). In this case, Orinocan and Andean-Brazilian are each other's closest relatives (PP of 1.00), whereas Amazonian and Interior branch off in intermediate positions, somehow connecting the two different *D. paulistorum* clades.

The average PBS scores presented by each partition across this total evidence phylogeny (Table 4) confirm the scenario in which even incongruent data sets can provide additional support to a combined tree. In this sense, although standardized PBS values are greater for the *per* and *Hb* partitions, only *COII* conflicts with more than three of the nine resolved TET nodes. Furthermore, only node three, which joins the Interior *D. paulistorum* semispecies with the Amazonian, Andean-Brazilian and Orinocan ones, is supported by less than two independent partitions.

## 5. The temporal pattern

After confirming by an hLRT that the molecular clock applies to the *willistoni* group *Adh* sequences ( $-2 \log\Delta = 21.81$ , 24 df,  $p > 0.5$ ), the timing of the divergence events within this group could be inferred according to Russo et al. (1995). The obtained estimations showed that the divergence between the *willistoni* subgroup and the *bocainensis* subgroup occurred at about 15 mya, whereas the *willistoni* subgroup diversification started at about 5.6 mya. On the other hand, *D. willistoni* and *D. equinoxialis* branched from the stem of the *willistoni* subgroup at about 4.6 and 3.4 mya, respectively, whereas the *D. paulistorum* radiation was even more recent, occurring at about 0.9 mya.

## 6. Molecular genetic differentiation

Testing for the molecular divergence level presented within and between different taxa provided some interesting results (Table 5), although they should be interpreted with caution given the extremely limited sample size applied at the inter-populational and inter-subspecific levels. In general, there is no strict correlation between the level of reproductive isolation and the average neutral genetic distance between taxa, at least concerning the nuclear partitions. Hence, little nuclear genetic differentiation seems to have taken place at the semispecific level. The mitochondrial partitions, on the other hand, appear to better conform the rule of increasing levels of reproductive isolation in association to the accumulation of genetic distance.

## DISCUSSION

### 1. The *willistoni* subgroup as a challenge to systematics and evolutionists

Phylogenetic relationships within the *willistoni* subgroup, and especially within the *D. paulistorum* complex, have proven here difficult to resolve. In general, the performed phylogenetic analyses were unable to produce an unambiguous phylogenetic hypothesis, especially as concerns the *D. paulistorum* superspecies internal composition. Some levels of incongruence were found between all the obtained topologies, but the mitochondrial derived hypothesis detached as the most incongruent ones. Although the phenomenon of ancestral polymorphism incompletely sorted and asymmetrical introgression can provide explanations for these kinds of observations, high levels of saturation not easily circumvented for render the immediate conclusion. This expands the poor performance presented by mitochondrial genes in recovering the phylogenetic relationships within

*Drosophila* previously reported by Goto and Kimura (2001) and Robe et al. (2005) to even lower taxonomic levels. All these results are probably related to the highly skewed base composition and to the asymmetrical transformation rate matrix presented by the two mitochondrial protein-coding genes.

Additional tests for the hypotheses of introgression and incomplete lineage sorting by *COI* and *COII* could be conducted if both the mitochondrial genes were isolated from identical lineages, since as linked markers, they should share identical histories unless saturation is the main factor. As this is not the case, saturation still provides the best explanation until now available, even in the case of *COI* which is considered the mitochondrial most conserved gene in terms of amino acid evolution (Simon et al. 1994), and the mitochondrial gene which possess the greater range of phylogenetic signal (Hebert et al. 2003). Indeed, *COI* failed to be a good reference gene to DNA barcoding species within the *willistoni* group, as is the original proposal of Hebert et al. (2003). Although the major outlying sequence (*D. paulistorum* Amazonian semispecies) could be attributed to rate acceleration instead of secondary convergence, *COI* also failed in the differentiation between *D. paulistorum* and its two allied species, *D. equinoxialis* and *D. pavlovskiana*. Moreover, *COI* diagnose abilities were also put in check by the 100% similarity presented between *D. paulistorum* Andean-Brazilian semispecies and the Carmody strain, which is supposedly a member of the *D. equinoxialis* species (Gleason et al. 1998). It appears that also within this subgroup, in agreement with Moritz and Cicero (2004), mitochondrial DNA divergence is neither necessary nor sufficient as a criterion for delineating species.

If saturation seems to be masking phylogenetic information at the mitochondrial markers, the picture looks quite different for the nuclear partitions. These markers do not present high levels of saturation, although they retained some incongruency, especially at

the inter-semispecific level. In this case, nevertheless, the problem of randomly shared ancestral polymorphisms, compounded by the scarcity of informative changes on the short internal branches of the tree seems to be the main source of conflict. Pamilo and Nei (1988) showed that, for neutral loci, the probability  $P$  that a gene tree equals a species tree is inversely proportional to the population size and directly proportional to the time between speciation events. Although we can not estimate the population size of the semispecies ancestral, our recovered trees present extremely short internal branches at their divergence points, which provides an explanation to the notion that their genomes are mosaics of conflicting genealogies as an outcome of subsequent divergence bursts or radiations. The small average genetic distance between these semispecies further suggests that these radiations are quite recent.

In our case, it is also possible that the molecular phylogenies obtained are unveiling other phenomena, as the possibility that some hybridization is taking place in the semispecies evolution. Since most of the crosses between semispecies produce sterile hybrid males but fertile hybrid females (revision in Ehrman and Powell 1982), introgression may result from backcrosses between these hybrid females and one of their parental semispecies - although also these result in male sterility (Ehrman 1960).

## 2. The *willistoni* subgroup evolutionary history and taxonomy

Despite these troubles and pitfalls, we obtained a confident description of the *willistoni* subgroup evolution. In general, the sibling species appear to have branched out in the following nested succession: *D. insularis* > *D. tropicalis* > *D. willistoni* > *D. equinoxialis* > *D. paulistorum* and *D. pavlovskiana*. Similar relationships were previously reported by Gleason et al. (1998) with the use of the *per* and the *Adh* data sets, and by



Tarrío et al. (2000) with the use of *Xdh* sequences. Nevertheless, according to Spassky et al. (1971), *D. insularis* is probably an insular offshoot of *D. tropicalis*, whereas Ayala et al. (1974) placed *D. willistoni* as the *D. tropicalis* closest sister-group, to the exclusion of *D. insularis* which is positioned as the *willistoni* subgroup oldest lineage. Otherwise, Cordeiro and Winge (1995) believed that *D. willistoni* represents a “presumed remainder of the primitive species”. The work of Rohde et al. (2006) provides cytogenetic support to our phylogenetic hypothesis in presenting *D. insularis* as the most rearranged species, at least as the IIR chromosome arm is concerned. Furthermore, that study also presents *D. willistoni* and *D. tropicalis* as independent lineages, since they do not share any inversion, in contrast to *D. equinoxialis* and *D. paulistorum*, which share inversion IIRd. Thus, not only molecular, but also cytogenetic markers seem to support our working hypothesis.

These phylogenetic evidences, nevertheless, do not reflect some taxonomic designations within the *willistoni* subgroup, especially as regard to *D. pavlovskiana* status. This taxon is characterized by the production of sterile hybrid males and fertile hybrid females when crossing with any of the *D. paulistorum* semispecies (see Ehrman and Powell 1982), so it was originally included as another semispecies (Guianan). Kastritsis and Dobzhansky (1967) further argued that it should be raised to the species level because it is genetically more distant from the other semispecies than they are from each other. Since this is not confirmed by any of our phylogenies that includes *D. pavlovskiana* sequences neither by Gleason et al. (1998), this seems to be neither a good biological species nor a phylogenetic species. Thus, we suggest that these individuals should return to the semispecies taxonomic level.

As concerns the Carmody strain, Gleason et al. (1998) found that, from a phylogenetic species perspective, the Carmody strain would be classified as *D.*

*equinoxialis*, but their *COI* sequences did not support this assertion. The available sequences only further support the Carmody strain (Carmody 1965) as intermediate between *D. paulistorum* and *D. equinoxialis*, the first being supported by *COI* and the second by *per*. Thus, its taxonomic and evolutionary status remain open questions.

Within the *D. paulistorum* complex TET, a clade comprising the Centro-American and Transitional semispecies is the oldest lineage, whereas its sister clade encompasses the other four semispecies: Interior, Amazonian, Andean-Brazilian and Orinocan. Hence, the allied relationships between the Centro-American and Transitional semispecies were previously suggested by Ayala et al. (1974) and Spassky et al. (1971), although both these studies also clustered the Andean-Brazilian strain with them. These results are in accordance with the observation that the Transitional semispecies is able to produce viable and fertile hybrids in crosses with the Andean-Brazilian and Centro-American semispecies (Dobzhansky et al. 1964, Ehrman and Powell 1982). However, Rohde et al. (2006) pointed out that the Andean-Brazilian and the Orinocan semispecies are homosequential for inversion IIRdg, supporting our conclusion that these are sister taxa. According to Ayala et al. (1974) and Spassky et al. (1971), otherwise, the Interior semispecies is Orinocan nearest relative, since they appear to exhibit a lower degree of isolation. As besides the cytogenetic data, three different partitions (*Adh*, *Amd* and *Per*) support our result, while none of the other three individual data sets recover the sister relationships between Interior and Orinocan, molecular data seem to better conform our total evidence hypothesis. In this case, Amazonian is clustered with the Orinocan and Andean-Brazilian semispecies, comprising a clade that presents the Interior semispecies as the closest outgroup.

Although robust and confident results were obtained, we need to be aware that some of the presented relationships, especially within the *D. paulistorum* complex are

incongruent between partitions and individually only tenuously supported. Furthermore, the two semispecies for which more than a single strain was sequenced were paraphyletic in most of the recovered trees. These data considered in the context of the small genetic divergence presented among different semispecies endorses the hypotheses of introgression and/or incomplete lineage sorting. Moreover, they raise the question if these individual semispecies are monophyletic at all. Perhaps, the sexual isolation presented by them is only a result of recent reinforcement (see Ehrman 1965) induced by natural selection against hybridization of incompatible lineages, whose incompatibility arose mainly through infection by different endosymbiont lineages (see Ehrman and Powell 1982). Only the inclusion of different strains encompassing the different semispecies, strengthened by the study of their endosymbiont lineages can prove this hypothesis true or false. Thus, although not totally definitive, our results also question some of the previous conclusions concerning the evolution and taxonomy of the *willistoni* subgroup.

Despite these uncertainties, taken as a whole our data support the notion that the origin and diversification of both the *willistoni* subgroup and the *D. paulistorum* complex took place in areas near the North of South America, or in the Central America. This inference is based on the fact that *D. insularis*, the most primitive member of the *willistoni* subgroup, is restricted to a few islands of the Antilles (Bächli 2008), whereas all the *willistoni* subgroup species overlap near the Central America area (Figure 4A). Moreover, the unique *D. paulistorum* semispecies which seems to reach other areas is the Andean-Brazilian one (Ehrman and Powell 1982) (Figure 4B), which occupies a derived positioning within the *D. paulistorum* complex tree. However, according to our dating estimates, the *willistoni* subgroup started its diversification at about 5.6 Mya, whereas the *D. paulistorum* semispecies diverged even more recently, at about 0.9 Mya. This places the *D. paulistorum*

radiation in the Pleistocene, conforming the hypothesis first presented by Spassky et al. (1971) that the semispecies origin may be correlated with Haffer's (1967, 1969) Amazonian refuge areas. Furthermore, these datings also raise the possibility that the formation of Panama Isthmus during the Pliocene may have influenced the evolution of the *willistoni* subgroup.

### **Acknowledgments**

We would like to thank Paloma Rubin for performing the sequencing reactions and Nina R. Mota for assistance with the English language. This study was supported by research grants and fellowships from CNPq, FAPERGS, CAPES and PROPESQ-UFRGS.

### **REFERENCES**

- Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Contr 19:716-723.
- Ayala FJ, Tracey ML (1973) Genetic differentiation and reproductive isolation between two subspecies of *Drosophila willistoni*. J Hered 64:120-124.
- Ayala FJ, Tracey ML, Hedgecock D, Richmond RC (1974) Genetic differentiation during the speciation process in *Drosophila*. Evolution 28:576-592.
- Bächli G. (2008) TaxoDros: The Database on Taxonomy of Drosophilidae, v.1.03, Database 2008/01. <http://taxodros.unizh.ch/>. Last accessed on 08/01/2008.
- Baker RH, DeSalle R (1997) Multiple sources of character information and the phylogeny of Hawaiian drosophilids. Syst Biol 46:654-673.
- Bremer K (1988) The limits of amino acid sequence data in angiosperm phylogenetics reconstruction. Evolution 42:795-803.

- Bremer K (1994) Branch support and tree stability. *Cladistics* 10:295-304.
- Burla H, DaCunha AB, Cordeiro AR, Dobzhansky T, Malogolowkin C, Pavan C (1949) The *willistoni* group of sibling species of *Drosophila*. *Evolution* 3:300-314.
- Carmody G (1965) Two unusual strains of the *D. willistoni* sibling species group. *Drosophila Inf Serv* 40:53.
- Dobzhansky T, Ehrman L, Pavlovsky O, Spassky B (1964) The superespecies *D. paulistorum*. *Proc Natl Acad Sci* 51:3-9.
- Dobzhansky T, Pavlovsky O, Ehrman L (1969) Transitional Populations of *Drosophila paulistorum*. *Evolution* 23:482-492.
- Dobzhansky T, Spassky B (1959) *Drosophila paulistorum* a cluster of species in statu nascendi. *Proc Natl Acad Sci* 45:419-128.
- Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113.
- Ehrman L (1960) Genetics of hybrid sterility in *Drosophila paulistorum*. *Evolution* 14:212-223.
- Ehrman L (1965) Direct observation of sexual isolation between allopatric and between sympatric strains of the different *Drosophila paulistorum* races. *Evolution* 19:459-464.
- Ehrman L, Powell JR (1982) The *Drosophila willistoni* species group. In: Ashburner M, Carson HL, Thompson JN (eds.), *The Genetics and Biology of Drosophila*. Academic Press, New York, pp 193-225.
- Farris JS, Källersjö M, Kluge A, Bult C (1994) Testing significance of congruence. *Cladistics* 10:315-320.
- Farris JS, Källersjö M, Kluge A, Bult C (1995) Constructing a significance test for incongruence. *Syst Biol* 44:570-572.

- Felsenstein J (1985) Confidence limits on phylogenies: an approach using bootstrap. *Evolution* 39:783-791.
- Garcia ACL, Rohde C, Audino GF, Valente VLS, Valiati VH (2006) Identification of the sibling species of the *Drosophila willistoni* subgroup through the electrophoretic mobility of acid phosphatase-1. *J Zool Syst Evol Res* 44:212-216.
- Garcia RN, D'Avila MF, Robe LJ, Loreto EL, Panzera Y, de Heredia FO, Valente VL. (2007) First evidence of methylation in the genome of *Drosophila willistoni*. *Genetica* 131:91-105.
- Gleason JM, Griffith EC, Powell JR (1998) A molecular phylogeny of the *Drosophila willistoni* group: conflicts between species concepts? *Evolution* 52:1093-1103.
- Goto SG, Kimura MT (2001) Phylogenetic utility of mitochondrial *COI* and nuclear *Gpdh* genes in *Drosophila*. *Mol Phylogenet Evol* 18:404-422.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696-704.
- Haffer J (1967) Speciation in Colombian forest birds west of the Andes. *Am Mus Novit* 294:1-57.
- Haffer J (1969) Speciation in Amazonian forest birds. *Science* 165:131-137.
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proc R Soc Lond* 270:313-321.
- Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.
- Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HM (ed.), *Mammalian protein metabolism*. Academic Press, New York, pp 21-132.

- Kastritsis CD, Dobzhansky, Th (1967) *Drosophila pavlovskiana*, a race or a species? *Am Midl Nat* 78:244-247.
- Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5:150-163.
- Moritz C, Cicero C (2004) DNA Barcoding: promise and pitfalls. *PLoS Biol* 2:e354.
- Mota NR, Robe LJ, Valente VLS, Budnik M, Loreto ELS (2008) Phylogeny of the *Drosophila mesophragmatica* group (Diptera, Drosophilidae): an example of Andean evolution. *Zool Sci.* (in press)
- Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Center, Uppsala University.
- O'Grady PM, Kidwell MG (2002) Phylogeny of the subgenus *Sophophora* (Diptera: Drosophilidae) based on combined analysis of nuclear and mitochondrial sequences. *Mol Phylogenet Evol* 22:442-453.
- Pagel M, Meade A (2004) A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Syst Biol* 53:571-581.
- Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Mol Biol Evol* 5:568-583.
- Pérez-Salas S, Richmond RC, Pavlovsky OA, Kastritsis CD, Ehrman L, Dobzhansky T (1970) The Interior semispecies of *Drosophila paulistorum*. *Evolution* 24:519-527.
- Posada C, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14:817-818.
- Powell JR, Sezzi E, Moriyama EN, Gleason JM, Caccone A (2003) Analysis of a shift in codon usage in *Drosophila*. *J Mol Evol* 57:214-225.

- Remsen J, DeSalle R (1998) Character congruence and multiple data partitions and the origin of the Hawaiian Drosophilidae. *Mol Phylogenet Evol* 9:225-235.
- Robe LJ, Valente VLS, Budnik M, Loreto ELS (2005) Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: a nuclear versus mitochondrial gene approach. *Mol Phylogenet Evol* 36:623-640.
- Rodriguez F, Oliver JL, Marin A, Medina JR (1990) The general stochastic model of nucleotide substitution. *J Theor Biol* 142:485-501.
- Rodríguez-Trelles F, Tarrío R, Ayala FJ (2000a) Fluctuating mutation bias and the evolution of base composition in *Drosophila*. *J Mol Evol* 50:1-10.
- Rodríguez-Trelles F, Tarrío R, Ayala FJ (2000b) Evidence for a high ancestral GC content in *Drosophila*. *Mol Biol Evol* 17:1710-1717.
- Rohde C, Garcia ACL, Valiati VH, Valente VLS (2006) Chromosomal evolution of sibling species of the *Drosophila willistoni* group. I. Chromosomal arm IIR (Muller's element B). *Genetica* 126:77-88.
- Rokas A, Williams BL, King N, Carroll SB (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425:798-804.
- Russo CAM, Takezaki N, Nei M (1995) Molecular phylogeny and divergence times of Drosophilid species. *Mol Biol Evol* 12:391-404.
- Sassi AK, Heredia F, Loreto ELS, Valente VLS, Rohde C (2005) Transposable elements P and gypsy in natural populations of *Drosophila willistoni*. *Genet Mol Biol* 28:734-739.
- Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16:1114-1116.



- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Entomol Soc Am* 87: 651-701.
- Sorenson MD, Franzosa EA (2007) TreeRot version 3. Boston University, Boston, MA.
- Spassky B (1957) Morphological differences between sibling species of *Drosophila*. *Univ Texas Publ* 5721:48-61.
- Spassky B, Richmond RC, Perez-Salas S, Pavlovsky O, Mourão CA, Hunter AS, Hoentgsberg H, Dobzhansky T, Ayala FJ (1971) Geography of the sibling species related to *Drosophila willistoni*, and of the semispecies of the *Drosophila paulistorum* complex. *Evolution* 25:129-143.
- Staden R (1996) The Staden sequence analysis package. *Mol Biotechnology* 5:233-241.
- Swofford DL (2003) PAUP: Phylogenetic Analysis using Parsimony (and other methods). Version 4. Sinauer Associates, Massachusetts.
- Tarrío R, Rodriguez-Trelles F, Ayala FJ (2000) Tree rooting with outgroups when they differ in their nucleotide composition from the ingroup: The *Drosophila saltans* and *willistoni* groups, a case study. *Mol Phylogenet Evol* 16:344-349.
- Templeton AR (1983) Convergent evolution and non-parametric inferences from restriction fragment and DNA sequence data. In: Weir B (ed.), *Statistical Analysis of DNA Sequence Data*. Marcel Dekker, New York, pp 151-179.
- Throckmorton LH (1975) The phylogeny, ecology and geography of *Drosophila*. In: King RC (ed.), *Handbook of Genetics*. Plenum, New York, pp 421-469.
- Ting CT, Tsaur SC, Wu CI (2000) The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. *Proc Natl Acad Sci* 97: 5313–5316.

Townsend JL (1954) Cryptic subspeciation in *Drosophila* belonging to the subgenus *Sophophora*. *Amer Nat* 88:339-351.

Xia X, Xie Z (2001) DAMBE: software package for data analysis in molecular biology and evolution. *J Hered* 92:371-3.

### Figure legends

**Figure 1:** Bayesian majority-rule consensus trees obtained with the A) *COI* data set; B) *COII* data set; and C) mitochondrial combined data set, and saturation plots presented by each individual partition. As regards the trees, numbers above the nodes are the Bayesian posterior probabilities for each clade (the combined analysis in C first presents the values regarding MrBayes 3.1.2 output, and then those resulting from BayesPhylogenies), whereas numbers below the nodes are the bootstrap supports for each clade in the ML, NJ and MP trees, respectively. In this case, branch lengths are proportional to the scale given in substitutions per nucleotide. The saturation plots show the transition (ts) and transversion (tv) rates plotted against the JC distances (Jukes and Cantor 1969).

**Figure 2:** Bayesian majority-rule consensus trees obtained with the A) *Adh* data set; B) *Amd* data set; C) *Hb* data set; D) *per* data set; and E) nuclear combined data set, and saturation plots presented by each individual partition. As regards the trees numbers above the nodes are the Bayesian posterior probabilities for each clade (the combined analysis in E first presents the values regarding MrBayes 3.1.2 output, and then those resulting from BayesPhylogenies), whereas numbers below the nodes are the bootstrap supports for each clade in the ML, NJ and MP trees, respectively. In this case, branch lengths are proportional to the scale given in substitutions per nucleotide. The saturation plots show the

transition (ts) and transversion (tv) rates plotted against the JC distances (Jukes and Cantor 1969).

**Figure 3:** *D. willistoni* subgroup total evidence tree. Numbers above the nodes are the Bayesian posterior probabilities for each clade in the MrBayes 3.1.2 and Bayesphylogenies outputs, respectively, whereas numbers below the nodes are the bootstrap supports for each clade in the ML, NJ and MP trees, respectively. Numbers to the right of each node are those referred to in Table 4 Branch lengths are proportional to the scale given in substitutions per nucleotide.

**Figure 4:** Known geographic distribution of the *D. willistoni* subgroup taxa, that is, the six sibling species (A) and the six *D. paulistorum* semispecies (B).

Notes: The distribution of the *D. willistoni* subgroup siblings followed the data presented by Bächli (2008), whereas the distribution of the *D. paulistorum* semispecies is in accordance to Ehrman and Powell (1982).

Table 1

List of *Drosophila* species and strains used in this study, with their source and GenBank accession numbers

Group	Subgroup	Species	Semi or Subspecies	Initials	Source	<i>Adh</i>	<i>Amd</i>	<i>COI</i>	<i>COII</i>	<i>Hb</i>	<i>Per</i>		
<i>willistoni</i>	<i>willistoni</i>	<i>D. paulistorum</i>	Amazonian	AM	Belém, Pa, Brazil	<u>EU532122</u>		U51594	<u>EU532085</u>	<u>EU532110</u>	U51080		
				Para	Belém, Pa, Brazil	<u>EU532128</u>		-	<u>EU532091</u>	<u>EU532105</u>	-		
			Andean-Brazilian	AB	Mesitas, Colombia	<u>EU532121</u>		U51595	<u>EU532084</u>	<u>EU532107</u>	U51081		
				JAI	São Paulo, SP, Brazil	<u>EU532125</u>		-	<u>EU532088</u>	<u>EU532108</u>	-		
				MLC	Morro da Lagoa da Conceição, SC, Brazil	<u>EU532126</u>		-	<u>EU532089</u>	<u>EU532111</u>	-		
				RP	Ribeirão Preto, SP, Brazil	<u>EU532129</u>		-	<u>EU532092</u>	<u>EU532109</u>	-		
			Centroamerican	CA	Lancetilla, Honduras	<u>EU532123</u>		U51596	<u>EU532086</u>	<u>EU532102</u>	U51082		
			Interior	Int	Llanos, Colombia	<u>EU532124</u>		U51597	<u>EU532087</u>	<u>EU532103</u>	U51083		
			Orinocan	Ori	Georgetown, Guyana	<u>EU532127</u>		U51598	<u>EU532090</u>	<u>EU532106</u>	U51084		
			Transitional	Trans	Santa Marta, Colombia	<u>EU532130</u>		U51599	<u>EU532093</u>	<u>EU532104</u>	U51085		
			<i>D. pavlovskiana</i>	-	-	-	-	AF264079	-	U51600	AF474101	-	U51086
			<i>D. willistoni</i>	<i>willistoni</i>	1	Belém, Pa, Brazil	<u>EU532119</u>		-	<u>EU532097</u>	<u>EU532112</u>	U51057	
					2	-	L08648		AF293730	U51590	-	*	U51064
		<i>D. equinoxialis</i>	<i>equinoxialis</i>	1	Tefê, AM, Brazil	<u>EU532118</u>		-	<u>EU532080</u>	<u>EU532101</u>	U51073		
				2	-	U95269		-	U51592	-	-	U51076	
			Carmody strain	-	-	-		-	U51593	-	-	U51079	
		<i>D. insularis</i>	-	1	St Kitts, Lesser Antilles	<u>EU532120</u>		-	<u>EU532082</u>	<u>EU532098</u>	-		
				2	-	U95273		-	U51602	-	-	U51088	
		<i>D. tropicalis</i>	<i>tropicalis</i>	1	San Salvador, El Salvador	<u>EU532131</u>		-	<u>EU532095</u>	<u>EU532100</u>	-		
				2	-	U95274		-	U51601	-	-	U51087	
			<i>cubana</i>	cub	Trinidad, Venezuela	<u>EU532132</u>		-	<u>EU532096</u>	<u>EU532099</u>	-		
		<i>bocainensis</i>	<i>D. capricorni</i>	1	Florianópolis, SC, Brazil	-		-	<u>EU532079</u>	<u>EU532113</u>	-		
				2	-	AY335196		-	AY335202	-	-	U51092	
<i>D. fumipennis</i>	-			1	Florianópolis, SC, Brazil	<u>EU532133</u>		-	<u>EU532081</u>	<u>EU532115</u>	-		
	-			2	-	AF264075		-	-	-	-		
<i>D. nebulosa</i>	-			1	Porto Alegre, RS, Brazil	-		-	<u>EU532083</u>	<u>EU532116</u>	-		
	-			2	-	U95275		AF293717	U51603	-	-	U51090	
<i>D. sucinea</i>	-			1	México City, Mexico	-		-	<u>EU532094</u>	<u>EU532114</u>	-		
<i>saltans</i>	<i>D. sturtevantii</i>	-	-	2	-	AY335197	-	AY335203	-	-	U51091		
		-	-	-	AY335201		-	AY335205	AF045084	<u>EU532117</u>	19		

Newly obtained sequences are underlined. (2) Population origin refers only to samples which were the source of new sequences. (3) The *Amd* sequences were not yet submitted.

\* The *D. willistoni\_2* sequence was obtained through Blast searches in the *D. willistoni* genome (<http://insects.eugenes.org/species/blast/>).

Table 2  
Properties of the molecular data sets

	<i>Adh</i>	<i>Amd</i>	<i>Hb</i>	<i>Per</i>	Combined nuclear	<i>COI</i>	<i>COII</i>	Combined mitochondrial	Total dataset
N <sup>o</sup> of positions in the alignment	675	882	753	1,249	3,445	495	672	1,167	4,612
N <sup>o</sup> of OTU's	26	19	21	18	12	16	21	15	12
N <sup>o</sup> of variable sites	196	231	211	389	801	83	108	180	952
N <sup>o</sup> of parsimoniously informative sites	131	109	111	224	296	50	68	111	363
Total Nucleotide frequencies <sup>a</sup>									
- A	25.2	25.3	26.1	24.7	25.1	29.3	32.4	31.1	26.7
- C	22.8	20.6	30.4	28.1	25.8	13.1	13.3	13.2	22.6
- G	21.6	27.8	23.2	27.0	25.6	15.9	13.1	14.3	22.7
- T	30.5	26.2	20.3	20.1	23.5	41.7	41.1	41.4	28.1
Transition/Transversion ratio <sup>a</sup>	1.3	1.6	1.6	1.4	1.5	1.1	1.3	1.1	1.5
X <sup>2</sup> test	18.4504 (df=75), ns	16.6009 (df=54), ns	30.9267 (df=60), ns	42.6879 (df=51), ns	33.1674 (df=33), ns	3.3131 (df=45), ns	3.3322 (df=60), ns	3.1353 (df=42), ns	29.0558 (df=33), ns
MrModelTest best fit model	HKY + I + G	GTR + I + G	GTR + G	HKY + I	GTR + I + G	GTR + G	GTR + I	GTR + I + G	GTR + I + G
ModelTest best fit model	TrN + I + G	TrN + I + G	GTR + G	TVM + I	GTR + I + G	GTR + G	GTR + I	GTR + I + G	GTR + I + G
Proportion of invariable sites <sup>b</sup>	0.4077	0.3817	-	0.5602	0.3497	-	0.7764	0.7504	0.4489
Gamma distribution shape parameter <sup>b</sup>	0.6862	0.7806	0.3540	-	0.7215	0.0804	-	1.9212	0.7487
Rate matrix <sup>b</sup>									
- rAC	1.0000	1.0000	1.6022	1.0042	0.7850	14.1908	15.0255	14.5795	1.0970
- rAG	2.0801	3.2972	2.8549	3.5802	2.3545	14.2812	52.3595	33.7375	3.2626
- rAT	1.0000	1.0000	1.3107	1.2932	1.0259	23.4564	33.8078	26.9169	1.5397
- rCG	1.0000	1.0000	0.7294	0.6815	0.6111	5.6948	0.0000	0.0000	1.0388
- rCT	3.4028	5.4891	6.0628	3.5802	3.7855	165.2966	199.7041	198.8293	5.4761

<sup>a</sup> Empirical values as obtained by Mega 3.1 program. <sup>b</sup>Maximum likelihood settings from the best fit model selected by AIC in ModelTest 3.7.

Abbreviations: ns, not significant; HKY, Hasegawa-Kishino-Yano model (Hasegawa et al. 1985); TrN, Tamura-Nei model (Tamura & Nei, 1993); GTR, General Time Reversible (Rodriguez *et al.*, 1990); TVM, transversional model; I, proportion of invariable sites; G, shape parameter of gamma distribution.

Table 3

p values resulting from pairwise PHT performed between each of the six data partitions included in the total evidence analyses

	<i>Adh</i>	<i>Amd</i>	<i>Hb</i>	<i>Per</i>	<i>COI</i>	<i>COII</i>
<i>Adh</i>	-	0.385 <sup>Ns</sup>	0.933 <sup>Ns</sup>	0.134 <sup>Ns</sup>	0.002**	0.001**
<i>Amd</i>		-	1.000 <sup>Ns</sup>	0.004**	0.006**	0.001**
<i>Hb</i>			-	0.061 <sup>Ns</sup>	0.001**	0.002**
<i>Per</i>				-	0.001**	0.001**
<i>COI</i>					-	0.001**
<i>COII</i>						-

<sup>Ns</sup> Not significant; \* Significant at the 0.05 level; \*\* Significant at the 0.01 level.

Table 4

Average Partitioned Bremer Support (PBS) scores for each of the six gene partitions across the *willistoni* subgroup TET (Fig. 3)

Clade	<i>Adh</i>	<i>Amd</i>	<i>Hb</i>	<i>Per</i>	<i>COI</i>	<i>COII</i>
1	1.0	1.0	0.0	8.0	-1.0	-1.0
2	-0.5	1.0	0.0	0.0	1.0	-0.5
3	0.0	-2.0	0.0	-1.0	-2.0	7.0
4	-1.0	1.0	0.0	0.0	0.0	2.0
5	6.0	1.5	2.0	5.5	0.5	-3.5
6	1.75	3.0	7.75	19.0	2.0	3.5
7	1.0	-1.0	-1.0	6.0	3.0	-2.0
8	4.0	2.0	9.0	15.0	-2.0	-6.0
9	15.0	14.0	25.0	42.0	9.0	20.0
$\Sigma$ PBS	27.25	20.5	42.75	94.5	10.5	19.5
Min steps	290	280	256	490	94	125
* $\Sigma$ PBS/ Min steps	0.094	0.073	0.167	0.193	0.112	0.156

\* PBS values summed across the tree and standardized by the minimum number of steps for each partition in their corresponding Bayesian Tree (Figs. 1 and 2).

Table 5

Average genetic distance between taxa at different taxonomic levels within the *willistoni* group, for each of the six different data partition included in the total evidence analyses

Taxonomic level	<i>Adh</i>	<i>Amd</i>	<i>Hb</i>	<i>Per</i>	<i>COI</i>	<i>COII</i>
Between populations	0.016	0.014	0.014	0.002	-	-
Between subspecies	0.018	0.018	0.004	0.014	-	0.002
Within semispecies	0.010	0.0095	0.000	-	-	0.0015
Between semispecies	0.010	0.0125	0.005	0.0084	0.024	0.014
Between sibling species	0.044	0.0533	0.045	0.051	0.028	0.026
Between non-sibling	0.117	0.120	0.077	0.090	0.049	0.057
Between sibling and non-sibling species	0.138	0.120	0.114	0.150	0.067	0.068



Supplementary Material - Table 1S  
 Summary of the Templeton (T) and Shimodaira-Hasegawa (SH) Tests results

Partitions	Topologies																		
	COI		COII		COI+COII		Adh		Amd		Hb		Per		Nuclear		TET		
	T	SH	T	SH	T	SH	T	SH	T	SH	T	SH	T	SH	T	SH	T	SH	
COI	B	B	X	X	X	X	X	X	X	✓	X	X	X	X	X	X	X	X	
COII	X	X	B	B	✓	✓	X	X	X	X	X	X	X	X	X	X	X	✓	
Combined mitochondrial	✓	X	✓	✓	B	B	X	X	X	X	X	X	X	X	X	X	X	✓	
Adh	X	X	X	X	X	X	B	B	X	X	X	X	✓	✓	✓	✓	✓	✓	
Amd	X	X	X	X	X	X	X	X	✓	✓	✓	X	X	✓	B	B	X	✓	
Hb	X	X	X	X	X	✓	✓	✓	✓	✓	B	B	✓	✓	✓	✓	✓	✓	
Per	X	X	X	X	X	X	✓	✓	X	X	X	X	B	B	✓	✓	✓	✓	
Combined Nuclear	X	X	X	X	X	X	X	✓	X	X	X	X	✓	✓	B	B	X	✓	
All data	X	X	X	X	X	X	X	✓	X	X	X	X	X	✓	✓	✓	✓	B	B

Notes: The letter B denotes the best tree found under the maximum parsimony (for the Templeton Test) or under the maximum likelihood (for the Shimodaira-Hasegawa Test) criterion for the different partitions. X indicates that the topology was rejected, whereas ✓ indicates that the topology was accepted.

Figure 1

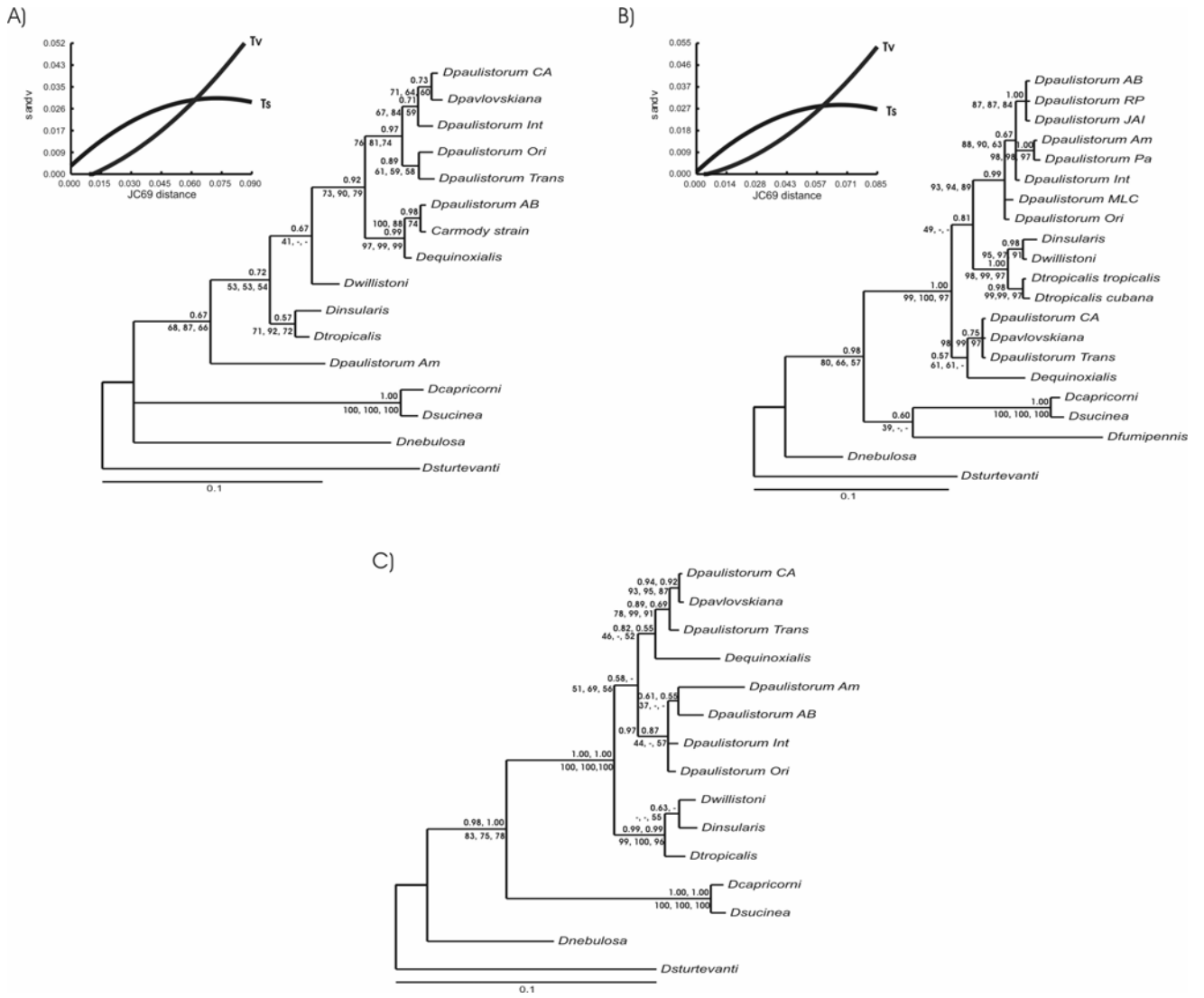


Figure 2

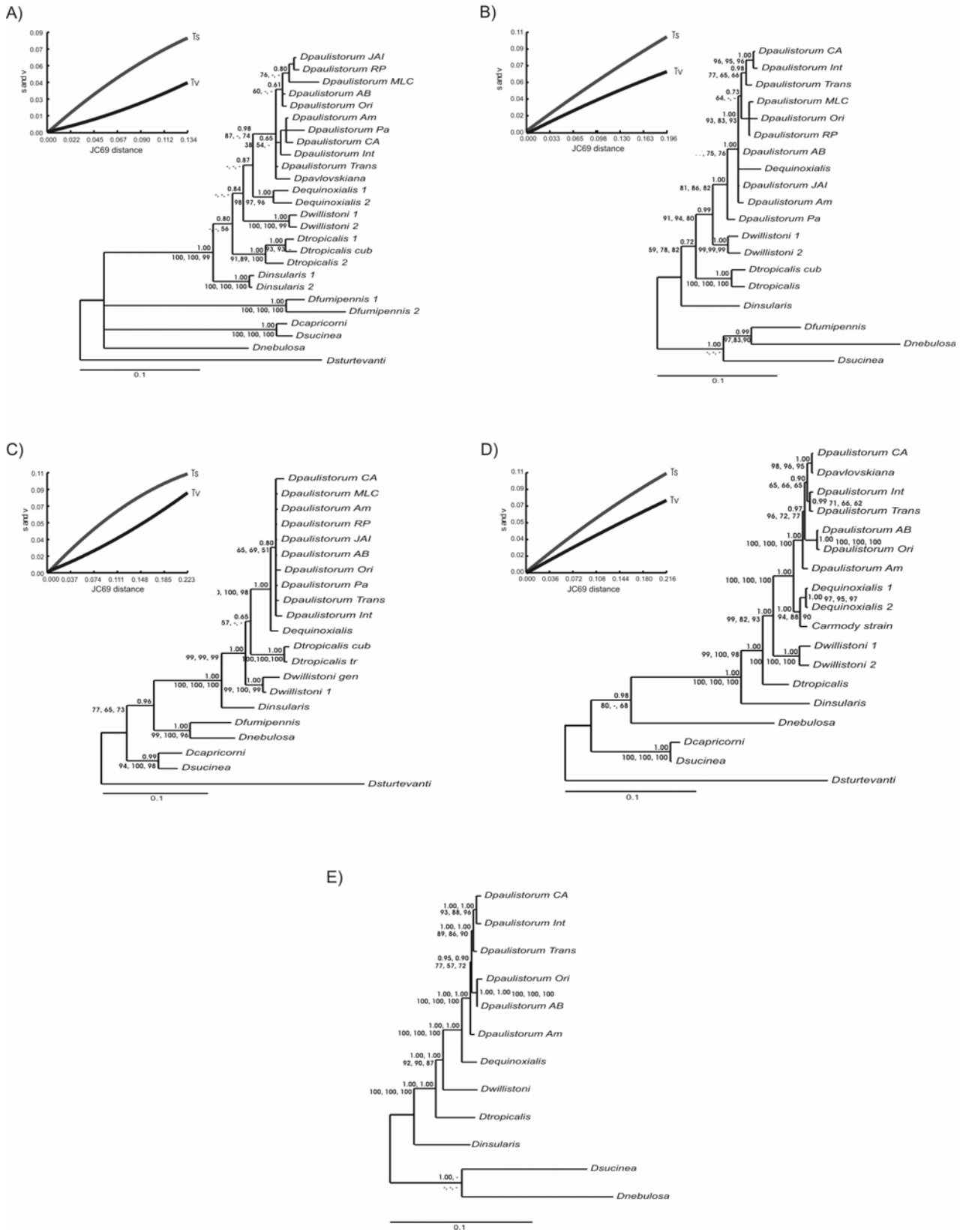


Figure 3

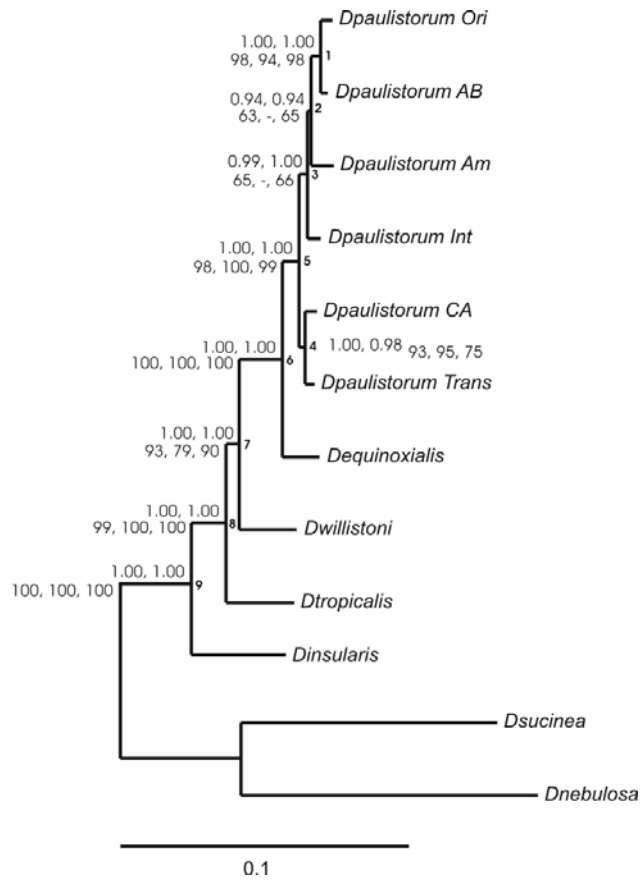


Figure 4A

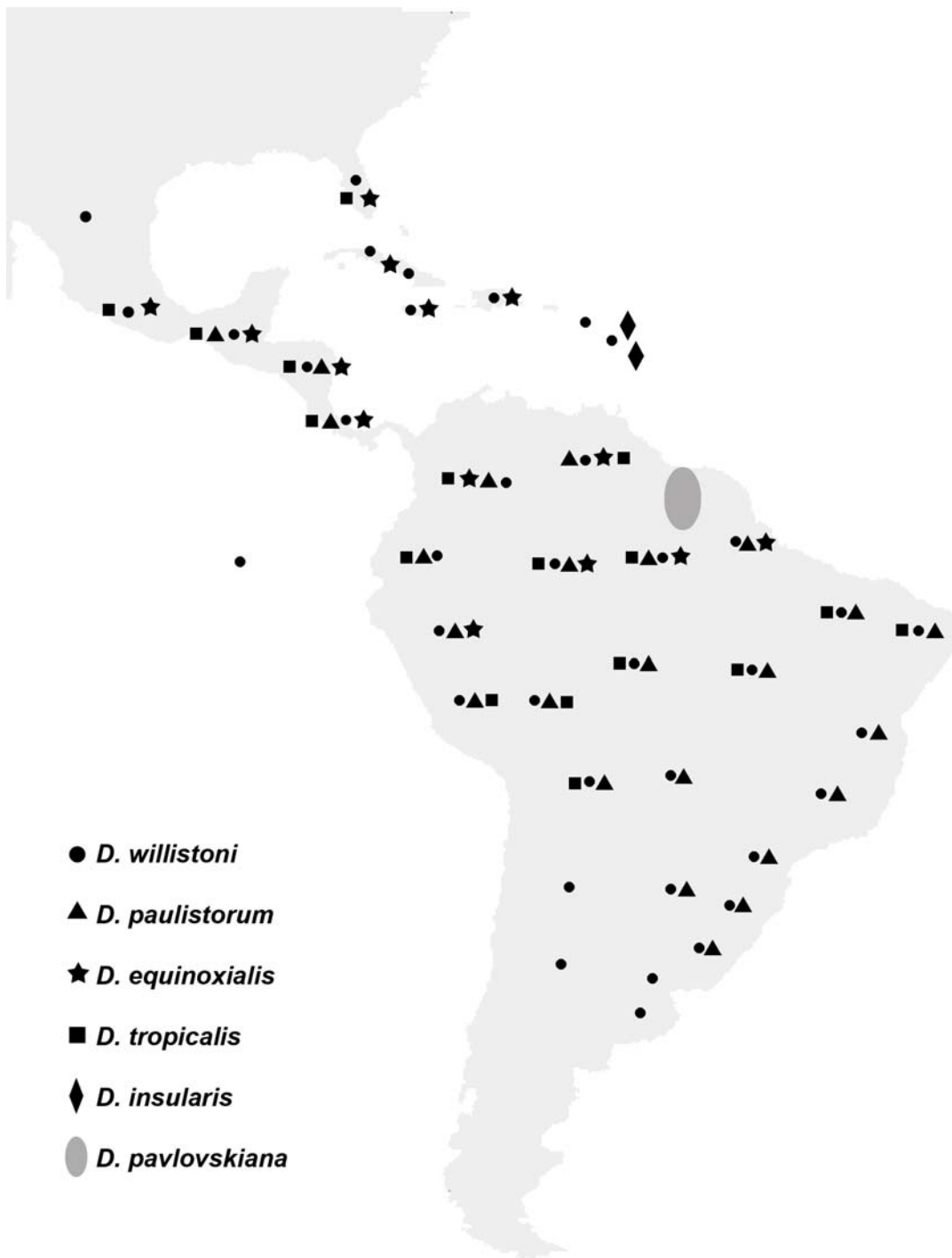


Figure 4B



## CAPÍTULO VI

---

### CONCLUSÕES GERAIS E PERSPECTIVAS

Com mais de 3.900 espécies distribuídas em 76 gêneros (Bächli, 2008), a família Drosophilidae representa um dos mais diversos e amplamente distribuídos grupos de insetos pertencentes à ordem Diptera. Ao longo de sua distribuição, os drosofilídeos desempenham um papel fundamental nas cadeias alimentares saprofíticas dos mais diversos ecossistemas. Além disso, diferentes espécies pertencentes a esta família tornaram-se, ao longo dos últimos 100 anos, organismos modelo ideais nas mais diversas áreas da pesquisa biológica, principalmente no âmbito genético e evolutivo.

A família Drosophilidae subdivide-se internamente em duas subfamílias: Steganinae, que apresenta 28 gêneros; e Drosophilinae, que perfaz outros 47 gêneros (Toda, 2007; Bächli, 2008). Grimaldi (1990), com base na análise de caracteres morfológicos, confirmou a monofilia da família Drosophilidae e das suas subfamílias. Posteriormente, Remsen & O'Grady (2002) obtiveram resultados semelhantes a partir da análise combinada de caracteres morfológicos e moleculares.

O gênero *Drosophila* destaca-se como membro mais notável da família Drosophilidae, uma vez que, sozinho, ele abrange cerca de um terço de suas espécies, desempenhando ainda um papel fundamental no desenvolvimento científico biológico. Entretanto, o próprio posicionamento do gênero *Drosophila* dentro de Drosophilidae é questionado, sendo que diferentes trabalhos têm evidenciado sua composição parafilética com relação a diversos outros gêneros (Throckmorton, 1975; Thomas & Hunt, 1991 e

1993; DeSalle, 1992a e b; Pélandakis & Solignac, 1993; Russo *et al.*, 1995; Kwiatowski *et al.*, 1997; Remsen & DeSalle, 1998; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999a e 2001; Katoh *et al.*, 2000; Remsen & O'Grady, 2002; Zilversmit *et al.*, 2002; Yotoko *et al.*, 2003; DaLage *et al.*, 2007).

A situação agrava-se quando parte dos oito subgêneros pertencentes ao gênero *Drosophila* são considerados. Quando incluídos em análises mais amplas, estes subgêneros revelam-se, muitas vezes, mais distantes entre si do que aos demais gêneros da família (Russo *et al.*, 1995; Katoh *et al.*, 2000; Kwiatowski & Ayala, 1999; Tatarenkov *et al.*, 1999 e 2001; Remsen & O'Grady, 2002; DaLage *et al.*, 2007). Apesar disso, nenhuma proposição efetiva de revisão taxonômica foi até o momento apresentada para o gênero *Drosophila*.

Existem, ainda, muitos grupos de espécies pertencentes ao gênero *Drosophila* que foram subrepresentados em análises filogenéticas prévias. O subgênero *Drosophila*, por exemplo, apesar de constituir o maior representante do gênero *Drosophila* (com 721 espécies – Bächli, 2008) e ocupar o *status* de maior representante da família Drosophilidae na região Neotropical (Val *et al.*, 1981), não apresenta, até o momento suas relações filogenéticas esclarecidas. Apenas Throckmorton (1975), e mais recentemente Remsen & O'Grady (2002) e DaLage *et al.* (2007) buscaram avaliar relações de amplo nível dentro deste subgênero. Além disso, diversos autores levantaram a possibilidade de que o subgênero *Drosophila* (ou parte dele) represente um grupo de evolução complexa, que se originou e se diversificou através de sucessivos eventos de radiação (Throckmorton, 1962 e 1975; Yotoko *et al.*, 2003).



Embora o subgênero *Sophophora* tenha sido mais amplamente amostrado do que os demais, ele também apresenta uma série de questões em aberto. O grupo *willistoni* em particular representa um complexo taxonômico cuja evolução ainda não se encontra totalmente esclarecida, a despeito de diversas tentativas prévias (Spassky *et al.*, 1971; Ayala *et al.*, 1974; Gleason *et al.*, 1998; Tarrío *et al.*, 2000).

Todas estas inconsistências, incongruências e irresoluções contrapõem-se ao fato de que a obtenção de uma hipótese filogenética robusta para *Drosophila* apresenta uma importância fundamental (Powel & DeSalle, 1995), não só para o entendimento da evolução de características já conhecidas para o grupo, como também para a interpretação e para o delineamento de estudos futuros. Considerando-se este fato, a presente tese foi desenvolvida com o objetivo principal de auxiliar na resolução de diversas questões relacionadas à história evolutiva do gênero *Drosophila*, principalmente no que diz respeito a alguns de seus táxons menos compreendidos filogeneticamente, mas que se destacam pela diversidade apresentada no Novo Mundo: o subgênero *Drosophila* e o grupo *willistoni* do subgênero *Sophophora*. Neste sentido, buscou-se não só ao estabelecimento das relações filogenéticas entre as espécies, mas também à proposição dos períodos e cenários que possivelmente atuaram na geração da atual diversidade.

Para tanto, o **Capítulo II** desta tese apresenta os resultados obtidos mediante utilização dos genes *Amd* e *COII* na busca pelo estabelecimento das relações filogenéticas entre espécies Neotropicais do subgênero *Drosophila*. De modo geral, as árvores obtidas com *Amd* e as árvores de evidências totais recapitulam Throckmorton (1975) em diversos aspectos, sendo que ambas reconstruções apresentam definições monofiléticas para as radiações *immigrans-Hirtodrosophila* (dada a exclusão do subgênero *Dorsilopha* e a

inclusão de *D. funebris* e do gênero *Liodrosophila*) e *virilis-repleta*, que representam a grande subdivisão interna do subgênero *Drosophila*. Além disso, e também em concordância com Throckmorton (1975) estas duas abordagens apresentam o subgênero *Sophophora* como uma ramificação basal dentre os Drosophilinae amostrados, de modo que o subgênero *Drosophila* adquira um caráter parafilético. *COII*, porém, revelou alguns níveis de incongruência com relação não só a *Amd*, como também com relação a Throckmorton (1975). A alta saturação apresentada por este marcador parece ter prejudicado sua performance, principalmente no que diz respeito às relações de nível mais inclusivo. Neste sentido, ainda que a radiação *immigrans-Hirtodrosophila* tenha sido recuperada nas análises realizadas a partir da matriz de *COII* (dada a inclusão de *D. funebris*, *D. nanoptera* e do gênero *Liodrosophila*), a radiação *virilis-repleta* revelou-se inteiramente parafilética para as seqüências de *COII*, e o subgênero *Sophophora* misturou-se com alguns de seus membros. Ambos os marcadores, entretanto, concordaram no estabelecimento de um grupo *tripunctata* parafilético, embora *Amd* tenha suportado a monofilia da radiação *tripunctata*, diferentemente de *COII*.

No **Capítulo III**, por outro lado, buscou-se resolver algumas destas incongruências ao ampliar a amostragem de espécies pertencentes ao subgênero *Drosophila*, bem como o número de nucleotídeos analisados. Neste sentido, além de novas seqüências para o gene *Amd*, seu parálogo *Ddc* também foi seqüenciado. Adicionalmente, seqüências do gene para fator de transcrição *Hb* também foram incluídas nas análises visando um refinamento das relações filogenéticas dentro da radiação *immigrans-tripunctata*. Como no **Capítulo II** *COII* demonstrou maiores níveis de congruência no clado referente aos membros desta radiação, suas seqüências também foram incluídas nesta parte da análise. De modo geral, os

resultados obtidos mediante análise individual e combinada dos genes *Amd* e *Ddc* favorecem a hipótese de que o subgênero *Sophophora* compreende uma ramificação basal em Drosophilinae, enquanto que um subgênero *Drosophila* parafilético subdivide-se em três maiores clados: a radiação *immigrans-tripunctata*, a radiação *virilis-repleta*, e a radiação dos drosofilídeos hawaianos. É possível que a inclusão de novas espécies nestas análises revele, ainda, a presença de outros grandes clados, como por exemplo, aquele que compõe *Zaprionus* e *Liodrosophila*.

Ainda no **Capítulo III**, estimativas bayesianas de datação dos eventos de divergência dentro destas três radiações demonstram que pelo menos suas ramificações basais se deram ao longo do Oligoceno e no início do Mioceno, épocas de relativa estabilidade climática. Desta forma, é bastante provável que a invasão do Novo Mundo pelo ancestral destas linhagens tenha fornecido oportunidades de diversificação e especiação a estas espécies, ao apresentar uma grande variedade de nichos alternativos disponíveis. Além disso, a busca por uma filogenia robusta para a radiação *immigrans-tripunctata* foi parcialmente alcançada pela análise combinada dos genes *Amd*, *Ddc*, *Hb* e *COII*. Diferentemente das análises moleculares anteriores (**Capítulo II**, Yotoko *et al.*, 2003), as reconstruções filogenéticas obtidas com estes dados concatenados forneceram altos níveis de suporte às relações internas à radiação *immigrans-tripunctata*, que se apresentou subdividida em três linhagens principais: as linhagens *tripunctata*, *cardini* e *mediostriata*. Entretanto, a sobreposição das estimativas de tempo de diversificação destas três linhagens em cerca de 18,5 milhões de anos atrás, bem como o curto tamanho dos ramos recuperados para alguns de seus ramos internos suportam o cenário em que radiações rápidas e sucessivas levaram à origem e diversificação da radiação *immigrans-tripunctata*.

Desta forma, nos **Capítulos II e III** o objetivo de obter uma descrição confiável da história evolutiva do subgênero *Drosophila* foi alcançado com um sucesso até maior do que o esperado, dadas as estimativas prévias de que este grupo de drosofilídeos representa um cenário complexo (Powell & DeSalle, 1995), em que rápidas e sucessivas radiações (Throckmorton, 1962 e 1975) atuaram na geração da diversidade atual. Na verdade, embora níveis confiáveis de resolução tenham sido alcançados, principalmente no que diz respeito ao **Capítulo III**, ainda é bastante provável que o acúmulo de novos marcadores forneça hipóteses alternativas, uma vez que eventos de *incomplete lineage sorting* e introgressão parecem ser possíveis complicadores na obtenção de cenários compartilhados mesmo entre marcadores não saturados.

De modo geral, os resultados apresentados nos **Capítulos II e III** embasam-se amplamente nas relações inferidas a partir de *Amd* e *Ddc*, que, como genes ligados, forneceram hipóteses congruentes. Mais do que isto, entretanto, *Amd* e *Ddc* se mostraram amplamente adequados para análises filogenéticas envolvendo este nível de divergência em *Drosophila*. Embora alguns níveis de saturação tenham sido evidenciados em ambos os casos, a adoção de modelos e métodos mais complexos (Nei & Kumar, 2000) parece ter recuperado com êxito a maior parte do sinal apresentado por estes genes. Desta forma, como muitas relações obtidas apresentam suporte de outros marcadores, e, mais do que isto, tem suporte morfológico (Throckmorton, 1962 e 1975), acredita-se que níveis adicionais de confiabilidade podem ser acrescentados aos dados apresentados. E por apresentar sinais filogenéticos tão fortes mesmo em cenários evolutivos complexos, como aqueles que possivelmente deram origem ao subgênero *Drosophila*, se levanta a possibilidade de que estes genes estejam atuando, em alguns casos, como genes de especiação. Talvez,

mudanças particulares na esclerotização e/ou pigmentação da cutícula (ver Anexo I) tenham sido fatores importantes na obtenção de isolamento reprodutivo pré-zigótico entre algumas populações, uma vez que estas propriedades podem alterar padrões de reconhecimento do parceiro sexual, como também modificar a morfogênese da genitália, que, ao adquirir novas formas, pode levar a eventos de isolamento reprodutivo. Entretanto, somente a continuidade dos estudos filogenéticos relacionados ao subgênero *Drosophila*, bem como a continuidade dos estudos relacionados à evolução molecular dos genes *Amd* e *Ddc* poderá testar com maior eficiência os pontos fortes e fracos de todas as hipóteses aqui discutidas.

Seja como for, já é evidente, pelos dados atuais, que a taxonomia de Drosophilidae encontra-se um pouco defasada. O problema é que uma mudança nesta área acarretaria problemas futuros na comunicação e mesmo no entendimento de milhões de manuscritos já publicados com a taxonomia atual. Mais do que isto, pelos dados atuais, é provável que uma das maiores mudanças taxonômicas se desse no próprio subgênero *Sophophora*, que, pelo atual *status* filogenético, deveria ser alocado a um gênero à parte. Isto faria com que, por exemplo, a familiar *Drosophila melanogaster* fosse renomeada à desconhecida *Sophophora melanogaster*, o que certamente mexeria com os alicerces científicos. Como a proposição de tamanha mudança não é tarefa fácil, dados são acumulados sem efeitos na classificação, que permanece estável a despeito dos avanços no conhecimento. Mais recentemente, movimentos estão sendo feitos no sentido de mudar a espécie tipo do gênero de *D. funebris* para *D. melanogaster* (Linde *et al.*, 2007), de forma que, se a atualização taxonômica ocorrer, não teria maiores efeitos na nomenclatura de sua espécie mais popular.

Resta saber o que aconteceria neste caso com a nomenclatura subgenérica. Aparentemente, esta nova medida parece complicar ainda mais o cenário.

Enquanto questões de ordem maior perduram no que diz respeito à filogenia e à taxonomia de grandes grupos, como por exemplo, o gênero e o subgênero *Drosophila*, resolver questões de nível menos inclusivo não deixam de ter sua importância. Neste sentido, o **Capítulo IV** foi delineado com o objetivo de resolver algumas questões acerca das relações filogenéticas e da própria taxonomia de um grupo Neotropical, fundamentalmente Andino, pertencente à radiação *virilis-repleta* do subgênero *Drosophila*: o grupo *mesophragmatica*. Embora houvessem ainda algumas inconsistências acerca das relações internas deste interessante grupo (Nacur, 1958; Brncic *et al.*, 1971; Nair *et al.*, 1971), as análises realizadas a partir de mais de 4.300 pb, compreendendo cinco diferentes marcadores (*Adh*, *Amd*, *Ddc*, *Hb* e *COII*) analisados individual e simultaneamente, apresentaram suporte para sua subdivisão em três linhagens: o primeiro, compondo a ramificação basal, e formado por *D. viracochi*; o segundo constituído pelas espécies crípticas *D. gaucha* e *D. pavani*; e o terceiro misturando representantes das espécies *D. gasici* com *D. brncici*, apresentando *D. mesophragmatica* como divergência basal. Mais uma vez, os genes *Amd* e *Ddc*, foram inteiramente congruentes com este cenário, também suportado pela árvore de evidências totais, enquanto *Adh*, *Hb* e *COII* apresentaram pequenas e isoladas incongruências. A partir destes resultados, se propôs uma mudança na taxonomia do grupo, de forma que esta passasse a refletir esta subdivisão em três, e não dois subgrupos. Também foi possível inferir a ação dos fenômenos de *incomplete lineage sorting* e introgressão neste pequeno grupo de espécies, dado que principalmente no que se refere ao subgrupo composto por *D. gasici*, *D. brncici* e *D. mesophragmatica* relações

incongruentes e espécies parafiléticas foram definidas. Desta forma, os complicadores representados por estes fatores parecem se fazer presentes nos mais diversos níveis de divergência dentro do subgênero *Drosophila*. Além disso, pela datação dos eventos de divergência e especiação dentro do grupo, foi possível evidenciar a ação que os fenômenos andinos de glaciação, e a conseqüente formação de refúgios, tiveram na evolução do grupo, principalmente durante o Plioceno e o Pleistoceno. Desta forma, os eventos de especiação dentro do grupo *mesophragmatica* parecem não fugir à regra geral de que o padrão temporal da evolução de Drosophilidae encontra-se intimamente correlacionado à fragmentação de habitat decorrente do resfriamento paleoclimático do Cenozóico (Tamura *et al.*, 2004).

Enquanto todos os capítulos anteriores buscaram auxiliar no estabelecimento da história evolutiva entre grupos e espécies pertencentes ao subgênero *Drosophila*, o **Capítulo V** foi designado no sentido de buscar níveis adicionais de resolução a questões referentes à evolução de um grupo de espécies Neotropicais pertencentes ao subgênero *Sophophora* que também apresenta uma série de lacunas no conhecimento de sua história evolutiva: o grupo *willistoni*. Este grupo destaca-se por sua complexidade taxonômica e evolutiva, sendo composto por táxons em variados níveis de diferenciação: espécies morfologicamente distintas, espécies crípticas, semi-espécies, subespécies e populações. Neste caso, embora amplos níveis de incongruência tenham sido detectados majoritariamente entre as partições nucleares (*Adh*, *Amd*, *Hb* e *Per*) e as mitocondriais (*COI* e *COII*), bem como no que se refere às partições mitocondriais comparadas entre si, nenhuma das nove topologias construídas foi idêntica a qualquer uma das outras. As incongruências e inconsistências dos genes mitocondriais podem, sem qualquer dúvida, ser

atribuídas à sua evidente saturação. Desta forma, mais uma vez, marcadores mitocondriais tiveram uma performance filogenética pobre, principalmente se comparados aos marcadores nucleares, o que também pôde ser constatado nos **Capítulos II, III e IV**. A composição nucleotídica enviesada, a sua matriz de transformação fortemente assimétrica e a heterogeneidade de taxas evolutivas entre sítios apresentada pelos genes mitocondriais, parecem constituir limitadores consideráveis no que diz respeito ao seu desempenho na busca por reconstruções filogenéticas no gênero *Drosophila*. Mais do que isso, estes fatores parecem colocar em cheque a própria utilidade do gene *COI* no processo de *DNA Barcoding* (Hebert *et al.*, 2003) dentro do gênero. Como fortes sinais de saturação não foram evidenciados para os genes nucleares, que também apresentaram algumas incongruências, é possível que processos complexos atribuídos à própria evolução do grupo *willistoni* (*incomplete lineage sorting* e introgressão) também atuem, mais uma vez, como fatores complicadores na obtenção de uma história filogenética robusta e única.

No caso do **Capítulo V**, entretanto, diferentemente do que fora apresentado para as espécies pertencentes ao subgênero *Drosophila*, *Amd* não se destacou como gene mais informativo, principalmente no que se refere às OTU's mais próximas (semiespécies de *D. paulistorum* e *D. equinoxialis*), de forma que seu limiar de utilidade filogenética parece ter sido ultrapassado. Mesmo assim, *Amd* contribuiu para o estabelecimento das relações filogenéticas do subgrupo *willistoni* em níveis mais inclusivos e, juntamente com os outros genes nucleares - *Adh*, *Per* e *Hb*, parece ter fornecido sinais suficientes para a reconstrução de uma filogenia relativamente confiável, dada sua concordância com uma série de resultados prévios (Gleason *et al.*, 1998; Tarrío *et al.*, 2000; Rohde *et al.*, 2006). Com base na filogenia de evidências totais foi possível, inclusive, fazer inferências acerca do cenário



evolutivo que levou à origem e diversificação do subgrupo *willistoni*, bem como fazer proposições de ordem taxonômica. De modo geral, ancestrais de subgrupo *willistoni* e da própria super-espécie *D. paulistorum*, parecem ter iniciado sua diversificação e dispersão a partir de centros localizados ao norte da América do Sul e na América Central. Mais uma vez, refúgios Pleistocênicos parecem ter sido importantes na divergência das semi-espécies, cuja monofilia foi, inclusive, questionada. Levantou-se, neste sentido, a possibilidade de que o simples compartilhamento de linhagens de endossimbiontes semelhantes por populações diferentes possa estar levando à diferenciação destas populações com relação às demais. Mais uma vez, somente a continuidade dos estudos deverá esclarecer de maneira mais satisfatória o cenário subjacente.

De modo geral, tanto nos **Capítulos II e III**, quanto nos **Capítulos IV e V**, as árvores reconstruídas a partir dos caracteres concatenados foram consideradas as melhores hipóteses evolutivas para as OTU's em questão, mesmo quando sinais significativos de incongruência foram detectados entre as partições individuais. Embora esta seja uma abordagem controversa (Kluge, 1989; Queiroz *et al.*, 1995; Huelsenbeck *et al.*, 1996), considera-se realísticas as defesas apresentadas por Baker & DeSalle (1997), Remsen & DeSalle (1998) e Remsen & O'Grady (2002), por exemplo, que demonstram na prática que a combinação de diferentes partições, ainda que heterogêneas, amplifica o sinal filogenético presente nas mesmas em detrimento do ruído. Isto por que, conforme esperado, quando o ruído é devido a um excesso de homoplasias e substituições múltiplas, o mesmo se dispersa em diversas direções, enquanto o sinal filogenético tende a ser sempre o mesmo. Por outro lado, quando o "ruído" deve-se ao fenômeno de *incomplete lineage sorting*, concatenar fornece evidências de qual o sinal mais freqüente, o que pode, em boa parte dos casos,

fornecer a história “verdadeira”. Neste caso, parece ser particularmente adequado o uso de estimativas bayesianas, que além de permitirem a atribuição de modelos de substituição específicos a cada partição (Programa MrBayes – Huelsenbeck & Ronquist, 2001), também permitem a subdivisão *a posteriori* da matriz concatenada (Programa BayesPhylogenies – Pagel & Meade, 2004).

Enfim, embora a presente tese tenha contribuído para o estabelecimento das relações filogenéticas e mesmo da história evolutiva de algumas dentre as 3.900 que atualmente compõem a família Drosophilidae, é importante mencionar que, dada a impossibilidade de se conhecer a filogenia verdadeira das espécies, apenas o acúmulo de novos estudos poderá nos aproximar de um cenário mais verossímil e confiável, que pode ou não ser semelhante ao aqui apresentado. Entretanto, além de fornecer argumentos que fortalecem a justificativa da continuidade dos estudos relacionados a estes grupos de espécies, esta tese fornece inúmeras novas perspectivas, tais como:

- A expansão das análises envolvendo *Amd* e *Ddc* para outras espécies e gêneros de Drosophilidae;
- A inclusão de marcadores de evolução lenta, tais como o gene 28S, nestas análises;
- A realização de análises filogeográficas em algumas espécies estudadas, como por exemplo, o grupo *cardini*, a fim de analisar diversos aspectos de sua história evolutiva e sua estruturação populacional;
- A realização de uma análise filogeográfica envolvendo as semi-espécies de *D. paulistorum*, para acessar sua diferenciação genética e sua estruturação;

- A realização de uma análise filogenética de endossimbiontes do grupo *willistoni*, a fim de evidenciar a correlação entre a filogenia dos hospedeiros e aquela de seus parasitas;

- O estudo dos padrões de evolução molecular de *Amd* e *Ddc*, bem como outros genes, em *D. immigrans* e *D. mediopicta*, que se destacaram como as espécies que apresentam as menores e maiores taxas evolutivas no subgênero *Drosophila*, respectivamente;

- Evidenciar os padrões de expressão dos genes *Amd* e *Ddc* no subgênero *Drosophila*, a fim de inferir mudanças em seu padrão espaço-temporal de expressão;

- Analisar a continuidade da ligação dos genes *Amd* e *Ddc* no cluster gênico *Ddc*, nos genomas disponíveis de *Drosophila*;

- Analisar a evolução do tamanho dos íntrons de *Amd* e *Ddc* ao longo do subgênero *Drosophila*, avaliando seus padrões de evolução no grupo.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

- Adams MD, et al. (2000) The genome sequence of *Drosophila melanogaster*. Science 287:2185-2195.
- Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Contr 19:716-723.
- Ashburner M, Bodmer M, Lemeunier F (1984) On the evolutionary relationships of *Drosophila melanogaster*. Dev Genet 4:295-312.
- Ayala FJ, Mourão CA, Pérez-Salas S, Richmond R, Dobzhansky T (1970) Enzyme variability in the *Drosophila willistoni* group. I – Genetic differentiation among sibling species. Proc Natl Acad Sci 67:225-232.
- Ayala FJ, Tracey ML (1973) Genetic differentiation and reproductive isolation between two subspecies of *Drosophila willistoni*. J Hered 64:120-124.
- Ayala FJ, Tracey ML, Hedgecock D, Richmond RC (1974) Genetic differentiation during the speciation process in *Drosophila*. Evolution 28:576-592.
- Bächli G (2008) TaxoDros: The Database on Taxonomy of Drosophilidae, v.1.03, Database 2008/01. <http://taxodros.unizh.ch/>. Last accessed on 08/01/2008.
- Bächli G, Vilela CR, Ratcov V (2000) Morphological differences among *D. paraguayensis* Duda, 1927 and its close relatives (Diptera, Drosophilidae). Mitt Schweiz Ent Ges 73:67-92.

- Baker RH, DeSalle R (1997) Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Syst Biol* 46:654-673.
- Baker RH, Yu X, DeSalle R (1998) Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. *Mol Phylogenet Evol* 9:427-436.
- Barrio E, Ayala FJ (1997) Evolution of the *Drosophila obscura* species group inferred from *Gpdh* and *Sod* genes. *Mol Phylogenet Evol* 7:79-93.
- Bernardi N (1981) Phylogenetic relationships, monophyletic group and related concepts. *Revta Bras Entom* 25:323-326.
- Bonacum J, O'Grady PM, Kambysellis M, DeSalle R (2005) Phylogeny and age diversification of the *planitibia* species group of the Hawaiian *Drosophila*. *Mol Phylogenet Evol* 37:73-82.
- Bremer K (1988) The limits of amino acid sequence data in angiosperm phylogenetics reconstruction. *Evolution* 42:795-803.
- Bremer K (1994) Branch support and tree stability. *Cladistics* 10:295-304.
- Breuer ME, Pavan C (1950) Genitália masculina de *Drosophila* (Diptera): grupo *annulimana*. *Rev Brasil Biol* 10:469-488.
- Brisson JA, Wilder J, Hollocher H (2006) Phylogenetic analysis of the *cardini* group of *Drosophila* with respect to changes in pigmentation. *Evolution* 60:1228-1241.
- Brcic D (1957) A comparative study of chromosomal variation in species of the *mesophragmatica* group of *Drosophila*. *Genetics* 42:789-805

- Brcic D (1958) The *mesophragmatica* group as an example of the speciation phenomena in *Drosophila*. Proceedings of the X International Congress of Genetics:420-433.
- Brcic D, Koref-Santibañez S (1957) The *mesophragmatica* species group of *Drosophila*. Evolution 11:300-310.
- Brcic D, Nair PS, Wheeler MR (1971) Cytotaxonomic relationships within the *mesophragmatica* species group of *Drosophila*. Univ Texas Publ 7103:1-16.
- Brcic D (1978) A note on the *flavopilosa* group of species of *Drosophila* in Rio Grande do Sul, Brazil, with the description of two new species (Diptera, Drosophilidae). Rev Brasil Biol 38:647-651.
- Burla H, DaCunha AB, Cordeiro AR, Dobzhansky T, Malogolowkin C, Pavan C (1949) The *willistoni* group of sibling species of *Drosophila*. Evolution 3:300-314.
- Burla H, Pavan C (1953) The *calloptera* group of species (*Drosophila*, Diptera). Rev Bras Biol 13:291-314.
- Caccone A, Moriyama EN, Gleason JM, Nigro L, Powell JR (1996) A molecular phylogeny for the *Drosophila melanogaster* subgroup and the problem of polymorphism data. Mol Biol Evol 13:1224-1232.
- Caletka BC, McAllister BF (2004) A genealogical view of chromosomal evolution and species delimitation in the *Drosophila virilis* species subgroup. Mol Phylogenetic Evol 33:664-670.
- Carmody G (1965) Two unusual strains of the *D. willistoni* sibling species group. *Drosophila Inf Serv* 40:53.

- Carrasco SF, Prado LF, Godoy-Herrera R (2003) Molecular phylogeny of the *mesophragmatica* species group inferred from *Cytochrome Oxidase II* sequence. *Drosophila Inf Serv* 86: 72-75
- Clark AG, et al. (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450:203-218.
- Clary DO, Goddard JM, Martin SC, Fauron CM, Wolstenholme DR (1982) *Drosophila* mitochondrial DNA: a novel gene order. *Nucleic Acids Res* 10:6619-6637.
- DaLage JL, Dergoat GJ, Maczkowiak F, Silvain JF, Cariou ML, Lachaise D (2007) A phylogeny of Drosophilidae using the *Amyrel* gene: questioning the *Drosophila melanogaster* species group boundaries. *J Zool Syst Evol Res* 45:47-63.
- DeSalle R (1992a) The origin and possible time of divergence of the Hawaiian Drosophilidae: Evidence from DNA sequences. *Mol Biol Evol* 9: 05-916.
- DeSalle R (1992b) The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. *Mol Phylogenet Evol* 1:31-40.
- DeSalle R, Grimaldi D (1992) Characters and the Systematics of Drosophilidae. *J Hered* 83:182-188.
- Dobzhansky, T.; Ehrman, L.; Pavlovsky, O.; Spassky, B. (1964) The superespecies *D. paulistorum*. *Proc. Natl. Acad. Sci.* 51: 3-9.
- Dobzhansky T, Pavan C (1943) Studies on Brazilian species of *Drosophila*. *Bol Facul Fil Ciên Letr Univ S Paulo, São Paulo*.
- Dobzhansky T, Pavlovsky O, Ehrman L (1969) Transitional Populations of *Drosophila paulistorum*. *Evolution* 23:482-492.

- Dobzhansky T, Spassky B (1959) *Drosophila paulistorum* a cluster of species in *statu nascendi*. Proc Natl Acad Sci 45:419-128.
- Duda, O (1927) Die Sudamerikanischen Drosophiliden unter Berücksichtigung auch der anderen Neotropischen sowie Nearktischen Arten. Archiv für Naturgeschichte 91(A)11/12: 1-229
- Durando CM, Baker RH, Etges WJ, Heed WB, Wasserman M, DeSalle R (2000) Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. Mol Phylogenet Evol 16:296-307.
- Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.
- Ehrman L (1960) Genetics of hybrid sterility in *Drosophila paulistorum*. Evolution 14:212-223.
- Ehrman L (1965) Direct observation of sexual isolation between allopatric and between sympatric strains of the different *Drosophila paulistorum* races. Evolution 19:459-464.
- Ehrman L, Powell JR (1982) The *Drosophila willistoni* species group. In: Ashburner M, Carson HL, Thompson JN (eds) The Genetics and Biology of *Drosophila*. Academic Press, New York, pp 193-225.
- Eveleth DD, Marsh JL (1986) Evidence for evolutionary duplication of genes in the dopa decarboxylase region of *Drosophila*. Genetics 114:469-483.
- Farris JS, Källersjö M, Kluge A, Bult C (1994) Testing significance of congruence. Cladistics 10:315-320.



- Farris JS, Källersjö M, Kluge A, Bult C (1995) Constructing a significance test for incongruence. *Syst Biol* 44:570-572.
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using bootstrap. *Evolution* 39:783-791.
- Franc G, Napp M (1992) Electrophoretic analysis of interspecific relationships in six species of the *Drosophila tripunctata* group from South of Brazil. *Evol Biol* 6:121-133.
- Fruta-Pessoa O (1954) Revision of the *tripunctata* group of *Drosophila* with description of fifteen new species (Drosophilidae, Diptera). *Arq Mus Paranaense* 10:253-304.
- Galtier N, Gouy M (1995) Inferring phylogenies from DNA sequences of unequal base compositions. *Proc Natl Acad Sci* 92:11317-11321.
- Galtier N, Gouy M (1998) Inferring pattern and process: maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. *Mol Biol Evol* 15:871-879.
- Gao J, Watabe H, Aotsuka T, Pang J, Zhang Y (2007) Molecular phylogeny of the *Drosophila obscura* species group, with emphasis on the Old World species. *BMC Evol Biol* 7(87).
- Garcia ACL, Rohde C, Audino GF, Valente VLS, Valiati VH (2006) Identification of the sibling species of the *Drosophila willistoni* subgroup through the electrophoretic mobility of acid phosphatase-1. *J Zool Syst Evol Res* 44:212-216.

- Garcia RN, D'Avila MF, Robe LJ, Loreto EL, Panzera Y, de Heredia FO, Valente VL. (2007) First evidence of methylation in the genome of *Drosophila willistoni*. *Genetica* 131:91-105.
- Gleason JM, Caccone A, Moriyama EN, White KP, Powell JR (1997) Mitochondrial DNA phylogenies for the *Drosophila obscura* group. *Evolution* 51:433-440.
- Gleason JM, Griffith EC, Powell JR (1998) A molecular phylogeny of the *Drosophila willistoni* group: conflicts between species concepts? *Evolution* 52:1093-1103.
- Gleason JM, Powell JR (1997) Interspecific and intraspecific comparisons of the *period* locus in the *Drosophila willistoni* sibling species. *Mol Biol Evol* 14:741-753.
- Goto SG, Kimura MT (2001) Phylogenetic utility of mitochondrial *COI* and nuclear *Gpdh* genes in *Drosophila*. *Mol Phylogenet Evol* 18:404-422.
- Grimaldi DA (1987) Amber fossil Drosophilidae (Diptera), with particular reference to the Hispaniolan taxa. *Am Mus Novitates* 2880:1-23.
- Grimaldi DA (1988) Relicts in the Drosophilidae (Diptera). In: Liebherr JK (ed) *Zoogeography of Caribbean Insects*. Comstock Publishing Association, Cornell University Press, New York, pp 183-312.
- Grimaldi DA (1990) A phylogenetic, revised classification of the genera in the Drosophilidae (Diptera). *Bull Am Mus Natl Hist* 197:1-139.
- Grimaldi DA, James AC, Jaenike J (1992) Systematics and modes of reproductive isolation in the holartic *Drosophila testacea* species group (Diptera: Drosophilidae). *Ann Ent Soc Am* 85:671-685.

- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696-704.
- Haffer J (1967) Speciation in Colombian forest birds west of the Andes. *Am Mus Novit* 294:1-57.
- Haffer J (1969) Speciation in Amazonian forest birds. *Science* 165:131-137.
- Hall BG (2001) *Phylogenetic trees made easy – A how to manual for molecular biologists*. 2<sup>nd</sup> ed. Sinauer Associates, Massachusetts.
- Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22:160-174.
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proc R Soc Lond* 270:313-321.
- Heed WB (1962) Genetic characteristics of island populations. *Univ Texas Publ* 6205:173-206.
- Heed WB, Krishnamurthy NB (1959) Genetic studies on the *cardini* group of *Drosophila* in the West Indies. *Univ Texas Publ* 5914:155-179.
- Heed WB, Russel JS (1971) Phylogeny and population structure in island and continental species of the *cardini* group of *Drosophila* studied by inversion analysis. *Univ Texas Publ* 7103: 91-130.
- Hennig W (1965) *Die Acalyptratae des Baltischen Bernsteins*. *Stuttg Beitr Naturkd* 145.
- Hennig W (1981) *Insect Phylogeny*. John Wiley & Sons, New York.

- Hofmann PRP (1985) Variabilidade genética em espécies de nível ecológico restrito. *Ciênc. Cult.* 37:579-581.
- Huelsenbeck JP, Bull JJ, Cunningham CW (1996) Combining data in phylogenetic analysis. *Trends Ecol Evol* 11:152–158.
- Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.
- Hunter AS, Hunter RA (1964) The *mesophragmatica* species group of *Drosophila* in Colombia. *Ann Entomol Soc Am* 57: 732-736
- Jaeger CP, Salzano FM (1953) *Drosophila gaucha*, a new species from Brazil. *Rev Bras Biol* 13: 205-208
- Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple Sequence Alignment with ClustalX. *Trends Biochem Sci* 23:403-405.
- Jowett T (1986) Preparation of nucleic acids. In: Roberts DB *Drosophila: a practical approach*. IRL Press, Oxford, pp 275-276.
- Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HM (ed) *Mammalian protein metabolism*. Academic Press, New York, pp 21-132.
- Kastanis P, Eliopoulos E, Goulielmos GN, Tsakas S, Loukas M (2003) Macroevolutionary relationships of species of *Drosophila melanogaster* group based on mtDNA sequences. *Mol Phylogenet Evol* 28:518-528.
- Kastritsis CD (1969) The chromosomes of some species of the *guarani* group of *Drosophila*. *J Heredity* 60:51-57.

- Kastritsis CD, Dobzhansky, Th (1967) *Drosophila pavlovskiana*, a race or a species? *Am Midl Nat* 78:244-247.
- Katoh T, Tamura K, Aotsuka T (2000) Phylogenetic position of the subgenus *Lordiphosa* of the genus *Drosophila* (Diptera: Drosophilidae) inferred from Alcohol Dehydrogenase (*Adh*) gene sequences. *J Mol Evol* 51:122-130.
- Katoh T, Nakaya D, Tamura T, Aotsuka T (2007) Phylogeny of the *Drosophila immigrans* species group (Diptera: Drosophilidae) based on *Adh* and *Gpdh* sequences. *Zool Science* 24:913-921.
- Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16: 111-120
- King JC (1947) Interspecific relationships within the *guarani* group of *Drosophila*. *Evolution* 1:143-153.
- Kluge AG (1989) A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Syst Zool* 38:7–25.
- Kumar S, Gadagkar SR (2001) Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. *Genetics* 158:1321-1327.
- Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA: Molecular Evolutionary Genetics Analysis Software. Version 2. Pennsylvania State University, University Park, and Arizona State University, Tempe.

- Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5:150-163.
- Kwiatowski J, Ayala FJ (1999) Phylogeny of *Drosophila* and related genera: Conflict between molecular and anatomical analyses. *Mol Phylogenet Evol* 13:319-328.
- Kwiatowski J, Krawczyk M, Jaworski M, Skarecky D, Ayala FJ (1997) Erratic Evolution of Glycerol-3-Phosphate Dehydrogenase in *Drosophila*, *Chymomyza*, and *Ceratitis*. *J Mol Evol* 44:9-22.
- Kwiatowski J, Skarecky D, Bailey K, Ayala FJ (1994) Phylogeny of *Drosophila* and related genera inferred from the nucleotide sequence of the Cu, Zn *Sod* gene. *J Mol Evol* 38:443-454.
- Latorre A, Barrio E, Moya A, Ayala FJ (1988) Mitochondrial DNA evolution in the *Drosophila obscura* group. *Mol Biol Evol* 5:717-728.
- Lakovaara S, Saura A (1982) Evolution and speciation in the *Drosophila obscura* group. In: Ashburner M, Carson HL, Thompson JN (eds) *The Genetics and Biology of Drosophila*. Academic Press, New York, pp 1-59.
- Lewis RL, Beckenbach AT, Mooers AO (2005) The phylogeny of the sobgroups within the *melanogaster* species group: Likelihood tests on *COI* and *COII* sequences and a Bayesian estimate of phylogeny. *Mol Phylogenet Evol* 37:15-24.
- Lin CP, Danforth BN (2004) How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analysis of combined datasets. *Mol Phylogenet Evol* 30:686-702.

- Linde K, Bächli G, Toda MJ, Zhang WX, Hu YG, Spicer GS (2007) Case 3407: *Drosophila* Fallén, 1832 (Insecta, Diptera): proposed conservation of usage. *Bull Zool Nomencl* 64.
- Lockhart PJ, Stell MA, Hendy MD, Penny D (1994) Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol Biol Evol* 11:605-612.
- Malogolowkin C (1953) Sobre a genitália dos drosofilídeos. IV. A genitália masculina no subgênero *Drosophila* (Diptera, Drosophilidae) *Rev Bras Biol* 13:245-264.
- Markow TA, O'Grady PM (2006) Phylogenetic relationships of Drosophilidae. In: Markow TA, O'Grady PM *Drosophila: A guide to species identification and use*. Academic Press, London, pp 3-64.
- McAlpine JF (1989) Phylogeny and classification of the Muscomorpha. In: McAlpine JF, Wood DM (eds) *Manual of Nearctic Diptera v3*. Research Branch, Agriculture Canada, Monograph 32:1397-1518.
- Meier R (2005) Role of Dipterology in phylogenetic systematics: the insights of Willi Hennig. In: Yeates DK, Wiegmann BM (eds) *The evolutionary biology of flies*, Columbia University Press, New York, pp 45-62.
- Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear gene-trees. *Evolution* 49:718-726
- Morán T, Fontdevila A (2007) On the phylogeny of the *Drosophila hydei* subgroup: New insights from combined analyses of nuclear and mitochondrial data. *Mol Phylogenet Evol* 43:1198-1205.
- Moritz C, Cicero C (2004) DNA Barcoding: promise and pitfalls. *PLoS Biol* 2:e354.

- Mota NR, Robe LJ, Valente VLS, Budnik M, Loreto ELS (2008) Phylogeny of the *Drosophila mesophragmatica* group (Diptera, Drosophilidae): an example of Andean evolution. *Zool Sci* (in press)
- Nacur J (1958) Genitalia masculina de *Drosophila* do grupo *mesophragmatica* (Diptera). *Rev Brasil Biol* 18:243-249.
- Nair PS, Brncic D, Kojima KI (1971) II. Isozyme variations and evolutionary relationships in the *mesophragmatica* species group of *Drosophila*. *Stud Genet* VI:17-28.
- Napp M, Cordeiro AR (1981) Interspecific relationships in the *cardini* group of *Drosophila* studied by electrophoresis. *Rev Brasil Biol* 4:537-547.
- Nei M, Kumar S (2000) *Molecular evolution and phylogenetics*. Oxford, New York, 333 pp.
- Nurminsky DI, Moriyama EN, Lozovskaya ER, Hartl DL (1996) Molecular phylogeny and genome evolution in the *Drosophila virilis* species group: duplications of the *Alcohol Dehydrogenase* gene. *Mol Biol Evol* 13:132-149.
- Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Center, Uppsala University.
- O'Grady PM (1999) Reevaluation of phylogeny in the *Drosophila obscura* species group based on combined analysis of nucleotide sequences. *Mol Phylogenet Evol* 12:124-139.
- O'Grady PM, Baker RH, Durando CM, Etges WJ, DeSalle R (2001) Polytene chromosomes as indicators of phylogeny in several species groups of *Drosophila*. *BMC Evol Biol* 1(6).



- O'Grady PM, Clark JB, Kidwell MG (1998) Phylogeny of the *Drosophila saltans* species group based on combined analysis of nuclear and mitochondrial DNA sequences. *Mol Biol Evol* 15:656-664.
- O'Grady PM, Kidwell MG (2002) Phylogeny of the subgenus *Sophophora* (Diptera: Drosophilidae) based on combined analysis of nuclear and mitochondrial sequences. *Mol Phylogenet Evol* 22:442-453.
- Oliveira DCS, Wintermute EH, Etges WJ, Heed WB, O'Grady PM, DeSalle R (2003) Hierarchical structure in the *Drosophila mojavensis* cluster (Diptera: Drosophilidae). *Hereditas* 139:223-227.
- Pagel M, Meade A (2004) A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Syst Biol* 53:571-581.
- Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Mol Biol Evol* 5:568-583.
- Patterson JT (1943) The Drosophilidae of the Southwest. *Univ Texas Publ* 4313:7-216
- Patterson JT (1957) A study of interspecific hybridization between members of the *tripunctata* group of *Drosophila*. *Univ Texas Publs* 5721:7-14.
- Pélandakis M, Higgins DG, Solignac M (1991) Molecular phylogeny of the subgenus *Sophophora* of *Drosophila* derived from large subunit of ribosomal RNA sequences. *Genetica* 84:87-94.
- Pélandakis M, Solignac M (1993) Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. *J Mol Evol* 37:525-543.

- Pérez-Salas S, Richmond RC, Pavlovsky OA, Kastritsis CD, Ehrman L, Dobzhansky T (1970) The Interior semispecies of *Drosophila paulistorum*. *Evolution* 24:519-527.
- Perlman SJ, Spicer GS, Shoemaker D, Jaenike J (2003) Associations between mycophagous *Drosophila* and their *Howardula* nematode parasites: a worldwide phylogenetic shuffle. *Mol Ecol* 12:237-249.
- Pollard DA, Iyer VN, Moses AM, Eisen MB (2006) Widespread discordance of gene trees with species trees in *Drosophila*: Evidence for incomplete lineage sorting. *PLoS Genet* 2:e173.
- Posada C, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14:817-818.
- Powell JR, DeSalle R (1995) *Drosophila* molecular phylogenies and their uses. In: Hecht MK et al. (eds) *Evolutionary Biology*. Plenum, New York, pp. 87-139.
- Powell JR, Sezzi E, Moriyama EN, Gleason JM, Caccone A (2003) Analysis of a shift in codon usage in *Drosophila*. *J Mol Evol* 57:214-225.
- Pringent S, Renard E, Cariou ML (2003) Electrophoretic mobility of amylase in *Drosophilids* indicates adaptation to ecological diversity. *Genetica* 119:133-145.
- Queiroz A, Donoghue MJ, Kim J (1995) Separate versus combined analysis of phylogenetic evidence. *Annu Rev Ecol Syst* 26:657-681.
- Remsen J, DeSalle R (1998) Character congruence and multiple data partitions and the origin of the Hawaiian *Drosophilidae*. *Mol Phylogenet Evol* 9:225-235.

- Remsen J, O'Grady PO (2002) Phylogeny of Drosophilinae (Diptera: Drosophilidae) with comments on combined analysis and character support. *Mol Phylogenet Evol* 24:249-264.
- Robe LJ, Silva LB, Loreto ELS (2002) Phylogenetic relationships among four species of the *guarani* group of *Drosophila* (Diptera, Drosophilidae) as inferred by molecular and morphological analyses. *Rev Bras Ent* 46:515-519.
- Robe LJ, Valente VLS, Budnik M, Loreto ELS (2005) Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: a nuclear versus mitochondrial gene approach. *Mol Phylogenet Evol* 36:623-640.
- Rodriguez F, Oliver JL, Marin A, Medina JR (1990) The general stochastic model of nucleotide substitution. *J Theor Biol* 142:485-501.
- Rodríguez-Trelles F, Tarrío R, Ayala FJ (1999a) Molecular evolution and phylogeny of the *Drosophila saltans* species group inferred from the *Xdh* gene. *Mol Phylogenet Evol* 13:110-121.
- Rodríguez-Trelles F, Tarrío R, Ayala FJ (1999b) Switch in codon bias and increased rates of amino acid substitution in the *Drosophila saltans* species group. *Genetics* 153:339-350.
- Rodríguez-Trelles F, Tarrío R, Ayala FJ (2000a) Fluctuating mutation bias and the evolution of base composition in *Drosophila*. *J Mol Evol* 50:1-10.
- Rodríguez-Trelles F, Tarrío R, Ayala FJ (2000b) Evidence for a high ancestral GC content in *Drosophila*. *Mol Biol Evol* 17:1710-1717.

- Rodríguez-Trelles F, Alarcón L, Fontdevila A (2000c) Molecular evolution and phylogeny of the *buzzatii* complex (*Drosophila repleta* group): a maximum-likelihood approach. *Mol Biol Evol* 17:1112-1122.
- Rohde C, Garcia ACL, Valiati VH, Valente VLS (2006) Chromosomal evolution of sibling species of the *Drosophila willistoni* group. I. Chromosomal arm IIR (Muller's element B). *Genetica* 126:77-88.
- Rokas A, Williams BL, King N, Carroll SB (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425:798-804.
- Ruiz A, Wasserman M, Heed WB (1990) Evolution of the *mojavensis* cluster of cactophilic *Drosophila* with descriptions of two new species. *J Heredity* 81:30-42.
- Russo CAM (2001) Como escolher genes para problemas filogenéticos específicos. In: Matioli SR (ed) *Biologia molecular e evolução*. Holos, Ribeirão Preto.
- Russo CAM, Takezaki N, Nei M (1995) Molecular phylogeny and divergence times of *Drosophilid* species. *Mol Biol Evol* 12:391-404.
- Rutschmann F (2005) Bayesian molecular dating using PAML/multidivtime. A step-by-step manual. University of Zurich, Switzerland. Available at <http://www.plant.ch/>.
- Salgado-Laboriau ML (1994) O Período Quaternário. In: Salgado-Laboriau ML (ed) *História Ecológica da Terra*. Editora Edgard Blücher, São Paulo, pp 255-280.
- Salzano FM (1955) Chromosomal polymorphism in two species of the *guarani* group of *Drosophila*. *Chromosoma* 7:39-50.
- Saitou N, Nei M (1987) The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 9:406-425.

- Santos RCO, Vilela CR (2005) Breeding sites of Neotropical Drosophilidae (Díptera). IV. Living and fallen flowers of *Sessea brasiliensis* and *Cestrum* spp. (Solanaceae). Rev Bras Entomol 49:544-551.
- Sassi AK, Heredia F, Loreto ELS, Valente VLS, Rohde C (2005) Transposable elements P and gypsy in natural populations of *Drosophila willistoni*. Genet Mol Biol 28:734-739.
- Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18:502-504.
- Schneider H (2003) Métodos de análise filogenética – Um guia prático. Holos and SBG, Ribeirão Preto.
- Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol 16:1114-1116.
- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am 87:651-701.
- Singh ND, Arndt PF, Petrov DA (2006) Minor shift in background substitutional patterns in the *Drosophila saltans* and *willistoni* lineages is insufficient to explain GC content of coding sequences. BMC Biol 4(37).
- Sorenson MD, Franzosa EA (2007) TreeRot version 3. Boston University, Boston, MA.
- Spassky B (1957) Morphological differences between sibling species of *Drosophila*. Univ Texas Publ 5721:48-61.

- Spassky B, Richmond RC, Perez-Salas S, Pavlovsky O, Mourão CA, Hunter AS, Hoentgsberg H, Dobzhansky T, Ayala FJ (1971) Geography of the sibling species related to *Drosophila willistoni*, and of the semispecies of the *Drosophila paulistorum* complex. *Evolution* 25:129-143.
- Spicer GS (1991) Molecular evolution and phylogeny of the *Drosophila virilis* species group as inferred by two-dimensional electrophoresis. *J Mol Evol* 33:379-394.
- Spicer GS (1995) Phylogenetic utility of the mitochondrial *cytochrome oxidase* gene: molecular evolution of the *Drosophila buzzatii* species complex. *J Mol Evol* 41:749-759.
- Spicer GS, Jaenike J (1996) Phylogenetic analysis of breeding site use and alpha-amanitin tolerance within the *Drosophila quinaria* species group. *Evolution* 50:2328-2337.
- Spicer GS, Pitnick S (1996) Molecular systematics fo the *Drosophila hydei* subgroup as inferred from mitochondrial DNA sequences. *J Mol Evol* 43:281-286.
- Staden R (1996) The Staden sequence analysis package. *Mol Biotechnology* 5:233-241.
- Stalker H (1966) The phylogenetic relationships of the species in the *Drosophila melanica* group. *Genetics* 53:327-342.
- Stalker H (1973) Intergroup phylogenies in *Drosophila* as determined by comparisons of salivary banding patterns. *Genetics* 70:457-474.
- Sturtevant AH (1921) The North American Species of *Drosophila*. *Publs Carnegie Instn* 301:1-150.
- Sturtevant AH (1939) On the subdivision of the genus *Drosophila*. *Proc Natl Acad Sci* 25:137-141.

- Sturtevant AH (1942) The classification of the genus *Drosophila* with the description of nine new species. Univ Texas Publ 4213:5-51.
- Swofford DL (2003) PAUP: Phylogenetic Analysis using Parsimony (and other methods). Version 4. Sinauer Associates, Massachusetts.
- Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: Hillis DM, Moritz C, Mable BK (eds) Molecular Systematics. Sinauer, Sunderland pp. 407-514.
- Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C biases. Mol Biol Evol 9:678-687.
- Tamura K, Nei M (1993) Estimation of the number of base nucleotide substitution in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512-526.
- Tamura K, Subramanian S, Kumar S (2004) Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. Mol Biol Evol 21:36-44.
- Tamura K, Toba G, Park J, Aotsuka T (1995) Origin of Hawaiian drosophilids inferred from alcohol dehydrogenase gene sequences. In: Nei M, Takahata N (eds) Current topics on molecular evolution. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park. pp 9-18.
- Tarrío R, Rodríguez-Trelles F, Ayala FJ (2000) Tree rooting with outgroups when they differ in their nucleotide composition from the ingroup: The *Drosophila saltans* and *willistoni* groups, a case study. Mol Phylogenet Evol 16:344-349.

- Tarrío R, Rodríguez-Trelles F, Ayala FJ (2001) Shared nucleotide composition biases among species and their impact on phylogenetic reconstructions of the Drosophilidae. *Mol Biol Evol* 18:1464-1473.
- Tatarenkov A, Ayala FJ (2001) Phylogenetic relationships among species groups of the *virilis-repleta* radiation of *Drosophila*. *Mol Phylogenet Evol* 21:327-331.
- Tatarenkov A, Kwiatowski J, Sharecky D, Barrío E, Ayala FJ (1999a) On the evolution of *Dopa decarboxylase (Ddc)* and *Drosophila* systematics. *J. Mol. Evol.* 48:445-462.
- Tatarenkov A, Sáez AG, Ayala FJ (1999b) A compact gene cluster in *Drosophila*: the unrelated *Cs* gene is compressed between duplicated *Amd* and *Ddc*. *Gene* 231:111-120.
- Tatarenkov A, Zurovcová M, Ayala FJ (2001) *Ddc* and *Amd* sequences resolve phylogenetic relationships of *Drosophila*. *Mol Phylogenet Evol* 20:321-325.
- Templeton AR (1983) Convergent evolution and non-parametric inferences from restriction fragment and DNA sequence data. In: Weir B (ed) *Statistical Analysis of DNA Sequence Data*. Marcel Dekker, New York, pp 151-179.
- Thomas RH, Hunt JA (1991) The molecular evolution of the Alcohol Dehydrogenase locus and the phylogeny of Hawaiian *Drosophila*. *Mol Biol Evol* 8:687-702.
- Thomas RH, Hunt JA (1993) Phylogenetic relationships in *Drosophila*: a conflict between molecular and morphological data. *Mol Biol Evol* 10:362-374.
- Thompson FC (ed) (2005) *Biosystematic Database of World Diptera*. <http://www.diptera.org/biosys.htm>. Last accessed on 10/08/2007.



- Thorne JL, Kishino H (2002) Divergence time and evolutionary rate estimation with multilocus data. *Syst Biol* 51:689–702.
- Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. *Mol Biol Evol* 15:1647–1657.
- Throckmorton LH (1962) The problem of phylogeny in the genus *Drosophila*. *Stud. Genet. II. Univ Texas Publ* 6205:207-343.
- Throckmorton LH (1966) The relationships of the endemic Hawaiian Drosophilidae. *Univ Texas Publ* 6615:335-396.
- Throckmorton LH (1975) The phylogeny, ecology and geography of *Drosophila*. In: King, R. C. (ed) *Handbook of Genetics*. Plenum, New York, pp 421-469.
- Throckmorton LH (1982) The *virilis* species group. In: Ashburner M, Carson HL, Thompson JNJ (eds) *The genetics and biology of Drosophila*. Academic Press, New York, pp 227-296.
- Throckmorton LH, Magalhães LE (1962) Changes with evolution of pteridine accumulations in species of the *saltans* group of the genus *Drosophila*. *Univ Texas Publ* 6205:489-505.
- Ting CT, Tsaur SC, Wu CI (2000) The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. *Proc Natl Acad Sci* 97: 5313–5316.
- Townsend JL (1954) Cryptic subspeciation in *Drosophila* belonging to the subgenus *Sophophora*. *Amer Nat* 88:339-351.

- Toda MJ (2007) DrosWLD-Species: Taxonomic Information Database for World Species of Drosophilidae. <http://bioinfo.lowtem.hokudai.ac.jp/db/modules/stdb/>. Last accessed on 15/08/2007.
- Tosi D, Pereira MAQR (1993) Karyotypes and phylogenetic relationships in the *Drosophila* species of the *annulimana* group (Diptera: Drosophilidae). Rev Bras Genet 16:321-333.
- Treier M, Pfeifle C, Tautz D (1989) Comparison of the gap segmentation gene *hunchback* between *Drosophila melanogaster* and *Drosophila virilis* reveals novel modes of evolutionary change. EMBO J 8:1517-1525.
- Tsacas L, Cariou ML, Lachaise D (1985) Le groupe *Drosophila obscura* en Afrique de l'Est, description de trois nouvelles especes (Diptera: Drosophilidae). Ann Soc Ent Fr 21:413-424.
- Val FC, Marques MD, Vilela CR (1981) Drosophilidae of Neotropical Region. In: Ashburner M, Carson HL, Thompson JN (eds) The Genetics and Biology of *Drosophila*. Academic Press, New York, pp 123-167.
- Vela D, Rafael V (2004) Three new andean species of *Drosophila* (Diptera, Drosophilidae) of the *mesophragmatica* group. Iheringia 94:295-299.
- Vilela CR (1983) A revision of the *Drosophila repleta* species group (Diptera, Drosophilidae). Rev Bras Entomol 27: 1-114
- Vilela CR, Bächli G (1990) Taxonomic studies on Neotropical species of seven genera of Drosophilidae (Diptera). Mitt Schweiz Ent Ges 63 (suppl.):1-332.

- Vilela CR, Bächli G (2000) Five new species of Neotropical *Drosophila* (Diptera, Drosophilidae). Mitt Schweiz Ent Ges 73:49-65.
- Vilela CR, Pereira MAQR (1982) A new species of the *annulimana* group of the genus *Drosophila* (Diptera, Drosophilidae). Rev Bras Ent 26:237-240.
- Vilela CR, Pereira MAQR (1985) On the male genitalia of five species of the *tripunctata* group of *Drosophila* (Diptera, Drosophilidae). Rev Bras Ent 29:453-461.
- Vilela CR, Pereira MAQR (1986) The male genitalia of four species of the *tripunctata* group of *Drosophila* (Diptera, Drosophilidae). Rev Bras Ent 30:213-219.
- Vilela CR, Val FC (1985) The male genitalia of types of six members of the *Drosophila tripunctata* species group (Diptera, Drosophilidae). Rev Bras Ent 29:503-513.
- Vilela CR, Valente VLS, Basso-da-Silva L (2004) *Drosophila angustibucca* Duda sensu Frota-Pessoa is an undescribed species (Diptera, Drosophilidae). Rev Bras Entomol 48:233-238.
- Wang B, Park J, Watabe H, Gao J, Xiangyu J, Aotsuka T, Chen H, Zhang Y (2006) Molecular phylogeny of the *Drosophila virilis* section (Diptera: Drosophilidae) based on mitochondrial and nuclear sequences. Mol Phylogenet Evol 40:484-500.
- Ward BL, Heed WB (1970) Chromosome phylogeny of *Drosophila pachea* and related species. J Hered 61:248-258.1
- Wasserman M (1992) Cytological evolution of the *Drosophila repleta* species group. In: Krimbas CB, Powell JR (eds) *Drosophila Inversion Polymorphism*. CRC Press, Boca Raton, pp 455-552.

- Winge H (1965) Inter-specific hybridization between the six cryptic species of *willistoni* group. *Heredity* 20:9-19.
- Wheeler MR (1981) The Drosophilidae: a taxonomic overview. In: Ashburner M, Carson HL, Thompson JNJ (eds) *The genetics and biology of Drosophila*. Academic Press, Orlando, pp 1-97.
- Wheeler MR (1986) Additions to the catalog of the world's Drosophilidae. In: Ashburner M, Thompson JN, Carson HL (eds) *The genetics and biology of Drosophila*. Academic Press, London, pp 395-409.
- Whiting MF (2005) Phylogenetic position of Diptera: Review of evidence. In: Yeates DK, Wiegmann BM (eds) *The evolutionary biology of flies*. Columbia University Press, New York, pp 3-13.
- Wong A, Jensen JD, Pool JE, Aquadro CF (2006) Phylogenetic incongruence in the *Drosophila melanogaster* species group. *Mol Phylogenet Evol* 43:1138-1150.
- Xia X, Xie Z (2001) DAMBE: Data Analysis in Molecular Biology and Evolution. *J Hered* 92:371-373.
- Yang Y, Zhang Y, Qian Y, Zeng Q (2004) Phylogenetic relationships of *Drosophila melanogaster* species group deduced from spacer regions of *histone* gene H2A-H2B. *Mol Phylogenet Evol* 30:336-343.
- Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *BACIOS* 13:555-556.
- Yeates DK, Wiegmann BM (1999) Congruence and Controversy: Towards a higher-level phylogeny of Diptera. *Annu Rev Entomol* 44:397-428.

- Yeates DK, Wiegmann BM (2005) Phylogeny and Evolution of Diptera: Recent insights and new perspectives. In: Yeates DK, Wiegmann BM (eds) The evolutionary biology of flies. Columbia University Press, New York, pp 14-44.
- Yotoko KSC, Medeiros HF, Solferini VN, Klaczko LB (2003) A molecular study of the systematics of the *Drosophila tripunctata* group and the *tripunctata* radiation. Mol Phylogenet Evol 28:614-619.
- Young I, Coleman AW (2004) The advantages of the ITS2 region of the nuclear rDNA cistron for analysis of phylogenetic relationships of insects: a *Drosophila* example. Mol Phylogenet Evol 30:236-242.
- Zharkikh A (1994) Estimation of evolutionary distances between nucleotide sequences. J Mol Evol 39: 315-329.
- Zilversmit M, O'Grady P, DeSalle R (2002) Shallow genomics, phylogenetics, and evolution in the family Drosophilidae. Pac Symp Biocomput:512-523.

## ANEXO I

### (Referente aos Capítulos II, III, IV e V)

---

#### OS MARCADORES UTILIZADOS

A escolha do gene adequado ao problema filogenético em questão é uma etapa crucial na realização de qualquer estudo evolutivo. Alinhamentos perfeitos, em que a maior parte dos sítios estão conservados entre as seqüências, costumam ser pouco informativos. Por outro lado, quando as seqüências analisadas são muito divergentes, as diferenças estarão sendo demasiadamente subestimadas, devido à ampla ocorrência de substituições múltiplas, reversas e paralelas. Segundo Russo (2001), um bom indicador de quão adequado é o gene utilizado é o resultado final do alinhamento das seqüências, quando se recomenda que a proporção de diferenças entre duas seqüências alinhadas não ultrapasse 20-30%. Hillis *et al.* (1996), entretanto, afirmam que o uso do percentual de divergência entre seqüências homólogas na estimativa da utilidade do gene para a resolução do problema em foco pode ser desastroso, pois é possível que uma pequena proporção dos sítios se encontre totalmente livre de pressões seletivas, e, desta forma, saturado, enquanto a maioria dos sítios permaneça invariável. De modo geral, genes com altas taxas evolutivas são úteis para o estudo de indivíduos, famílias e populações, enquanto que aqueles que evoluem mais lentamente são melhor utilizados no estudo de espécies ou táxons supra-específicos (Solé-Cava, 2001). Muitas considerações devem, ainda, ser feitas, no que diz respeito à neutralidade do marcador escolhido. Segundo Avise (2004), o sentimento de que marcadores moleculares devem, necessariamente, ser neutros é uma supersimplificação,

embora este autor ressalte que intensidades diferentes de seleção direcional ou diversificadora geram uma heterogeneidade de taxas mutacionais entre as OTU's, enquanto que a seleção balanceadora pode reter polimorfismos moleculares particulares após sucessivos eventos de especiação.

Como o principal foco dos estudos realizados nesta tese foram relações entre espécies e grupos próxima e medianamente relacionados, em sua maior parte pertencentes a um mesmo subgênero, buscou-se marcadores com taxas evolutivas intermediárias. Neste caso, os escolhidos foram:

#### **O GENE *ADH* (ÁLCOOL DESIDROGENASE)**

O produto do gene *Adh* é o maior responsável pela capacidade de *Drosophila*, em seus estágios larvais ou adultos, de explorar diferentes substratos em estado de fermentação, uma vez que este gene confere resistência ao etanol (Linsley & Zimm, 1992). Neste sentido, a enzima *Adh* metaboliza a maior parte do álcool produzido pelas leveduras a partir da fermentação do substrato, convertendo este material em substâncias não-tóxicas e permitindo sua utilização como fonte energética. Além disso, *Adh* também catalisa a oxidação de acetaldeído a acetato (Linsley & Zimm, 1992). Em ambas atividades, *Adh* funciona como um homodímero.

A atividade de *Adh* varia durante o desenvolvimento de *Drosophila* (Goldberg, 1980): a enzima é primeiramente detectada em larvas de primeiro instar, aumentando sua atividade ao longo do desenvolvimento larval, e gradualmente declinando durante o estágio de pupa. Antes da emergência dos adultos, entretanto, *Adh* volta à atividade e atinge seu pico 5-10 dias depois. Estima-se que *Adh* compreenda cerca de 1% das proteínas presentes

em indivíduos adultos (Sofer & Ursprung, 1968), mesmo que seja expressa apenas em nível de corpos gordurosos e intestinos (Lindsley & Zimm, 1992).

Pela utilização de sítios de iniciação da transcrição alternativos, somada ao uso de sítios alternativos de término da transcrição e ao processo de *splicing* alternativo, *Adh* produz cinco RNA's mensageiros diferentes, mas que apresentam o mesmo potencial codificador (Entrez Gene, site NCBI) (Fig. 01). Embora eventos de *splicing* alternativo demarquem a presença de diversos íntrons dentro da região não-codificadora, apenas dois íntrons são encontrados dentro da região codificadora (Goldberg, 1980).

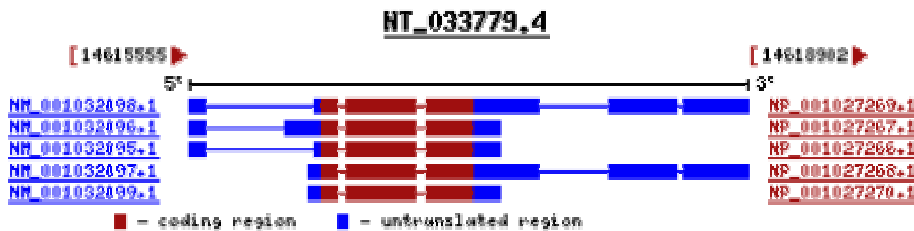


Figura 1 – Mapa genômico do gene *Adh* em *D. melanogaster*, conforme obtido no Entrez Gene, do site NCBI, contendo os números de acesso das seqüências de mRNA (à esquerda, em azul), das seqüências de proteínas (à direita, em vermelho), e da seqüência de nucleotídeos do gene inteiro (em cima, em preto), cuja extensão abrange os números posicionados acima da figura em ambas extremidades. Retângulos representam éxons enquanto que suas linhas conectoras indicam íntrons. Regiões não-traduzidas 5' e 3' estão demarcadas em azul, enquanto a região codificadora encontra-se demarcada em vermelho.

Duas características apresentadas pelo gene *Adh*, entretanto, desfavorecem sua utilização em estudos acerca das relações filogenéticas em *Drosophilidae*: (1) a inferência da atuação de uma seleção positiva, ao menos no que diz respeito ao grupo *melanogaster* de *Drosophila* (McDonald & Kreitman, 1991); (2) a existência de um grande número de duplicações envolvendo este gene (Russo *et al.*, 1995; Nurminsky *et al.*, 1996), o que dificulta a obtenção e o estabelecimento de ortologias.



## OS GENES PARÁLOGOS *AMD* (ALFA METILDOPA) E *DDC* (DOPA DESCARBOXILASE)

A cutícula dos insetos é uma estrutura altamente diferenciada que exibe um considerável grau de variabilidade e especialização. Parte destas propriedades é alcançada pela regulação da natureza e/ou extensão da esclerotização, estando os genes *Amd* e *Ddc* diretamente envolvidos nestes processos (Wang & Marsh, 1995). Estes são genes estrutural, funcional e evolutivamente relacionados, uma vez que apresentam uma relação de paralogia entre si, estando envolvidos no metabolismo de catecolaminas<sup>2</sup> durante a diferenciação terminal de *Drosophila* (Marsh *et al.*, 1986). Para tanto, codificam para descarboxilases que atuam nos eventos finais da diferenciação cuticular (Wang & Marsh, 1995). Segundo Lindsley & Zimm (1992), a duplicação que deu origem a estes genes precedeu a divergência Díptera-Lepidoptera, sendo que Tatarenkov *et al.* (1999b) chegam a extrapolar que esta duplicação pode ter, inclusive, precedido a divergência Protostoma-Deuterostoma.

*Amd* e *Ddc* desenvolveram, entretanto, funções não sobrepostas (Lindsley & Zimm, 1992; Wang & Marsh, 1995; Wang *et al.*, 1996), incluindo a produção de diferentes tipos de cutícula e, para *Ddc*, a produção dos neurotransmissores dopamina e serotonina.

### 1- O gene *Amd* (*alpha methyl dopa*)

Parálogo do gene *Ddc*, *Amd* é um gene essencial para a viabilidade e para a fertilidade de drosofilídeos. Heterozigotos para mutações em *Amd* caracterizam-se pela elevada mortalidade quando cultivados em níveis de alfa metildopa que não são letais para indivíduos selvagens, enquanto que a homozigose para mutações em *Amd* é letal (Lindsley & Zimm, 1992). Este gene parece ser particularmente ativo na produção de uma cutícula

---

<sup>2</sup> Catecolaminas servem como agentes de ligação cruzada para a esclerotização da cutícula, como precursores para a formação da melanina ou ainda como neurotransmissores.

flexível e incolor durante o desenvolvimento do embrião, quando é expresso primordialmente na epiderme, no intestino anterior e posterior, no proventrículo e no sistema traqueal (Wang *et al.*, 1996). Sua expressão é, ainda, requerida na linhagem germinativa e nas células foliculares femininas (onde *Amd* é responsável pela produção de uma membrana vitelínica flexível e impermeável, característica dos ovos de *Drosophila*), bem como nos discos imaginais, no sistema nervoso e nas glândulas linfáticas larvais.

O gene *Amd* apresenta um total de quatro éxons distintos, que, a partir dos processos de *splicing* alternativo e de sítio de início da transcrição alternativos, levam a formação de duas diferentes isoformas (Fig. 2). Segundo Wang & Marsh (1995) o íntron de *Amd* encontrado no transcrito primário NM\_057244 pode apresentar seqüências regulatórias importantes, tanto para alguns domínios de expressão embrionária, quanto para o desenvolvimento larval.



Figura 2 – Mapa genômico do gene *Amd* em *D. melanogaster*, conforme obtido no Entrez Gene, do site NCBI, contendo os números de acesso das seqüências de mRNA (à esquerda, em azul), das seqüências de proteínas (à direita, em vermelho), e da seqüência de nucleotídeos do gene inteiro (em cima, em preto), cuja extensão abrange os números posicionados acima da figura em ambas extremidades. Retângulos representam éxons enquanto que suas linhas conectoras indicam íntrons. Regiões não-traduzidas 5' e 3' estão demarcadas em azul, enquanto a região codificadora encontra-se demarcada em vermelho.

## 2- O gene *Ddc*

O gene *Ddc* codifica para dopa descarboxilase, enzima que catalisa a descarboxilação de dopa em dopamina, e de 5-hidroxi-triptofano em serotonina, mas não de tirosina em tiramina (Lindsley & Zimm, 1992). De acordo com estes autores, a maioria das

mutações em *Ddc* é letal em homozigose. *Ddc* funciona como um homodímero que requer piridoxal-5-fosfato para sua atividade. Atua na esclerotização da cutícula de drosofilídeos através da produção de dopamina, que após descarboxilada por *Ddc* é metabolizada para N-acetil-dopamina e N-β-alanil-dopamina, formas que são, então, oxidadas a suas respectivas quinonas, de forma a executar uma ligação cruzada com proteínas cuticulares (Lindsley & Zimm, 1992). Alguma atividade de *Ddc* é, ainda, encontrada junto ao sistema nervoso central (SNC) de adultos, onde *Ddc* atua na produção dos neuro-transmissores dopamina e serotonina. De acordo com Marsh & Wright (1980) cinco picos de atividade de *Ddc* são evidenciados durante o desenvolvimento *Drosophila*: no final da embriogênese, durante as duas mudas larvais, durante a puparização e durante a eclosão. Desta forma, diferentemente de *Amd*, *Ddc* especializou-se na esclerotização de uma cutícula pigmentada e na produção de neurotransmissores, sendo expresso em estágios posteriores, após uma cutícula inicial já ter sido depositada.

O gene *Ddc* codifica para três diferentes isoformas (A, B e C), que são geradas a partir de sítios de início da transcrição alternativos, bem como por *splicing* alternativo (Fig. 3). As isoformas, entretanto, diferem apenas pelo acréscimo de 35 aminoácidos à extremidade amino-terminal da proteína expressa no SNC. Segundo Lindsley & Zimm (1992), o maior mRNA codificado por *Ddc* é expresso na hipoderme.



acumular-se após cerca de 12 hs de embriogênese, atingindo um máximo por volta de 18 hs, sendo sua transcrição reiniciada no estágio de pupa e logo após a eclosão dos adultos (Eveleth & Marsh, 1986).

Atualmente, duas regiões de extensa homologia podem ser detectadas nas seqüências gênicas de *Amd* e *Ddc*: a primeira corresponde à seqüência encontrada em grande parte do éxon IV de *Amd*, homóloga ao éxon IV de *Ddc*; a segunda refere-se à parte da seqüência 5' UTR do transcrito NM\_057244 de *Amd*, homóloga ao éxon III e a ambas as bordas flanqueadoras do íntron III de *Ddc* (Eveleth & Marsh, 1986)<sup>3</sup>. A posição e a seqüência das regiões intervenientes é, entretanto, altamente divergente. O percentual geral médio de similaridade de seqüências de proteína com homologia estabelecida compartilhadas por estes dois parálogos é 61% em *D. melanogaster* (Wang *et al.*, 1996).

Tatarenkov & Ayala (2001) e Tatarenkov *et al.* (1999a e 2001) têm utilizado seqüências codificadoras dos genes *Amd* e *Ddc* no endereçamento de diversas questões filogenéticas relacionadas ao gênero *Drosophila*, tendo obtido resultados bastante consistentes e congruentes entre si. Dentre as características apresentadas por estes genes, e que os tornam confiáveis para a inferência de relações filogenéticas neste nível de divergência, pode-se destacar sua relativa restrição funcional, contrastada pelos altos níveis de divergência encontrados na terceira posição de seus códons e seu reduzido viés na utilização de códons (Tatarenkov *et al.*, 1999b). Além disso, estes genes apresentam uma distribuição heterogênea de substituições não-sinônimas, de modo que diferentes regiões dos genes podem ser utilizadas de acordo com o problema filogenético em questão. Para *Amd*, por exemplo, a região C-terminal destaca-se como a menos restrita funcionalmente

---

<sup>3</sup> Localização adaptada para nova anotação, conforme Adams *et al.* (2000).

(Tatarenkov *et al.*, 1999b). Outra característica notável de *Amd* e *Ddc* refere-se ao seu elevado ENC (número efetivo de códons) (Tatarenkov *et al.*, 1999b), reduzindo, desta forma, o viés introduzido pelo uso preferencial de códons.

### **O GENE *Hb* (*HUNCHBACK*)**

O gene *Hb* foi inicialmente caracterizado em *D. melanogaster* como um gene *gap* de segmentação (Tautz *et al.*, 1987), quando se demonstrou seu papel chave na padronização espacial anterior durante o desenvolvimento de insetos. Neste caso, um gradiente antero-posterior da proteína *hunchback*, estabelecido a partir de gradientes de genes maternos (*bicoid* e *nanos*, por exemplo), atua na regulação da expressão de genes fundamentais no estabelecimento do processo de segmentação (Wolpert *et al.*, 2000). Para desempenhar esta atividade, *Hb* conta com motivos de ligação ao DNA do tipo ddo de zico, que parecem mediar interações não apenas com ácidos nucléicos, mas também com proteínas (Pinnell *et al.*, 2006).

Pinnell *et al.* (2006) defendem que além de apresentar o domínio *gap* anterior de expressão, *Hb* também é expresso em diversos outros tecidos embrionários e larvais (tais como membranas epiteliais extra-embrionárias e neuroblastos), além de apresentar um domínio materno de expressão, de forma que o mesmo parece desempenhar outros papéis durante o desenvolvimento de protostômios<sup>4</sup>. Entretanto, segundo estes autores, enquanto a expressão nos oócitos e nos tecidos extra-embrionários parece ser plesiomórfica, a função

---

<sup>4</sup> O homólogo de *Hb* em mamíferos – *Ikaros* -, entretanto, parece ser expresso apenas no timo, de forma a atuar como um fator de transcrição específico de células linfóides, mas não apresentando qualquer destas funções.

de gene *gap* é autopomórfica para insetos, enquanto a expressão em neuroblastos é sinapomórfica em insetos e nematódeos. Neste caso, a expressão nos tecidos extra-embrionários parece ser importante para uma correta morfogênese e alongamento do embrião, enquanto que a expressão em neuroblastos parece atuar na padronização do sistema nervoso central. Além disso, Treier *et al.* (1989) já haviam demonstrado que, enquanto os padrões iniciais de expressão de *Hb* em *Drosophila* parecem estar consideravelmente modulados, algumas diferenças espaço-temporais de expressão podem ser detectadas posteriormente entre *D. melanogaster* e *D. virilis*, levantando a hipótese da existência de uma flexibilidade importante para a geração de novos padrões evolutivos.

Segundo Tautz *et al.* (1987) o gene *hb* contém dois diferentes promotores, que levam a dois diferentes transcritos, que codificam, entretanto, para a mesma proteína (Fig. 4). O transcrito P2 (NM\_169233) parece ser regulado por *bicoid*, desempenhando a função *gap* (Treier *et al.*, 1989).

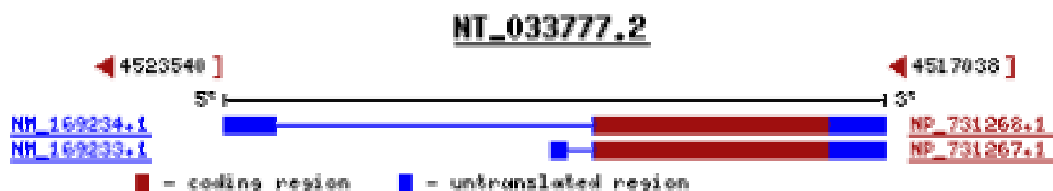


Figura 4 – Mapa genômico do gene *Hb* em *D. melanogaster*, conforme obtido no Entrez Gene, do site NCBI, contendo os números de acesso das seqüências de mRNA (à esquerda, em azul), das seqüências de proteínas (à direita, em vermelho), e da seqüência de nucleotídeos do gene inteiro (em cima, em preto), cuja extensão abrange os números posicionados acima da figura em ambas extremidades. Retângulos representam éxons enquanto que suas linhas conectoras indicam íntrons. Regiões não-traduzidas 5' e 3' estão demarcadas em azul, enquanto a região codificadora encontra-se demarcada em vermelho.

Embora segundo Pinnell *et al.* (2006) o gene *Hb* ancestral possivelmente apresentasse cerca de nove dedos de zinco independentes, estes autores defendem que parece haver uma tendência de perda lenta, mas progressiva, de alguns destes domínios a

partir de hemimetábolos em direção a holometábolos. Ainda segundo eles, todos ortólogos de *hb* seqüenciados até o momento apresentam uma região altamente variável entre os dois domínios dedo de zinco mais conservados. Esta região parece apresentar um perfil de simplicidade crítica (Treier *et al.*, 1989), uma vez que eventos de deriva mutacional causados pelo processo de escorregamento entre fitas parece causar permutações cíclicas do motivo (CAG)<sub>n</sub>, que acabam por levar a inserções e deleções envolvendo principalmente os aminoácidos glicina (CAG), serina (AGC) e alanina (códon GCA).

### **O GENE *PER* (*PERIOD*)**

O gene *period* é considerado um gene de comportamento, uma vez que apresenta efeitos tanto no ritmo circadiano quanto na produção espécie-específica do som de corte (Gleason & Powell, 1997). Muitos autores o consideram, inclusive, um gene de especiação (Coyne, 1992). Recentemente Sakai *et al.* (2004) também demonstraram que o gene *Per* parece, inclusive, estar relacionado à formação da memória de longo prazo em *Drosophila*.

Durante sua atuação no ritmo circadiano, *Per* liga-se à proteína *Tim* (*timeless*), de modo a formar um heterodímero, que tem sua expressão regularmente ativada por um outro heterodímero, constituído pelas proteínas *Clk* (*clock*) e *Cyc* (*cycle*). Através de um processo de *feedback* auto-regulatório, entretanto, os próprios complexos *Per-Tim* acabam por inativar a atividade de *Clk-Cyc* ao final do ciclo. Sinais ambientais tais como luz, temperatura e interações sociais parecem modular o processo. (Kaushik *et al.*, 2007)

Estruturalmente, o gene *Per* apresenta regiões com variados graus de conservação e divergência (Gleason & Powell, 1997). A região mais conservada corresponde, neste caso ao domínio PAS (de *per*, aryl hydrocarbon receptor e *single-minded*), que contando com



dois motivos repetidos de 51 aminoácidos cada, media a dimerização de *Per* com *Tim*. Algumas mutações neste domínio, que abrange o final do éxon 3 até o início do éxon 5 (Fig. 5), resultam na produção de ritmos circadianos alterados em *D. melanogaster* (Baylies *et al.*, 1987 e 1992). Cinco regiões divergentes, nomeadas n1 (encontrada no início do terceiro éxon), n2-n4 (encontrados no quinto éxon) e n5 (encontrada no sexto éxon), são também características do gene *Per* segundo Gleason & Powell (1997). Em *D. melanogaster* a região n2 é a mais variável, contendo repetições dos aminoácidos threonina e glicina (Thr-Gly), cujo comprimento parece estar sob influência de uma seleção térmica, uma vez que o mesmo apresenta uma clina latitudinal (Costa *et al.*, 1991 e 1992). No grupo *willistoni*, por outro lado, n2 encontra-se conservada, sendo n3 a região mais polimórfica e variável entre as espécies (Gleason & Powell, 1997). Este fragmento apresenta em sua região inicial repetições polimórficas do aminoácido glicina, seguidas por uma região polimórfica rica em glicinas. Aparentemente, eventos de deriva mutacional causados por escorregamento entre fitas e *crossing-over* desigual parecem ser a fonte desta variabilidade.



Figura 5 – Mapa genômico do gene *Per* em *D. melanogaster*, conforme obtido no Entrez Gene, do site NCBI, contendo os números de acesso das seqüências de mRNA (à esquerda, em azul), das seqüências de proteínas (à direita, em vermelho), e da seqüência de nucleotídeos do gene inteiro (em cima, em preto), cuja extensão abrange os números posicionados acima da figura em ambas extremidades. Retângulos representam éxons enquanto que suas linhas conectoras indicam íntrons. Regiões não-traduzidas 5' e 3' estão demarcadas em azul, enquanto a região codificadora encontra-se demarcada em vermelho.

Algumas evidências de seleção positiva foram encontradas por Gleason & Powell (1997) ao longo do gene *Per*. Ainda segundo estes autores, este gene apresenta taxas

evolutivas superiores a *Adh*, e um reduzido viés na utilização de códons, ao menos no que se refere às espécies do grupo *willistoni* de *Drosophila*.

#### **OS GENES MITOCONDRIAIS *COI* (CITOCROMO OXIDASE I) E *COII* (CITOCROMO OXIDASE II)**

Genes mitocondriais são freqüentemente utilizados como marcadores em projetos que buscam a realização de inferências acerca das relações filogenéticas ente OTU's relativamente próximas. Isto porque eles apresentam uma série de características potencialmente úteis, tais como:

- são amplamente distribuídos, mas o seu conteúdo gênico é conservado: o genoma mitocondrial animal geralmente contém dois genes de RNA ribossomal (rRNA), 22 genes de RNA transportador (tRNA) e 13 genes codificadores de proteínas (Simon *et al.*, 1994);

- apresentam um elevado número de cópias por célula;

- apresentam, em geral, taxas de mutação superiores às apresentadas por genes nucleares (Satta *et al.*, 1987; Goto and Kimura, 2001);

- apresentam genes com taxas de evolução diferentes (Simon *et al.*, 1994);

- não costumam apresentar íntrons (Saccone *et al.*, 1999) (Figs. 6A e 6B);

- aparentemente não sofrem recombinação (Saccone *et al.*, 1999);

- apresentam geralmente uma herança uniparental (Saccone *et al.*, 1999);

- apresentam uma redução no efeito de *incomplete lineage sorting* em relação aos genes nucleares, por apresentarem um tamanho efetivo populacional menor e, desta forma, um menor tempo de coalescência (Templeton, 2006).

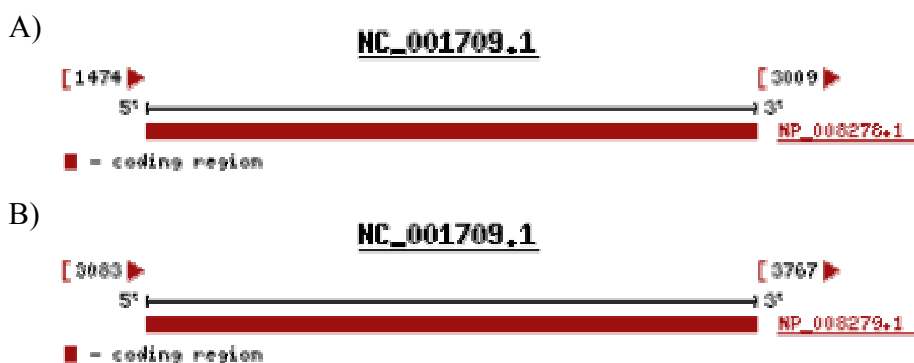


Figura 6 – Mapa genômico dos genes A) *COI*; e B) *COII* em *Drosophila*, conforme obtido no Entrez Gene, do site NCBI, contendo o número de acesso da seqüência de nucleotídeos do genoma mitocondrial de *D. melanogaster* (acima de cada figura) e a localização de *COI* e *COII* dentro deste genoma (à direita e à esquerda de cada figura). Retângulos vermelhos representam a região codificadora.

Os genes *COI* e *COII* codificam para as subunidades I e II do citocromo c, de modo a compor parte desta “proteína transportadora de elétrons que contém um grupamento hemo como grupo prostético” (Stryer, 1996). Desta forma, eles desempenham um papel fundamental na transferência de elétrons durante a cadeia respiratória da respiração celular.

De acordo com Simon *et al.* (1994), enquanto as taxas evolutivas dos genes mitocondriais de rRNA variam consideravelmente ao longo da molécula, genes de tRNA evoluem, em geral, mais lentamente que os genes mitocondriais codificadores. Também entre os genes codificadores, entretanto, pode haver grande variação de taxas evolutivas (Clark *et al.*, 2007). Neste sentido, *COI* é o gene mitocondrial mais conservado em termos de evolução em nível de aminoácidos (Simon *et al.*, 1994), enquanto *COII* apresenta taxas intermediárias de substituição. Entretanto, genes mitocondriais não apresentam sempre um bom desempenho em análises filogenéticas, uma vez que eles podem atingir níveis de saturação rapidamente, possivelmente devido a um sistema de reparo deficiente somado a sua composição nucleotídica altamente enviesada, rica em A+T (Goto & Kimura, 2001).

## Referências Bibliográficas

- Adams MD et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287:2185-2195.
- Avise JC (2004) Molecular markers, natural history, and evolution. 2<sup>nd</sup> ed. Sinauer Associates, Sunderland, 684 pp.
- Baylies MK, Bargiello TA, Jackson FR, Young MW (1987) Changes in abundance or structure of the *per* gene product can alter periodicity of the *Drosophila* clock. *Nature* 326:390-392.
- Baylies MK, Vosshall LB, Sehgal A, Young MW (1992) New short period mutations of the *Drosophila* clock gene *per*. *Neuron* 9:575-581.
- Clark AG, et al. (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450:203-218.
- Coyne JA (1992) Genetics and speciation. *Nature* 355:511-515.
- Eveleth DD, Marsh JL (1986) Evidence for evolutionary duplication of genes in the *dopa decarboxylase* region of *Drosophila*. *Genetics* 114:469-483.
- Gleason JM, Powell JR (1997) Interspecific and intraspecific comparisons of the *period* locus in the *Drosophila willistoni* sibling species. *Mol Biol Evol* 14:741-753.
- Goldberg DA (1980) Isolation and partial characterization of the *Drosophila* alcohol dehydrogenase gene. *Proc Natl Acad Sci USA* 77:5794-5798.
- Goto SG, Kimura MT (2001) Phylogenetic utility of mitochondrial *COI* and nuclear *Gpdh* genes in *Drosophila*. *Mol Phylogenet Evol* 18:404-422.

- Hillis DM, Mable BK, Larson A, Davis SK, Zimmer EA (1996) Nucleic Acids IV: Sequencing and Cloning. In: Hillis DM, Moritz C, Mable BK Molecular Systematics, 2<sup>nd</sup> Ed. Sinauer Associates, Sunderland, pp 321-381.
- Kaushik R, Nawathean P, Busza A, Murad A, Emery P, Rosbash M (2007) PER-TIM interactions with the photoreceptor cryptochrome mediate circadian temperature responses in *Drosophila*. PLOS Biol 5:e146.
- Lindsley DL, Zimm GG (1992) The genome of *Drosophila melanogaster*. Academic Press, Califórnia, 1117 pp.
- Marsh JL, Erfle MP, Leeds CA (1986) Molecular localization, developmental expression and nucleotide sequence of the *alpha-methyl dopa* gene of *Drosophila*. Genetics 114: 453-467.
- Marsh JL, Wright TRF (1980) Developmental relationship between dopa decarboxylase, dopamine acetyltransferase and ecdysone in *Drosophila*. Dev Biol 80:379-387.
- McDonald JH, Kreitman M (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. Nature 351:652-654.
- Nurminsky DI, Moriyama EN, Lozovskaya ER, Hartl DL (1996) Molecular phylogeny and genome evolution in the *Drosophila virilis* species group: duplications of the *Alcohol Dehydrogenase* gene. Mol Biol Evol 13:132-149.
- Pinnell J, Lindeman PS, Colavito S, Lowe C, Savage RM (2006) The divergent roles of the segmentation gene *hunchback*. Integr Comp Biol 46:519-532.
- Russo CAM (2001) Como escolher genes para problemas filogenéticos específicos. In: Mاتيoli SR (ed) Biologia Molecular e Evolução. Holos, Ribeirão Preto, pp 130-136.

- Russo CAM, Takezaki N, Nei M (1995) Molecular phylogeny and divergence times of *Drosophilid* species. *Mol Biol Evol* 12:391-404.
- Saccone C, DeGiorgi C, Gissi C, Pesole G, Reyes A (1999) Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. *Gene* 238:195-209.
- Sakai T, Tamura T, Kitamoto T, Kidokoro Y (2004) A clock gene, *period*, plays a key role in long-term memory formation in *Drosophila*. *Proc Natl Acad Sci USA* 101:16058-16053.
- Satta Y, Ishiwa H, Chigusa SI (1987) Analysis of nucleotide substitutions of mitochondrial DNAs in *Drosophila melanogaster* and its sibling species. *Mol Biol Evol* 4:638-650.
- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Ent Soc of America* 87:651-701.
- Sofer W, Ursprung H (1968) *Drosophila* alcohol dehydrogenase. Purification and partial characterization. *J Biol Chem* 243:3110-3115.
- Solé-Cava AM (2001) Biodiversidade molecular e genética da conservação. In: Matioli SR (ed) *Biologia Molecular e Evolução*. Holos, Ribeirão Preto, pp 171-192.
- Stathakis DG, Pentz ES, Freeman ME, Kullman J, Hankins GR, Pearlson NJ, Wright TRF (1995) The genetic and molecular organization of the *Dopa decarboxylase* gene cluster of *Drosophila melanogaster*. *Genetics* 141:629-655.
- Stryer L (1996) *Bioquímica*. Quarta Edição. Guanabara Koogan, Rio de Janeiro, 1000 pp.
- Tatarenkov A, Ayala FJ (2001) Phylogenetic relationships among species groups of the *virilis-repleta* radiation of *Drosophila*. *Mol Phylogenet Evol* 21:327-331.

- Tatarenkov A, Kwiatowski J, Sharecky D, Barrió E, Ayala FJ (1999a) On the evolution of *Dopa decarboxylase (Ddc)* and *Drosophila* systematics. *J. Mol. Evol.* 48:445-462.
- Tatarenkov A, Sáez AG, Ayala FJ (1999b) A compact gene cluster in *Drosophila*: the unrelated *Cs* gene is compressed between duplicated *Amd* and *Ddc*. *Gene* 231:111-120.
- Tatarenkov A, Zurovcová M, Ayala FJ (2001) *Ddc* and *Amd* sequences resolve phylogenetic relationships of *Drosophila*. *Mol Phylogenet Evol* 20:321-325.
- Tautz D, Lehmann R, Schnürch H, Schuh R, Seifert E, Keinlin A, Jones K, Jäckle H (1987) Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes. *Nature* 327:383-389.
- Templeton AR (2006) *Population Genetics and Microevolutionary Theory*. John Wiley & Sons, New Jersey, 705 pp.
- Treier M, Pfeifle C, Tautz D (1989) Comparison of the gap segmentation gene *hunchback* between *Drosophila melanogaster* and *Drosophila virilis* reveals novel modes of evolutionary change. *EMBO J* 8:1517-1525.
- Wang D, Marsh JL (1995) Developmental regulation of the  $\alpha$ -*Methyl*dopa *Hypersensitive* gene of *Drosophila melanogaster*. *Dev Biol* 168:598-612.
- Wang D, Marsh JL, Ayala FJ (1996) Evolutionary changes in the expression pattern of a developmentally essential gene in three *Drosophila* species. *Proc Natl Acad Sci USA* 93:7103-7107.
- Wolpert L, Beddington R, Brockes J, Jessell T, Lawrence P, Meyerowitz E (2000) *Princípios de Biologia do Desenvolvimento*. Artes Médicas do Sul, Porto Alegre, 484 pp.