UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE GENÉTICA

**Filogenia do complexo *Drosophila buzzatii* (grupo
repleta): inferências de análises multilocus
mitocondriais e nucleares**

RAFAEL FRANSACK FERREIRA

**RIBEIRÃO PRETO
2011**

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE GENÉTICA

**Filogenia do complexo *Drosophila buzzatii* (grupo
repleta): inferências de análises multilocus
mitocondriais e nucleares**

RAFAEL FRANSACK FERREIRA

Dissertação apresentada ao Departamento de Genética da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, como parte das exigências para obtenção do título de Mestre em Ciências Biológicas.

Área de Concentração: Genética.

Orientador: Prof^a. Dr^a. Maura Helena Manfrin

**RIBEIRÃO PRETO
2011**

Autorizo a reprodução total ou parcial deste trabalho para fins de estudo, desde que citada a fonte.

FICHA CATALOGRÁFICA

Fransak, Rafael Ferreira

Filogenia do complexo *Drosophila buzzatii* (grupo *repleta*): inferências de análises multilocus mitocondriais e nucleares. Orientador: Maura Helena Manfrin. Ribeirão Preto, 2011.

137 p.: 32 il.

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, Departamento de Genética.

1. Filogenia. 2. COI. 3. COII. 4. EF-1 α F1 5. *transformer*. 6. *period*. 7. *Drosophila*. 8. tempo de divergência.

Este trabalho foi realizado com auxílio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico, através do processo 134419/2009-0.

Dedico à minha companheira Caroline,
com amor.

AGRADECIMENTOS

Em especial à Prof^ª. Dr^ª. Maura Helena Manfrin (FFCLRP/USP), por ter confiado em meu potencial e me orientado de forma tão prestativa e paciente. Obrigado por todos os ensinamentos, pela motivação e, principalmente, pela amizade.

Ao Prof. Dr. Fábio de Melo Sene (FMRP/USP), por todos os ensinamentos, idéias, discussões e, principalmente, pelas histórias compartilhadas nos cafés do laboratório.

Ao Prof. Dr. Reinaldo O. A. A. de Brito (UFSCar), pela ajuda com o alinhamento das sequências e sugestões.

Ao Prof. Dr. Iderval da Silva Junior Sobrinho (UFSCar), pelas discussões sobre estatística e seleção.

Aos professores da banca examinadora, pela disponibilidade em analisar este trabalho.

Às amigas Carla Bruniera e Dr^ª. Érica Cristina de Carvalho, pela fundamental ajuda nas análises filogenéticas.

Aos amigos do laboratório, Paulo Epifânio, Mateus Henrique, Luís Bizzo, Camila Kokudai, Cintia Graziela, Gislaíne Rodrigues e Taís Lavagnini, por estarem sempre dispostos a me ajudar e por tornarem a convivência no laboratório tão agradável.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa de mestrado. À CAPES, FINEP, FAPESP, CNPq e USP que fornecem verbas para a manutenção do Laboratório de Genética Evolutiva.

Ao Departamento de Genética da FMRP/USP, pela disponibilização de recursos.

À toda a minha família, em especial aos meus pais, Manoel Ferreira Dionísio Jr. e Rosimeiry Ap. Fransak Davanso, ao meu padrasto, Carlos Roberto Davanso, e aos meus avós, Narciso Fransak, Olga Alves Fransak e Elza Peroni, por toda dedicação e por terem sempre me incentivado e investido em minha formação.

À minha querida companheira Caroline dos Santos Rodrigues, pela paciência, carinho, apoio e dedicação.

SUMÁRIO

RESUMO	01
ABSTRACT	02
1. INTRODUÇÃO	03
1.1. Considerações iniciais	03
1.2. O complexo <i>Drosophila buzzati</i>	04
1.3. O cluster <i>Drosophila buzzati</i>	07
1.4. Marcadores Moleculares	14
1.4.1. Citocromo oxidase I e II	15
1.4.2. EF-1 α F1	16
1.4.3. <i>Transaformer</i>	17
1.4.4. <i>Period</i>	21
1.5. Análise individual <i>versus</i> análise combinada	23
2. OBJETIVOS	26
3. MATERIAL E MÉTODOS	27
3.1. Material biológico	27
3.2. Extração de DNA genômico	28
3.3. Amplificação dos genes mitocondriais e nucleares	28
3.4. Eletroforese de verificação.....	29
3.5. Eletroforese para recuperação das bandas de interesse (EF-1 α F1 e <i>period</i>)	30
3.6. Purificação dos fragmentos dos genes COI, COII e <i>transformer</i>	30
3.7. Purificação dos fragmentos dos genes EF-1 α F1 e <i>period</i>	30
3.8. Sequenciamento direto	31
3.9. Análise das sequências de DNA	31
3.10. Reconstrução filogenética	32
3.11. Testes de neutralidade baseados em dn e ds	33
3.12. Tempo de divergência	34
4. RESULTADOS	35
4.1. Citocromo oxidase I	35
4.1.1. Máxima parcimônia	36
4.1.2. Máxima verossimilhança	38
4.1.3. Inferência bayesiana	40
4.2. Citocromo oxidase II	41
4.2.1. Máxima parcimônia	41
4.2.2. Máxima verossimilhança	43

SUMÁRIO

4.2.3. Inferência bayesiana	43
4.3. EF-1 α F1	45
4.3.1. Máxima parcimônia	46
4.3.2. Máxima verossimilhança	47
4.3.3. Inferência bayesiana	48
4.4. <i>Transformer</i>	48
4.4.1. Máxima parcimônia	49
4.4.2. Máxima verossimilhança	51
4.4.3. Inferência bayesiana	51
4.5. <i>Period</i>	52
4.5.1. Máxima parcimônia	53
4.5.2. Máxima verossimilhança	55
4.5.3. Inferência bayesiana	57
4.6. Teste de congruência	58
4.7. Análises combinadas	58
4.7.1. mtDNA	58
4.7.2. nDNA	60
4.7.3. mtDNA + nDNA	61
4.8. Testes de neutralidade	63
4.8.1. <i>Transfromer</i>	63
4.8.2. <i>Period</i>	65
4.9. Tempos de divergência	68
5. DISCUSSÃO	70
5.1. Análise filogenética do complexo <i>Drosophila buzzatii</i>	70
<i>Análises individuais</i>	70
<i>Análises combinadas</i>	76
5.2. Os genes <i>transformer</i> e <i>period</i>	78
5.3. Diversificação e distribuição geográfica do cluster <i>Drosophila buzzatii</i>	80
6. CONCLUSÕES	89
7. REFERÊNCIAS BIBLIOGRÁFICAS	90
APÊNDICE I	110

RESUMO

FRANSACK, R.F. **Filogenia do complexo *Drosophila buzzatii* (grupo *repleta*): inferências de análises multilocus mitocondriais e nucleares.** 2011. Xf. Dissertação de Mestrado. Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo.

O complexo *Drosophila buzzatii* (grupo *repleta*) compreende 13 espécies, divididas em três clusters, de acordo com o bandejamento observado nos cromossomos politênicos: cluster *D. stalker*, incluindo *D. richardsoni* e *D. stalker*, restrito às ilhas do Caribe e Flórida; cluster *D. martensis*, incluindo *D. martensis*, *D. uniseta*, *D. venezolana* e *D. starmeri*, encontrado em áreas desérticas da Colômbia e Venezuela; e cluster *D. buzzatii*, incluindo *D. buzzatii*, *D. koepferae*, *D. antonietae*, *D. serido*, *D. gouveai*, *D. borborema* e *D. seriema*, habitando regiões sazonalmente secas ao longo da diagonal de vegetação aberta da América do Sul. O presente estudo teve como objetivo inferir as relações filogenéticas entre as espécies do complexo *D. buzzatii*, dando ênfase ao cluster *D. buzzatii*, por meio de análises multilocus de genes mitocondriais (COI e COII) e nucleares (EF-1 α F1, *transformer* e *period*). Nas hipóteses filogenéticas estabelecidas, as espécies do complexo *D. buzzatii* constituíram um grupo monofilético, composto por dois subgrupos monofiléticos, os clusters *D. martensis* e *D. buzzatii*, e um parafilético, o cluster *D. stalker*. As relações de parentesco entre as espécies do cluster *D. buzzatii* foram estabelecidas. *Drosophila buzzatii* ocupou a posição mais basal dentro do cluster *D. buzzatii*, estando proximamente relacionada à espécie *D. koepferae*. *Drosophila antonietae* ocupou uma posição intermediária em relação às espécies *D. koepferae* e *D. serido*, que representa o táxon irmão do ramo formado por *D. gouveai*, *D. borborema* e *D. seriema*, com *D. gouveai* ocupando uma posição mais basal em relação às espécies irmãs *D. borborema* e *D. seriema*. Foi detectada seleção purificadora como a principal força dirigindo a evolução dos genes nucleares *transformer* e *period*, para as espécies do complexo *D. buzzatii*. O gene mitocondrial COI, por sua vez, foi utilizado para estimar os tempos de divergência para as espécies do cluster *D. buzzatii*, revelando que o processo de diversificação do grupo iniciou-se no período Plioceno, provavelmente em decorrência de eventos de vicariância associados à elevação dos Andes, sendo também influenciado pelo avanço e retração da vegetação xerófitas, nas flutuações climáticas do Pleistoceno.

Palavras-chave: Filogenia, COI, COII, EF-1 α F1, *transformer*, *period*, *Drosophila*, tempo de divergência.

ABSTRACT

FRANSACK, R.F. **Phylogeny of *Drosophila buzzatii* complex (repleta group): inferences from mitochondrial and nuclear multilocus analysis.** 2011. Xf. Dissertação de Mestrado. Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo.

Drosophila buzzatii complex (*repleta* group) consists of 13 species, divided into three clusters according to the banding seen in polytene chromosomes: *D. stalker*i cluster, including *D. richardsoni* and *D. stalker*i, restricted to the Caribbean Islands and Florida; *D. martensis* cluster, including *D. martensis*, *D. uniseta*, *D. venezuelana* and *D. starker*i, found in desert areas of Colombia and Venezuela, and *D. buzzatii* cluster, including *D. buzzatii*, *D. koepferae*, *D. antonietae*, *D. gouveai*, *D. borborema* and *D. seriema*, that inhabit seasonally dry regions along the open vegetation diagonal in South America. This study aimed to infer the phylogenetic relationships among the *D. buzzatii* species complex, emphasizing the *D. buzzatii* cluster, by multilocus analysis of mitochondrial (COI and COII) and nuclear (EF-1 α F1, *transformer* and *period*) genes. In established phylogenetic hypotheises, the species of the *D. buzzatii* complex formed a monophyletic group, composed of two monophyletic subgroups, the *D. martensis* and *D. buzzatii* clusters, and a paraphyletic one, the *D. stalker*i cluster. The relationships among the *D. buzzatii* species cluster were established. *Drosophila buzzatii* occupied the most basal position within the *D. buzzatii* cluster and is closely related to *D. koepferae*. *D. antonietae* occupied an intermediate position in relation to the *D. koepferae* and *D. serido* species. *D. serido* represents the sister taxon of the branch formed by the *D. gouveai*, *D. borborema* and *D. seriema* species, with *D. gouveai* occupying a basal position in relation to the sister species *D. borborema* and *D. seriema*. It was detected that purifying selection is the main force driving the evolution of *transformer* and *period* nuclear genes for the species of the *D. buzzatii* complex. The divergence time of the *D. buzzatii* species cluster was estimated by the COI gene analysis, revealing that the process of diversification of the group began in the Pliocene period, probably due to vicariant events associated with the uplift phase of the Andes, and it was also influenced by the advance and retraction of xerophytic vegetation in Pleistocene climatic fluctuations.

Key-words: Phylogeny, COI, COII, EF-1 α F1, *transformer*, *period*, *Drosophila*, divergence time.

1. INTRODUÇÃO

1.1. Considerações Iniciais

A organização da informação sobre a diversidade biológica, juntamente com a definição da história das ramificações que deram origem às diferentes espécies que compõe os grupos, é importante, pois pode auxiliar na elaboração de hipóteses sobre o processo evolutivo.

A evolução biológica consiste na mudança das características hereditárias de grupos de organismos ao longo das gerações, sendo que todos os seres, vivos ou extintos, resultam de uma história de descendência com modificação a partir de ancestrais comuns.

Segundo a sistemática filogenética, organismos que compartilham a condição modificada de um determinado caráter (apomorfia) descendem de uma mesma espécie ancestral, na qual essa condição surgiu a partir da modificação de uma condição anterior (plesiomórfica). Neste contexto, cladogramas são dendrogramas que expressam relações filogenéticas entre táxons terminais, evidenciadas por sinapomorfias, ou seja, apomorfias compartilhadas (Amorim, 1997).

A sistemática filogenética tem como objetivo a descrição de padrões hierárquicos de relações entre entidades biológicas, as hipóteses filogenéticas, que nem sempre resultam em árvores evolutivas bem definidas, mas adicionam informações a respeito da história de um grupo.

Hipóteses de relações filogenéticas entre espécies representam um registro indireto dos eventos de especiação e fornecem evidências para a investigação dos processos evolutivos que geram a diversidade biológica.

Além das diferentes abordagens disponíveis para inferências de hipóteses filogenéticas, estudos evolutivos mais acurados e que ajudem no discernimento do padrão de diversificação de um determinado grupo são melhores obtidos com base em conjuntos independentes de caracteres (Amorim, 1997).

Análises utilizando múltiplas fontes de informação têm se multiplicado em estudos filogenéticos. Além de caracteres morfológicos, tradicionalmente usados em análises filogenéticas, diferentes classes de marcadores moleculares, como DNA mitocondrial (mtDNA), DNA de cloroplasto (ctDNA) e DNA nuclear (nDNA), vêm sendo utilizados, pois diferentes partes do genoma, assim como os diferentes genes,

estão sujeitos a pressões seletivas e eventos demográficos distintos, podendo refletir diferentes aspectos da história de uma espécie (Huelsenbeck *et al.* 1996). Tal abordagem é essencial para o entendimento das relações filogenéticas, especialmente entre espécies de divergência recente, nas quais existe pouca divergência morfológica e genética, possibilitando a análise dos processos microevolutivos pelos quais a diversidade biológica é gerada em eventos de especiação (Machado & Hey, 2003).

1.2. O complexo *Drosophila buzzatii*

O gênero *Drosophila* é dividido hierarquicamente em “subgêneros”, “grupos”, “subgrupos”, “complexos” e “clusters” de espécies. Embora não reconhecida pelo Código Internacional de Nomenclatura Zoológica, essa classificação é amplamente utilizada em estudos envolvendo espécies desse gênero (Bächli, 2010).

O complexo *D. buzzatii* é um dos 10 complexos que compõem o subgrupo *mulleri*, grupo *repleta*, subgênero *Drosophila* (Wasserman, 1982). Compreende 13 espécies, identificadas pela morfologia do aparelho reprodutor masculino e divididas em três clusters, de acordo com o bandejamento observado nos cromossomos politênicos (Wasserman 1982; Ruiz & Wasserman 1993; Vilela, 1983): cluster *Drosophila stalker*, formado pelas espécies *Drosophila richardsoni* Vilela, 1983 e *Drosophila stalker* Wheeler, 1954; cluster *Drosophila martensis*, compreendendo as espécies *Drosophila uniseta* Wasserman *et al.* 1973, *Drosophila venezolana* Wasserman, Fontdevila & Ruiz, 1983, *Drosophila starmeri* Wasserman *et al.* 1973 e *Drosophila martensis* Wasserman & Wilson, 1957; e cluster *buzzatii*, agrupando as espécies *Drosophila buzzati* Patterson & Wheeler, 1942, *Drosophila koepferae* Fontdevila & Wasserman, 1988, *Drosophila antonietae* Tidon-Sklorz & Sene, 2001, *Drosophila serido* Vilela & Sene, 1977, *Drosophila gouveai* Tidon-Sklorz & Sene, 2001, *Drosophila borborema* Vilela & Sene, 1977 e *Drosophila seriema* Tidon-Sklorz & Sene, 1995.

Em relação à distribuição geográfica conhecida do complexo *D. buzzatii*, as espécies do cluster *D. stalker* encontram-se restritas às ilhas do Caribe e Flórida, as espécies do cluster *D. martensis* são encontradas em áreas desérticas da Colômbia e Venezuela, e as espécies do cluster *D. buzzatii* habitam regiões sazonalmente secas ao longo da diagonal de vegetação aberta da América do Sul, entre as florestas Amazônica e Atlântica (Vilela, 1983; Rodriguez-Trelles *et al.* 2000; Manfrin & Sene, 2006).

As relações de parentesco entre as espécies do complexo *D. buzzatii* foram estabelecidas por meio de estudos citológicos (Wasserman, 1982; Ruiz & Wasserman, 1993) e moleculares (Spicer, 1995; Rodriguez-Trelles *et al.* 2000; Durando *et al.* 2000), cujos resultados estão representados na Figura 1.

Os dados cromossômicos (Wasserman 1982; Ruiz & Wasserman, 1993) mostram o compartilhamento das inversões 2m e 2n entre todas as espécies dos três clusters. As espécies dos clusters *D. buzzatii* e *D. martensis* compartilham a inversão 2z⁷, e o último, com exceção da espécie *D. martensis*, apresenta as inversões fixas 2e² e 3v. Estes dados permitem inferir apenas a relação de proximidade entre os clusters *D. martensis* e *D. buzzatii*, mas não fornecem uma indicação da direção evolutiva das mudanças cromossômicas verificadas (Ruiz & Wasserman, 1993; Rodriguez-Trelles *et al.* 2000). Além das inversões citadas anteriormente, outras inversões fixas não filogeneticamente informativas são observadas nas 13 espécies que compõem o complexo *D. buzzatii* (Figura 1A).

A análise do bandejamento cromossômico ainda sugere que o cromossomo 2 de *D. stalker*i seja derivado do arranjo cromossômico designado como Primitivo I (Wasserman, 1960, 1982), a forma ancestral do grupo *repleta*, pelo menor número de inversões (apenas duas) em relação às demais espécies do complexo, o que pode ser considerado um forte indício de que o cluster *D. stalker*i seja mais basal dentro do complexo *D. buzzatii* (Ruiz & Wasserman, 1993; Rodriguez-Trelles *et al.* 2000).

Embora resultados de estudos de hibridação e análise da viabilidade de F1 (Patterson & Alexander, 1952; Ruiz & Wasserman 1993) contrastem, a princípio, com a hipótese proposta acima, indicando que cluster *D. buzzatii* é mais passível de intercruzamento com representantes do complexo *D. mulleri* do que o cluster *D. stalker*i, tal fato pode ser explicado pela existência de regiões em simpatria de *D. stalker*i e *D. richardsoni* com espécies do complexo *D. mulleri* (Wasserman & Wasserman, 1992), o que teria possibilitado um reforço na seleção de caracteres relacionados ao isolamento reprodutivo (Ruiz & Wasserman, 1993).

A primeira hipótese filogenética proposta para as espécies do complexo *D. buzzatii*, com base em sequências gênicas, utilizou as subunidades I, II e III do gene mitocondrial da citocromo oxidase (Spicer, 1995). Esta hipótese sugere a monofilia do complexo *D. buzzatii*, o posicionamento do cluster *D. stalker*i como o mais basal, e define algumas relações intra-cluster (Figura 1B). Para o cluster *D. martensis*, *D. venezolana* foi posicionada como espécie irmã de *D. starkeri*, e *D. martensis* foi

alocada como mais basal em relação ao clado formado pelas espécies anteriores. *Drosophila buzzatii* aparece como a espécie mais basal, entre as analisadas do cluster *D. buzzatii*.

Na hipótese filogenética proposta a partir da análise de sequências do gene *Xanthine dehidrogenase* (Rodriguez-Trelles *et al.* 2000), o monofiletismo do complexo *D. buzzatii* foi corroborado, além disso, os clusters *D. martensis* e *D. buzzatii* foram sugeridos como grupos monofiléticos (Figura 1C). O cluster *D. stalker* foi posicionado como o mais basal, com *D. richardsoni* representando a espécie mais ancestral dentro do cluster. *Drosophila venezolana* e *D. starmeri* foram agrupadas como espécies irmãs, constituindo a linhagem evolutiva mais derivada dentro do cluster *D. martensis*. Ainda em relação a este cluster, *D. martensis* foi posicionada como a espécie mais basal e *D. uniseta* ocupou uma posição filogenética intermediária. Para o cluster *D. buzzatii*, *D. serido* e *D. borborema* foram agrupadas como espécies irmãs, constituindo uma linhagem evolutiva intimamente relacionada à *D. koepferae*. *Drosophila buzzatii* foi posicionada como a espécie mais basal dentro do cluster.

Uma terceira hipótese filogenética para o complexo é proveniente de um estudo com espécies do grupo *repleta*, com base em sequências de várias partições gênicas: 16S, citocromo oxidase II, *nitrogen dehidrogenase 1* e o gene nuclear *hunchback* (Durando *et al.* 2000). Além do monofiletismo do complexo *D. buzzatii*, uma relação de monofilia recíproca entre os três clusters foi sugerida. Algumas incongruências são observadas em relação à hipótese filogenética proposta por Rodriguez-Trelles *et al.* (2000), *D. uniseta* foi posicionada como a espécie mais basal do cluster *D. martensis* e o cluster *D. buzzatii* foi dividido em duas linhagens evolutivas, uma representada pelo clado formado pelas espécies *D. buzzatii* e *D. koepferae*, e outra pelo clado formado pelas espécies *D. serido* e *D. borborema* (Figura 1D).

Análises filogenéticas utilizando sequências multilocus mitocondriais e nucleares de novas partições gênicas com um número representativo de espécies de cada cluster poderiam adicionar novas informações acerca das relações dentro do complexo *D. buzzatii* e, possivelmente, ajudar a esclarecer as incongruências observadas nos estudos abordados anteriormente.

Figura 1 – Relações de parentesco entre as espécies do complexo *D. buzzatii*. **A.** Hipótese baseada em dados de inversões cromossômicas (adaptada de Ruiz & Wasserman, 1993). **B.** Hipótese baseada em sequências dos genes da citocromo oxidase I, II e III (adaptada de Spicer, 1995). **C.** Hipótese baseada em sequências do gene nuclear *Xanthine dehydrogenase* (adaptada de Rodriguez-Trelles *et al.* 2000). **D.** Hipótese baseada em sequências de DNA ribossomal 16S, dos genes mitocondriais citocromo oxidase II e *nitrogen dehydrogenase 1*, e do gene nuclear *hunchback* (adaptada de Durando *et al.* 2000).

1.3. O cluster *Drosophila buzzatii*

As sete espécies nominais que compõem o cluster *Drosophila buzzatii*, *D. buzzatii*, *D. koepferae*, *D. antonietae*, *D. serido*, *D. gouveai*, *D. borborema* e *D. seriema*, são morfologicamente semelhantes (Vilela, 1983) e naturalmente endêmicas da América do Sul (Manfrin & Sene, 2006). Dentre estas, apenas *D. buzzatii* tornou-se cosmopolita, sendo encontrada em várias regiões do mundo, como Europa, África e Oceania (Fontdevila *et al.* 1981; Barker *et al.* 1985). Essa colonização ocorreu nos últimos 200 anos, devido ao transporte de seu cacto hospedeiro da América do Sul para as demais regiões por ação antrópica (Barker *et al.* 1985). A distribuição geográfica atual das espécies do cluster *D. buzzatii* na América do Sul está mostrada na Figura 2.

Na natureza, o desenvolvimento larval das espécies do cluster *D. buzzatii* ocorre exclusivamente em tecidos de cactos em decomposição (Pereira *et al.* 1983). Esta especificidade ecológica tem pelo menos três consequências, tornando estas espécies um excelente material biológico para estudos evolutivos (Manfrin & Sene, 2006):

Figura 2 – Distribuição geográfica conhecida das espécies do cluster *D. buzzatii* na América do Sul. As principais bacias hidrográficas, bem como a distribuição aproximada das principais áreas de florestas secas da América do Sul, estão apresentadas no mapa. FTSS: Florestas Tropicais Sazonalmente Secas (adaptado de Franco, 2009).

1- A distribuição geográfica das espécies do cluster *D. buzzatii* é determinada pela ocorrência de espécies de cactáceas na América do Sul. Devido a esta associação, elas ocorrem em áreas de vegetação xerófita ao longo da diagonal de vegetação aberta desse continente, entre as florestas Amazônica e Atlântica, em áreas de Florestas Tropicais

Sazonalmente Secas (FTSS), no Domínio do Chaco, e em redutos de cactáceas em regiões de Cerrado, Mata Atlântica e costa Atlântica.

No Brasil, o Domínio da Caatinga é uma região de endemismo para várias espécies de cactáceas, com grande diversidade de espécies e grande densidade populacional. Fora da Caatinga, as cactáceas são encontradas em áreas de solo adequado para seu desenvolvimento, tais como afloramentos rochosos e dunas (Taylor & Zappi, 2004). Nestas regiões, as populações e espécies do cluster *D. buzzatii* possuem distribuição fragmentada, o que pode ter propiciado condições para a diferenciação entre elas;

2- A exploração de diferentes cactos hospedeiros pode estar associada ao desenvolvimento de características adaptativas ao novo recurso, uma vez que os diferentes cactos apresentam composições químicas e microflora distintos (Starmer *et al.* 1986), produzindo ambientes ecológicos distintos para as espécies cactófilas de *Drosophila* (Matzkin, 2008).

3- Eventos demográficos de expansão populacional e os processos de diferenciação dentro do cluster podem estar relacionados a eventos não sincrônicos de expansão e retração da vegetação xerófita na América do Sul, durante alterações climáticas causadas por eventos paleoclimáticos (Ab' Saber, 1977; Sene *et al.* 1988; de Brito *et al.* 2002a; Moraes *et al.* 2009).

Sendo assim, o estudo das espécies do cluster *D. buzzatii* pode contribuir para o debate sobre o papel das flutuações climáticas do Quaternário em moldar a diversidade genética de táxons neotropicais.

A morfologia do edeago é considerada a principal característica diagnóstica para a identificação das espécies do grupo *D. repleta* (Vilela, 1983). Em relação a esse marcador, as espécies do cluster *D. buzzatii* podem ser divididas em dois grupos. Um deles é formado pelas populações de *D. buzzatii*, o outro grupo, denominado *D. serido* “sibling set” (Tidon-Sklorz & Sene, 2001), é formado pelas demais espécies do cluster *D. buzzatii*, as quais compartilham três inflexões pontiagudas no edeago (Figura 3A), resultando numa aparência falciforme na extremidade distal do edeago. Dentre as espécies do *D. serido* “sibling set”, *D. borborema* possui a forma do edeago facilmente distinguível das demais, que são discriminadas somente através de análises quantitativas (Silva & Sene, 1991; Tidon-Sklorz & Sene, 1995a; Tidon-Sklorz & Sene, 2001; Franco *et al.* 2006a). Uma análise morfométrica das asas das espécies do cluster *D. buzzatii*

também subdivide o cluster em *D. buzzatii* e *D. serido* “sibling set” (Moraes *et al.* 2004, Figura 3B).

Figura 3 – Relações de similaridade morfológica estabelecidas para as espécies do cluster *D. buzzatii*. **A.** Relações baseadas na morfologia do edeago (Tidon-Sklorz & Sene, 2001). As setas indicam as inflexões pontiagudas do edeago, características das espécies do *D. serido* “sibling set”. **B.** Relações fenéticas baseadas na morfologia da asa (Moraes *et al.* 2004).

Graus de isolamento reprodutivo pós-zigóticos entre as espécies do cluster *D. buzzatii* foram obtidos através de estudos experimentais, envolvendo o cruzamento entre diferentes linhagens evolutivas desse grupo (Bizzo, 1982; Marin *et al.* 1993; Madi-Ravazzi *et al.* 1997; Machado *et al.* 2002; Machado *et al.* 2006). A principal conclusão desses estudos é que existem diferentes níveis de isolamento reprodutivo entre as espécies do cluster *D. buzzatii*, com nenhum caso de isolamento reprodutivo completo. Por exemplo, Machado *et al.* (2006) realizaram uma série de cruzamentos entre estoques representativos de populações alopátricas de espécies do cluster *D. buzzatii*, mostrando que cerca de 50% dos cruzamentos interespecíficos resultaram em uma prole F1 fértil. Dentre os cruzamentos com F1 estéril, a maioria envolvia cruzamentos com a espécie *D. buzzatii*, indicando que essa é a espécie mais isolada reprodutivamente do cluster *D. buzzatii*, em congruência com resultados de estudos anteriores (Maddi-Ravazzi *et al.* 1997).

Além dos estudos empíricos, têm sido descritas evidências de cruzamentos interespecíficos, na natureza, envolvendo alguns pares de espécies do cluster *D. buzzatii* (Manfrin *et al.* 2001; de Brito *et al.* 2002a; Gomes & Hasson, 2003; Piccinali *et al.* 2004).

No limite sul da distribuição de *D. gouveai*, três de suas populações (Analândia/SP, Rifaina/SP e Brotas/SP) apresentam haplótipos mitocondriais

característicos de *D. antonietae* (Manfrin *et al.* 2001; de Brito *et al.* 2002a). Pelo fato dessas populações estarem localizadas próximas ao limite norte da distribuição de *D. antonietae*, foi levantada a hipótese de um evento de introgressão unidirecional de genes citoplasmáticos de *D. antonietae* nestas populações de *D. gouveai*, após um contato secundário entre essas espécies, que atualmente estão em alopatria (Manfrin *et al.* 2001; de Brito *et al.* 2002a).

No estado brasileiro de Santa Catarina foi descrita uma área de contato secundário entre as espécies *D. serido* e *D. antonietae* (Manfrin & Sene, 2006). Um extenso conjunto de dados, com análises de haplótipos mitocondriais e características morfológicas, sugere que ocorreram eventos de hibridação entre essas espécies com introgressão de haplótipos mitocondriais, uma vez que alguns indivíduos com enzimas, edeagos e asas, com o padrão típico de *D. serido*, apresentaram haplótipos de *D. antonietae*, embora essas espécies provavelmente não estejam cruzando atualmente nesta área de contato (Kokudai *et al.* no prelo).

As evidências de hibridação introgressiva envolvendo espécies do cluster *D. buzzatii* não estão restritas a genes mitocondriais. De acordo com dados dos genes nucleares Esterase-5 e *Xhantine dehidrogenase*, a melhor explicação para a grande quantidade de polimorfismos compartilhados, entre algumas populações de *D. buzzatii* e *D. koepferae*, é hibridação introgressiva, uma vez que o tempo de divergência entre essas espécies deve ter sido suficiente para perda da maioria dos polimorfismos compartilhados (Gomes & Hasson, 2003; Piccinali *et al.* 2004).

Em *Drosophila* algumas das características que podem levar ao isolamento reprodutivo pré-copulatório são a corte sonora, produzida pelos machos durante o cortejo das fêmeas, através do batimento coordenado de suas asas (Kyricacou & Hall, 1980; Ewing, 1983), e a composição dos hidrocarbonetos epicuticulares (Jallon & David, 1987; Coyne *et al.* 1994; Coyne & Oyama, 1995; Ferveur *et al.* 1996; Etges & Ahrens, 2001; Etges & Jackson 2001; Higgie & Blows 2007; Oliveira *et al.* em preparação). Ambos são caracteres espécie específico em várias espécies de *Drosophila*, podendo ser componentes utilizados pela fêmea para aceitar machos co-específicos (Etges & Jackson 2001; Etges, 2002; Oliveira *et al.* em preparação).

A duração média da corte sonora é significativamente diferente entre as espécies do cluster *D. buzzatii* (Machado *et al.* 2002). Oliveira (2008) analisou diversas características do padrão acústico da corte sonora, mostrando que esta é bastante distinta entre as espécies do cluster.

A análise da composição dos hidrocarbonetos epicuticulares revelou diferenças quantitativas significativas entre as espécies do cluster *D. buzzatii* (Oliveira *et al.* em preparação). Entretanto, experimentos de transferência de hidrocarbonetos epicuticulares entre pares de espécies do cluster não resultaram em alterações no padrão de acasalamento, sendo necessários outros testes para determinar se tais caracteres possuem atividade feromonal para o grupo (Oliveira *et al.* em preparação).

Hipóteses de relações filogenéticas entre as espécies do cluster *D. buzzatii* foram estabelecidas a partir de dados de inversões cromossômicas fixas (Ruiz & Wasserman 1993), do gene nuclear *Xhantine dehidrogenase* (Rodriguez-Trelles *et al.* 2000), ambas já abordadas anteriormente, do gene mitocondrial Citocromo Oxidase I (Manfrin *et al.* 2001) e do gene nuclear *period* (Franco *et al.* 2010), apresentadas na Figura 4. Embora as filogenias obtidas não sejam concordantes quanto às suas topologias, em todas foi inferida a origem monofilética do cluster *D. buzzatii*.

A hipótese filogenética baseada no gene mitocondrial COI sugere que as espécies do cluster *D. buzzatii* são divididas em três linhagens evolutivas principais (Manfrin *et al.* 2001). A mais basal é representada pelo clado formado pelas espécies *D. buzzatii* e *D. koepferae*, a segunda é representada pelo clado formado pela espécie *D. antonietae*, e a terceira linhagem evolutiva agrupa em uma politomia as espécies *D. gouveai*, *D. borborema*, *D. seriema* e *D. serido* (Figura 4C). O clado formado pelas espécies da terceira linhagem evolutiva foi denominado clado AB em de Brito *et al.* (2002a), em alusão aos morfótipos A e B dos edeagos, característicos das espécies *D. serido* e *D. gouveai*, respectivamente (Silva & Sene, 1991). Entre as espécies da segunda e terceira linhagens, *D. serido* e *D. antonietae* compartilham a inversão cromossômica fixa $2x^7$, em desacordo com a filogenia molecular. Por outro lado, as espécies *D. gouveai*, *D. seriema* e *D. borborema* compartilham a inversão cromossômica fixa $2e^8$, concordando com a filogenia molecular (Figura 6a; Tosi & Sene, 1989; Ruiz & Wasserman, 1993).

Na hipótese filogenética estabelecida a partir de fragmentos do gene *period* (Franco *et al.* 2010), as sequências de *D. koepferae*, provenientes de indivíduos de duas localidades no Chaco argentino, foram divididas em duas linhagens evolutivas, denominadas pelo autor como koep1 e koep2. A primeira foi agrupada em um mesmo clado com as sequências de *D. buzzatii*, e a segunda foi alocada na base do clado formado pelas demais espécies do cluster (Figura 4D). As sequências de *D. antonietae* e *D. serido* formaram clados robustos, no entanto, contrastando com a hipótese filogenética baseada no gene mitocondrial COI (Manfrin *et al.* 2001), *D. antonietae* foi

posicionada de forma mais relacionada filogeneticamente às espécies *D. gouveai*, *D. borborema* e *D. seriema*, e *D. serido* mais intimamente relacionada às espécies *D. buzzatii* e *D. koepferae*. *Drosophila gouveai*, *D. borborema* e *D. seriema* também formaram um clado robusto, as duas últimas sugeridas como espécies irmãs, ocupando a posição mais derivada do clado, porém, sem monofilia recíproca (Figura 4D).

Figura 4 – Relações de parentesco entre as espécies do cluster *D. buzzatii*. **A.** Hipótese baseada em dados de inversões cromossômicas (adaptada de Ruiz & Wasserman, 1993). **B.** Hipótese baseada em sequências do gene nuclear *Xanthine dehydrogenase* (adaptada de Rodriguez-Trelles *et al.* 2000). **C.** Hipótese baseada em haplótipos do gene mitocondrial COI (adaptada de Manfrin *et al.* 2001). **D.** Hipótese baseada em sequências do gene *period* (adaptada de Franco *et al.* 2010).

Os padrões contrastantes, encontrados quando diferentes marcadores são utilizados para estudar as relações filogenéticas entre as espécies do cluster *D. buzzatii*, parecem ser uma consequência das dificuldades intrínsecas à reconstrução filogenética de espécies filogeneticamente próximas (Machado & Hey, 2003). As espécies do cluster *D. buzzatii* apresentam um padrão complexo de divergência populacional (Sene *et al.* 1988), com ocorrência de hibridação introgressiva entre diferentes linhagens dentro do cluster (Manfrin *et al.* 2001; de Brito *et al.* 2002a; Gomes & Hasson, 2003; Piccinalli *et al.* 2004; Manfrin & Sene, 2006). Além disso, algumas espécies dentro desse grupo foram submetidas a eventos de expansão populacional, ao longo de sua história evolutiva (de Brito *et al.* 2002a; Piccinalli *et al.* 2007; Franco, 2010), o que pode ter dificultado a perda de polimorfismos compartilhados entre diferentes linhagens dentro do cluster *D. buzzatii*, pela manutenção do tamanho populacional grande, impedindo a aquisição de monofilia recíproca.

Datações moleculares sugerem que o início da divergência das espécies do cluster *D. buzzatii* ocorreu no Período Terciário. A separação do clado composto por *D. buzzatii* e *D. koepferae* teria ocorrido há cerca de 6-12 milhões de anos, enquanto a separação do clado composto por *D. antonietae* (Clado D) do clado formado pelas espécies do Clado AB teria ocorrido entre 3-6 milhões de anos atrás (Figura 4C; Manfrin *et al.* 2001). Essas datações contradizem a hipótese de que todas as espécies do cluster *D. buzzatii*, exceto *D. buzzatii*, teriam se diferenciado através do isolamento das populações durante os ciclos recorrentes de expansão/retração da vegetação cactácea durante eventos paleoclimáticos, que ocorreram na América do Sul durante o Período Quaternário (Sene *et al.* 1988). Porém, como o clado composto pelas espécies *D. serido*, *D. gouveai*, *D. seriema* e *D. borborema* provavelmente iniciou sua divergência há menos de três milhões de anos (Manfrin *et al.* 2001), próximo ao limite máximo do Pleistoceno, não é possível descartar a hipótese de que as linhagens mais recentes desse clado tenham divergido durante esse período.

Franco (2009) realizou estimativas de tempo para o ancestral comum mais recente (T_{ACMR}) das espécies do clado AB, a partir de análises com o gene COI, baseadas na teoria da coalescência. Segundo tais análises, o T_{ACMR} para as espécies do clado AB é cerca de 843 mil anos, 761 mil anos para a espécie *D. serido*, 697 mil anos para *D. gouveai*, 824 mil anos para *D. seriema* e 199 mil anos para *D. borborema*.

1.4. Marcadores moleculares

Análises usando múltiplas fontes de informação têm se multiplicado em estudos de evolução. Uma das razões para este fato é que quanto mais aprendemos sobre o processo evolutivo, torna-se evidente que diferentes conjuntos de dados não são equivalentes (Huelsenbeck *et al.* 1996; Avise, 2009). Considerando sequências de DNA, vários fatores são responsáveis por este fato. Entre eles, podemos citar: diferentes taxas de evolução entre diferentes sequências de DNA, genes e códons; transferência horizontal de genes; polimorfismo ancestral; paralogia e convergência. Todos esses fatores podem colaborar para observação de sinais filogenéticos conflitantes (Huelsenbeck *et al.* 1996; Machado & Hey, 2003). Dessa forma, informações mais acuradas, que ajudem no discernimento dos processos envolvidos na evolução de um determinado grupo de espécies, podem ser obtidas sobre conjuntos independentes de

dados, o que pode fortalecer, acrescentar ou definir as relações, eventos e processos relacionados à diversificação de um determinado grupo.

Basicamente duas classes de sequências de DNA têm sido utilizadas nos estudos evolutivos: sequências mitocondriais e nucleares. Estas duas classes de sequências possuem aspectos biológicos distintos e a integração das informações obtidas a partir de cada uma, sobre um mesmo conjunto de dados, é importante para o estabelecimento de hipóteses consistentes (Avice & Wollenberg, 1997; Ballard & Whitlock, 2004; Avice, 2009). Com relação aos genes nucleares, tem se destacado como marcadores moleculares genes potencialmente relacionados ao isolamento reprodutivo, os denominados “genes da especiação” (Coyne, 1992; Ting *et al.* 2000), uma vez que esses genes devem recuperar melhor a filogenia de um grupo do que genes não relacionados ao isolamento reprodutivo, pois geralmente são menos afetados por polimorfismos ancestrais e/ou introgressão (Ting *et al.* 2000).

1.4 .1. Citocromo oxidase I e II

O DNA mitocondrial apresenta muitas características favoráveis para sua utilização como ferramenta em estudos evolutivos (revisado em Avice, 2009). Por exemplo, as principais fontes de variação do DNA mitocondrial são mutações de ponto. Além disso, existe uma heterogeneidade nas taxas de evolução do DNA mitocondrial, permitindo que diferentes regiões da molécula possam ser utilizadas para analisar diferenciação em diferentes níveis taxonômicos (Simon *et al.* 1994; Li, 1997; Avice, 2009).

A molécula de DNA mitocondrial, geralmente, representa um caráter não recombinante, transmitido assexuadamente através da fêmea. Assim, os indivíduos retêm na sequência do DNA mitocondrial a história matrilinea de organismos co-específicos, o que faz dessa molécula um importante marcador para o estabelecimento de genealogias de genes e para o estudo filogeográfico (Avice, 2000, 2009).

Além disso, o DNA mitocondrial apresenta um tamanho efetivo populacional menor em relação aos genes nucleares. De acordo com Templeton (2006), o tempo de coalescência é dado por $2xNe$, em que x é o nível de ploidia e Ne é o tamanho efetivo populacional. Desse modo, para genes autossômicos ($x = 2$) e ligados ao cromossomo X ($x = 1,5$), o tempo de coalescência é $4Ne$ e $3Ne$, respectivamente. Para o DNA mitocondrial ($x = 1$), o tempo de coalescência é $2Ne/2$ ou simplesmente Ne (tamanho

efetivo populacional das fêmeas), pois essa molécula representa uma herança matrilinea. Dessa maneira, os tempos de coalescência para os genes mitocondriais são cerca de três e quatro vezes menores do que para os genes ligados ao X e autossômicos, respectivamente (Templeton, 2005; Templeton, 2006). Por esse motivo, é esperado menos efeito de *lineage sorting* incompleto em genes mitocondriais em relação aos genes nucleares, o que pode ser uma vantagem em estudos evolutivos com espécies de divergência recente.

A baixa fidelidade na replicação do DNA, em consequência da ausência de um mecanismo de reparo eficiente, também contribui para altas taxas evolutivas observadas para o mtDNA (Li, 1997).

Os genes mitocondriais Citocromo Oxidase I (COI) e Citocromo Oxidase II (COII) codificam as subunidades I e II do citocromo C, que é um componente fundamental na cadeia respiratória (Simon *et al.* 1994).

Estes genes são os mais frequentemente utilizados em estudos filogenéticos (O'Grady *et al.* 1998; Nagaraja *et al.* 2004; Silva-Bernardi *et al.* 2006) e populacionais (Hurtado *et al.* 2004; Pfeiler *et al.* 2007; Knowles *et al.* 2008; Duda Jr & Lee, 2009), sendo o gene COI, inclusive, considerado um dos genes relevantes para identificação de animais pelo chamado código de barras genético (Hebert *et al.* 2003). Além da utilização dos genes mitocondriais para a elaboração de hipóteses filogenéticas para o complexo *D. buzzatii* (Spicer, 1995, Durando *et al.* 2000) e para o cluster *D. buzzatii* (Manfrin *et al.* 2001), tais genes têm sido utilizados para o estabelecimento de hipóteses filogeográficas para as espécies desse cluster (de Brito *et al.* 2002a; de Brito *et al.* 2002b; Morales, 2005; Franco, 2009).

1.4.2. EF-1 α F1

No processo de tradução da informação genética do mRNA para a proteína, a elongação da cadeia de aminoácidos evolve uma série de complexos protéicos, que foram denominados fator de elongação 1 (EF-1) e fator de elongação 2 (EF-2). O complexo citoplasmático EF-1 é formado por três proteínas: EF-1 α , EF-1 β e EF-1 γ . A subunidade EF-1 α catalisa a ligação GTP dependente do aminoacil-tRNA ao sítio aceptor do ribossomo. Devido à importância fundamental para o metabolismo das células eucarióticas, o gene codificador da proteína EF-1 α é evolutivamente conservado (Hovemann *et al.* 1988).

Em *Drosophila melanogaster*, o gene EF-1 α ocorre como duas cópias parálogas, EF-1 α F1 e EF-1 α F2, distintas em relação à posição dos íntrons e composição de nucleotídeos. Essas duas cópias são expressas em diferentes fases do desenvolvimento, a proteína EF-1 α F1 está presente em todos os estágios do desenvolvimento, enquanto EF-1 α F2 ocorre em fases específicas. Os dados de expressão gênica sugerem que a separação genética de F1 e F2 é um evento antigo (Hovemann *et al.* 1988).

Cópias parálogas do gene EF-1 α também foram identificadas e caracterizadas em representantes de *Apis mellifera* (Hymenoptera, Apidae). Essas cópias do gene EF-1 α , semelhante ao verificado em *Drosophila*, diferem em composição das sequências nucleotídicas e posição dos íntrons; a localização dos íntrons da cópia F2 não corresponde à posição de nenhum dos íntrons da cópia F1 (Danforth & Ji, 1998).

A existência de duas cópias parálogas de EF-1 α em *Drosophila* e *Apis* sugere que elas podem ter surgido no ancestral dos insetos holometábolos (Danforth & Ji, 1998).

O gene EF-1 α , por ser altamente conservado, tem sido frequentemente utilizado em análises filogenéticas envolvendo grupos de divergência antiga (Brower & DeSalle, 1994). No entanto, genes conservados evolutivamente também podem ser filogeneticamente informativos em estudos envolvendo grupos de divergência recente, porque podem apresentar taxas de substituições sinônimas relativamente altas (Cho *et al.* 1995).

1.4 .3. *Transformer*

O desenvolvimento do fenótipo sexual em *Drosophila* é mediado por uma série de genes que convertem a relação cromossomo X / autossomo em uma célula de macho ou de fêmea. Quando a relação é 1 (i.e., quando existem dois cromossomos X por célula diplóide) o embrião se desenvolve em uma mosca fêmea. Por outro lado, quando a relação é de 0,5 (i.e., quando existe apenas um cromossomo X por célula diplóide) o embrião se desenvolve em um macho (Gilbert, 2003).

O gene *transformer* (*tra*) é um dos genes que controla a diferenciação sexual somática em *Drosophila*. É fundamental para a formação de fêmeas, e sua perda resulta em moscas macho, independentemente da relação cromossômica. Durante todo o período larval, o gene *tra* é transcrito ativamente em um pré-mRNA que é processado em um mRNA geral ou em um mRNA específico de fêmeas. O mRNA geral, encontrado tanto em machos como em fêmeas, contém um códon de parada precoce

(UGA) no segundo éxon, responsável pela tradução de uma proteína truncada não funcional. Portanto, o transcrito geral do gene *tra* não está relacionado com a determinação do sexo (Figura 2) (O'Neil & Belote, 1992; Gilbert, 2003).

Devido à atuação do produto do gene *Sex-lethal* (*Sxl*), um *splicing* alternativo sexo-específico do pré-mRNA do gene *tra* ocorre em fêmeas. Se a relação X / autossomo for 1, a proteína SXL funcional será produzida, e em embriões ou larvas XY a proteína SXL não é funcional. A ligação da proteína SXL funcional a um sítio específico do pré-mRNA do gene *tra* resulta na formação do mRNA específico de fêmeas, que é o único transcrito funcional desse gene. Se o gene *Sxl* for deletado ou mutado em uma mosca XX, haverá somente a presença do transcrito geral do gene *tra* e a mosca se tornará um macho (O'Neil & Belote, 1992; Gilbert, 2003).

A unidade de transcrição do gene *tra* é relativamente curta, com aproximadamente 1.100 pb, incluindo dois íntrons (O'Neil & Belote, 1992). Embora o gene *tra* seja transcrito a partir do mesmo sítio promotor em ambos os sexos, o pré-mRNA está sujeito a um *splicing* alternativo sexo-específico do primeiro íntron. O pré-mRNA do gene *tra* apresenta dois sítios de emenda 3' (aceptores) no íntron 1 (Figura 5) para a ligação do *spliceosome*, um com maior e outro com menor afinidade por esse complexo protéico. Na ausência da proteína SXL funcional, o *spliceosome* se liga ao sítio de maior afinidade, ocorrendo a incorporação do códon de parada precoce UGA no mRNA. O resultado é a formação do mRNA geral, que codifica um peptídeo truncado inativo. Por outro lado, a proteína funcional SXL presente em fêmeas, bloqueia o sítio usual (de maior afinidade) do pré-mRNA do gene *tra*, obrigando o *spliceosome* a se ligar ao sítio alternativo (de menor afinidade). Consequentemente, ocorre a formação do mRNA específico de fêmea, que codifica a proteína TRA funcional (Figura 6) (O'Neil & Belote, 1992; Gilbert, 2003).

Por sua vez, a proteína funcional TRA, juntamente com a proteína TRA-2, controla o processamento alternativo sexo-específico do pré-mRNA do gene *doublesex* (*dsx*), permitindo a formação de uma proteína DSX específica de fêmea, que atua como fator de transcrição, ativando genes específicos de fêmea e inativando genes masculinos. Em machos, a ausência da proteína TRA resulta na formação de uma proteína DSX específica de macho, outro fator de transcrição que ativa genes específicos de macho e suprime genes específicos de fêmeas (Figura 6) (Gilbert, 2003).

O gene *tra*, embora esteja envolvido nos processos primários de desenvolvimento, que são geralmente conservados, apresenta altas taxas evolutivas entre espécies de *Drosophila* (Kulathinal *et al.*, 2003).

A comparação entre sequências homólogas do gene *tra* de três espécies do subgênero *Sophophora*, *Drosophila erecta*, *D. simulans* e *D. melanogaster*, e duas espécies do subgênero *Drosophila*, *D. hydei* e *D. virilis*, revelou um grau elevado de divergência evolutiva incomum entre as regiões codificadoras (O'Neil & Belote, 1992).

Existem duas explicações possíveis para as altas taxas de substituições em sequências de aminoácidos: eventos múltiplos de evolução adaptativa ou acúmulo de mutações neutras em proteínas de baixa restrição funcional (McAllister & McVean, 2000).

Estudos dos padrões de polimorfismos em populações de *Drosophila americana* (subgênero *Drosophila*) sugerem que a deriva genética é responsável pela evolução rápida do gene *tra* (McAllister & McVean, 2000). Análises similares também sugerem que o locus do gene *tra* de espécies do complexo *melanogaster* (subgênero *Sophophora*) apresenta evolução neutra (Kulathinal *et al.* 2003).

Figura 5 – Estrutura do gene *transformer* (adaptado de McAllister & McVean, 2000). Os introns são representados por linhas enquanto os éxons por retângulos. NSS é a região não sexo-específica, que contém o códon de parada precoce UGA. UTR é a região não traduzida. Após a ligação da proteína SXL ao transcrito do gene *transformer*, ocorre o splicing no sítio alternativo, específico de fêmeas, removendo as regiões NSS e UTR, conjuntamente denominadas *female-specific intron* (FSI).

Figura 6 – Representação esquemática dos eventos de emenda alternativa na cascata gênica de determinação do sexo em *Drosophila*. A via feminina está esquematizada à esquerda, a via masculina à direita, e os genes *transformer* (*tra*) e *doublesex* (*dsx*) estão no centro (modificado de Gilbert, 2003: 470).

1.4.4. *Period*

Variações comportamentais rítmicas são comuns em praticamente todos os seres vivos (Panda *et al.* 2002). Quando estas variações permanecem em condições constantes, dependendo pouco de fatores externos, o ritmo é chamado circadiano (Konopka & Benzer, 1971). A manutenção destes ritmos está associada a um sistema genético integrado, sendo que a caracterização do gene *period* estabeleceu as primeiras bases genéticas de um comportamento animal (Panda *et al.* 2002).

Na espécie *Drosophila melanogaster*, o gene *period* é composto por cinco éxons (Figura 7) e codifica a proteína PER, que possui cerca de 1200 amino ácidos (Citri *et al.* 1987). A proteína PER interage com a proteína TIM, codificada pelo gene *timeless*, que é outro gene importante na regulação do ciclo circadiano. A expressão dos genes *period* e *timeless* é ativada pelo heterodímero CLK-CYC, cujas proteínas componentes são codificadas, respectivamente, pelos genes *clock* e *cycle*, através de um processo de *feedback* autoregulatório, em que o complexo PER-TIM inibe a expressão dos genes *clock* e *cycle* (Panda *et al.* 2002; Tauber & Kyriacou, 2008).

Além de seu papel na regulação do ciclo circadiano, o gene *period* também possui um efeito pleiotrópico no ritmo do IPI, que é um dos componentes do padrão acústico da corte sonora de machos de *Drosophila melanogaster* (Kyriacou & Hall, 1980). Portanto, alterações no *period* podem afetar o som dessa corte, que é espécie específico e tem um importante papel no reconhecimento e preservação do isolamento sexual entre as espécies (Kyriacou & Hall, 1980; Konopka *et al.* 1996).

Figura 7 – Estrutura do gene *period* em *Drosophila melanogaster* (retirado de Franco *et al.* 2010). Os introns são representados por linhas enquanto os éxons por retângulos. O domínio PAS (região C2) está colorido com a cor cinza. No éxon cinco estão indicadas às regiões não conservadas n2, n3 e n4, inter-espaçadas pelas regiões conservadas C3, C4 e C5 (Collot *et al.* 1988). A faixa listrada indica a região n2, formada pela região repetitiva Treonina-Glicina. O retângulo negro representa o domínio inibitório CCID (região C3 até região C4). Abaixo da estrutura do gene, uma barra negra indica a região do éxon cinco que foi analisada no presente trabalho.

Por estar envolvido na corte sonora de *D. melanogaster*, o gene *period* foi considerado um candidato a “gene da especiação” (Colot *et al.* 1988; Coyne, 1992). Em congruência com essa idéia, esse gene também controla o ritmo circadiano de acasalamento em *D. melanogaster* e *D. simulans* (Sakai & Ishida, 2001; Tauber *et al.* 2003), e está relacionado a outras características ligadas ao sucesso reprodutivo, tais como, fecundidade (Beaver *et al.* 2002; Beaver *et al.* 2003) e tempo de duração da cópula em *D. melanogaster* (Beaver & Giebultowicz, 2004). Além de *Drosophila*, foi demonstrado que os níveis de expressão do gene *period* podem causar isolamento reprodutivo pré-zigótico em moscas *Bactrocera cucurbitae* (Diptera: Tephritidae), por acarretarem alocria e alteração nos períodos do dia em que essas moscas estão receptivas à cópula (Miyatake *et al.* 2002).

O gene *period* possui regiões com taxas de substituição nucleotídicas diferentes, sendo formado por trechos altamente similares entre diferentes grupos de insetos, denominados regiões conservadas (C1-C5), e regiões em que não é possível o alinhamento entre algumas espécies de *Drosophila*, denominadas regiões não conservadas (N1-N5) (Colot *et al.* 1988; Kyriacou *et al.* 1996; Tauber & Kyriacou, 2008). Essa particularidade permite a utilização do gene *period* como marcador molecular tanto no nível intraespecífico quanto no interespecífico. De fato, esse gene tem sido utilizado em vários estudos comparativos em *Drosophila*, incluindo comparações entre espécies e populações dentro dos complexos *D. melanogaster* e *D. athabasca* (Kliman & Hey, 1993; Ford *et al.* 1994; Kyriacou *et al.* 2007) e dos grupos *D. virilis* e *D. willistoni* (Hilton & Hey, 1996; Gleason & Powell, 1997). Além disso, homólogos ao gene *period* têm sido descritos em outros grupos, incluindo outros insetos (Barr *et al.* 2005; Mazzotta *et al.* 2005; Mazzoni *et al.* 2006; Regier *et al.* 2008), plantas (McClung, 2006) e vertebrados (Cahill, 2002; Reppert & Weaver, 2002). Nesses táxons, o gene *period* também foi utilizado como marcador para inferências de hipóteses filogenéticas (Barr *et al.* 2005; Regier *et al.* 2008) e divergência populacional (Bauzer *et al.* 2002).

Em *Drosophila*, as regiões mais estudadas do gene *period* são as regiões C2 e N2, localizadas no éxon cinco (Figura 7). A primeira é altamente conservada e inclui o domínio PAS, que contém dois motivos repetidos de 51 pb e que mediam a dimerização das proteínas PER e TIM (Panda *et al.* 2002; Tauber & Kyriacou, 2008). A região N2 possui códons organizados em tandem, responsáveis pela codificação dos aminoácidos treonina e glicina. Essa região foi extensivamente estudada devido à sua alta

variabilidade e ao fato de estar potencialmente envolvida em mecanismos de compensação de temperatura, uma vez que populações naturais de *D. melanogaster* da Europa e Oceania apresentam uma clina latitudinal nas repetições da região N2 (Costa *et al.* 1992; Sawyer *et al.* 1997; Tauber & Kyriacou, 2008).

Na região C-terminal do éxon cinco, encontra-se o domínio CCID (*Clock Cycle Inhibitory Domain*) (Figura 6), responsável por uma etapa importante da maquinaria do relógio circadiano: a inibição da atividade transcricional dos genes *clock* e *cycle* (Crews *et al.* 1988; Sehgal *et al.* 1994; Chang & Reppert, 2003). Embora a região CCID tenha sido pouco utilizada como marcador molecular, ela também pode ser informativa para estudos comparativos, pois nesta região existem trechos conservados e não conservados, permitindo seu uso em diferentes níveis taxonômicos (Colot *et al.* 1988; Chang & Reppert, 2003; Barr *et al.* 2005). Embora os fragmentos do gene *period* utilizados por Gleason & Powell (1997) englobem o domínio CCID, Barr *et al.* (2005) foram os primeiros a utilizarem exclusivamente essa região com propósitos filogenéticos em espécies de *Anastrepha sp.* De acordo com os autores, o domínio CCID apresentou sinal filogenético, indicando que o mesmo pode ser considerado potencialmente informativo em estudos filogenéticos (Barr *et al.* 2005).

1.5. Análise individual *versus* análise combinada

A disponibilidade de um número cada vez maior de dados úteis para a estimação de filogenias intensificou a discussão sobre qual a melhor forma de análise desses dados provenientes de diferentes marcadores. Opiniões diferem sobre se grupos de dados distintos, relacionados a um o mesmo problema filogenético, devem ser analisados separadamente, combinados e analisados simultaneamente, ou analisados separadamente antes de combinar as estimativas independentes usando métodos consenso (Bull *et al.* 1993, De Queiroz *et al.* 1995).

Um argumento geral em favor da análise individual ou separada é que diferentes genes, na maioria das vezes, possuem diferentes níveis de *lineage sorting*, além de estarem sujeitos a eventos distintos de hibridização, duplicação gênica e perda diferencial, introgressão e transferência lateral (Leigh *et al.* 2008). Em consequência disso, simulações computacionais indicam que uma filogenia obtida a partir da análise combinada de grupos de caracteres heterogêneos pode ser pior do que filogenias dos

mesmos grupos de dados analisados separadamente, uma vez que seria mais informativo obter uma resposta certa e outra errada a partir da análise separada, ao invés de uma única resposta errada a partir da análise simultânea (Bull *et al.* 1993, Miyamoto & Fitch, 1995).

Bull *et al.* (1993) sugerem que grupos de dados distintos, antes de serem combinados e analisados simultaneamente, devem ser submetidos a um teste estatístico de incongruência filogenética. Se a hipótese nula de congruência é aceita, os dados podem ser combinados, mas se a hipótese nula é rejeitada, os grupos de dados não devem ser combinados em uma análise que assuma homogeneidade de caracteres.

No entanto, em muitos casos, grupos de dados heterogêneos são tratados como homogêneos, pois o nível de heterogeneidade observado no tamanho amostral considerado não é suficiente para rejeitar a hipótese nula (Bull *et al.* 1993).

Por outro lado, problemas com falso negativo também vêm sendo reportados em relação ao teste de congruência mais comumente usado, o ILD (*Incongruence Length Difference test*) (Farris *et al.*, 1995) ou PHT (*Partition Homogeneity Test*), que por ser um teste baseado em parcimônia, é muito sensível às características evolutivas das partições analisadas, isto é, à variação das taxas evolutivas e das taxas de substituição entre os sítios (Darlu & Lecointre, 2002; Barker & Lutzoni, 2002; Leigh *et al.* 2008).

A simples comparação entre filogenias obtidas a partir da análise individual de dois grupos de dados distintos também permite identificar, rapidamente, a presença de heterogeneidade. Sendo assim, a análise separada tem sido utilizada para explorar possíveis incongruências entre diferentes grupos de dados (De Queiroz *et al.* 1995, Huelsenbeck *et al.* 1996).

O resultado empírico que fundamenta a maioria dos argumentos em favor da análise individual é a observação de que estimativas filogenéticas derivadas de diferentes grupos de dados são frequentemente incongruentes (De Queiroz *et al.* 1995, Cunningham, 1997). Bull *et al.* (1993) defendem que se essa incongruência é maior do que a esperada devido a erro amostral, portanto, a combinação e análise simultânea dos grupos de dados é inapropriada.

Verificada a incongruência entre grupos de dados distintos, métodos consenso podem ser utilizados para sumarizar as hipóteses alternativas, obtidas a partir da análise individual de cada grupo de caracteres, em um cladograma consenso. O problema apresentado pelas metodologias consenso é que estas podem resultar na perda de

informações ao sumarizar grupos de dados individuais na forma de filogenias fundamentais (Barrett *et al.* 1991; De Queiroz *et al.* 1995).

É importante mencionar que ao se optar por analisar diferentes grupos de dados separadamente, o uso de métodos consenso não é obrigatório. Similarmente, uma análise simultânea não implica necessariamente na incorporação de todos os grupos de caracteres disponíveis em uma única análise (De Queiroz *et al.* 1995).

Um dos argumentos em favor da análise simultânea baseia-se no princípio filosófico de “evidência total” (Carnap, 1950), que é a recomendação do uso de todas as evidências disponíveis na estimação de uma probabilidade (Kluge, 1989).

Alguns autores também defendem que a análise simultânea evita as arbitrariedades inerentes aos métodos consenso, como, por exemplo, a necessidade de escolher entre as várias metodologias consenso para sumarizar as congruências entre os grupos de dados analisados (Kluge, 1989). Deste modo, a análise simultânea seria preferível, em detrimento dos métodos consenso, por gerar hipóteses filogenéticas que representam um sumário mais eficiente das evidências disponíveis, e, consequentemente, por apresentar maior poder descritivo e explicativo (De Queiroz *et al.* 1995, De Queiroz & Gatesy, 2007).

O argumento em favor da análise simultânea, que recebeu mais atenção recentemente, refere-se à capacidade dessa análise em descobrir grupos filogenéticos reais (De Queiroz *et al.* 1995, De Queiroz & Gatesy, 2007). De fato, com um número de caracteres maior, o sinal filogenético tem maior probabilidade de se impor sobre o ruído, resultando em uma estimativa mais acurada da filogenia verdadeira. Em outras palavras, o erro amostral é reduzido pelo aumento do número de dados. Esse aumento do sinal filogenético, em análises simultâneas, pode refletir em um aumento surpreendente no suporte de *bootstrap* ou das probabilidades posteriores para um ramo particular (De Queiroz *et al.* 1995, De Queiroz & Gatesy, 2007).

Diante do fato de que tanto a análise individual quanto a análise combinada apresentam benefícios potenciais, uma solução para a questão de como analisar da melhor forma diferentes grupos de dados relacionados a um problema filogenético, seria realizar ambos os tipos de análise.

2. OBJETIVOS

O presente trabalho tem como objetivos:

- Inferir as relações filogenéticas entre as espécies do complexo *Drosophila buzzatii*, dando ênfase ao cluster *D. buzzatii*, por meio de análises multilocus de sequências de genes mitocondriais (COI e COII) e nucleares (EF-1 α F1, *transformer* e *period*).
- Traçar parâmetros comparativos por meio de testes de homogeneidade por partição, das taxas evolutivas e das filogenias obtidas a partir das análises separadas, para avaliar a possibilidade de uma análise combinada de todos os genes ou de análises combinadas parciais (dois, três, quatro ou cinco dos genes).
- Avaliar se as regiões analisadas dos genes *period* e *transformer* estão sob pressões seletivas e, assim, contribuir para o entendimento da evolução molecular desses loci para o complexo *Drosophila buzzatii*.
- Realizar estimativas de tempo de divergência para as espécies do cluster *D. buzzatii*, na tentativa de inferir se os processos pelos quais o grupo passou ao longo de sua diversificação estão relacionados aos eventos paleogeográficos do Terciário ou às flutuações paleoclimáticas do Quaternário.

3. MATERIAL E MÉTODOS

3.1. Material biológico

Foram utilizadas neste trabalho 12 espécies do complexo *D. buzzatii* (Tabela 1). As espécies *D. virilis*, *D. hydei* e *D. mojavensis* foram utilizadas como grupos externos nas análises filogenéticas, e suas sequências foram geradas em laboratório ou obtidas no *GenBank*, quando disponíveis (números de acesso: *D. virilis* - BK006340.1; *D. hydei* - EU390734.1, GU597488.1; *D. mojavensis* - XM_002012150.1, XM_002010950.1).

Tabela 1. Localidades e respectivos códigos das amostras de espécimes do complexo *D. buzzatii* que foram utilizadas nas análises do presente estudo.

Espécie	Código	Localidade	Cordenadas
<i>D. richardsoni</i>	<i>D. richardsoni</i>	La Parguera-Porto Rico	17°58'27" N 67°02'09" W
<i>D. stalker</i>	<i>D. stalker</i>	Saint Petersburg-EUA	27°42'55" N 82°35'32" W
<i>D. martensis</i>	<i>D. martensis</i>	Guaca-Venezuela	10°38'12" N 63°29'38" W
<i>D. venezolana</i>	<i>D. venezolana</i>	Guaca-Venezuela	10°38'12" N 63°29'38" W
<i>D. starmeri</i>	<i>D. starmeri</i>	Riohacha-Colômbia	11°32'10" N 72°56'46" W
<i>D. seriema</i>	D62	Mucugê-BA	13°00'00" S 41°23'59" W
<i>D. seriema</i>	D40	Serra do Cipó-MG	19°18'58" S 43°36'30" W
<i>D. borborema</i>	1281	Milagres-BA	11°12'25" S 39°53'04" W
<i>D. borborema</i>	N70	Junco do Seridó- PB	06°59'45" S 36°42'26" W
<i>D. gouveai</i>	J78	Ibotirama-BA	12°06'00" S 43°18'00" W
<i>D. gouveai</i>	H6	Altinópolis-SP	21°06'00" S 47°35'59" W
<i>D. serido</i>	1431	Milagres-BA	11°12'25" S 39°53'04" W
<i>D. serido</i>	H49	Bertioga-SP	23°52'06" S 46°08'08" W
<i>D. serido</i>	N20	Arraial do Cabo-RJ	23°00'00" S 42°00'00" W
<i>D. serido</i>	N45	Mucugê-BA	13°00'00" S 41°23'59" W
<i>D. antonietae</i>	D93	Cianorte-PR	23°42'06" S 52°35'56" W
<i>D. antonietae</i>	H47	Santiago-RS	29°12'42" S 54°51'49" W
<i>D. koepferae</i>	B26	Famatina-Argentina	29°11'37" S 67°23'10" W
<i>D. koepferae</i>	Suy DK	Suyuque -Argentina	33°37'22" S 66°53'55" W
<i>D. buzzatii</i>	N36	Petrolina-PE	09°06'00" S 40°36'00" W
<i>D. buzzatii</i>	H98	San Juan-Argentina	31°34'47" S 68°31'12" W

As espécies analisadas do cluster *D. buzzatii* foram coletadas por armadilhas fechadas (Tidon & Sene, 1988) contendo banana, laranja e fermento (*Saccharomyces cerevisiae*), e as moscas identificadas por análise morfológica da genitália masculina

externa (*aedeagus*) (Vilela, 1983), sendo os machos identificados diretamente e as fêmeas pela sua progênie masculina.

O DNA genômico de um indivíduo macho de cada espécime listado acima foi utilizado na amplificação e sequenciamento dos genes mitocondriais, COI e COII, e dos genes nucleares, EF-1 α F1, *transformer* e *period*. Para alguns destes espécimes sequências dos genes mitocondriais e/ou nucleares já estavam disponíveis no Laboratório de Genética Evolutiva da USP – Campus Ribeirão Preto, SP.

3.2. Extração de DNA genômico

O DNA genômico foi extraído dos indivíduos utilizando-se o conjunto de reagentes Wizzard® Genomic DNA Purification Kit (Promega). Em um tubo de microcentrífuga contendo 60 μ L de EDTA 0,5 M e 250 μ L de Solução de Lise de Núcleo, um indivíduo macho adulto foi macerado e homogeneizado. Depois de adicionados 10 μ L de Proteinase K 20 mg/mL ao homogeneizado, a amostra foi incubada à 65°C por duas horas. Após este período, foram adicionados 100 μ L de Solução de Precipitação de Proteínas à amostra, que foi agitada vigorosamente por 20 segundos. O tubo de microcentrífuga contendo a amostra foi colocado no gelo por 10 minutos e em seguida centrifugado por quatro minutos a 14.000 rpm. O sobrenadante contendo o DNA foi transferido para um novo tubo contendo 300 μ L de isopropanol, sendo centrifugado a 14.000 rpm por um minuto para a precipitação do DNA. Para lavagem do DNA, foram adicionados 300 μ L de etanol 70%. Após a remoção e secagem do etanol, o DNA foi ressuspenso em 60-100 μ L de água bi-destilada autoclavada, e conservado a -20°C.

3.3. Amplificação dos genes mitocondriais e nucleares

Como citado anteriormente, os marcadores moleculares mitocondriais utilizados foram os genes das subunidade I e II da citocromo oxidase (COI e COII), e os genes nucleares utilizados foram Fator de Elongação 1 Alfa - cópia paráloga F1 (EF-1 α F1), *transformer* (tra) e *period* (per).

As amplificações dos genes foram realizadas por meio da reação em cadeia da polimerase (PCR). A PCR foi feita em um Termociclador PE Applied Biosystems (DNA Thermal Cycler 480). Para os genes COI, COII e transformer, cada amostra

continha 4 µL de DNA, 2,5 µL de dNTP (20 pmol/µL), 2,5 µL de cada um dos *primers* (0,4 pmol/µL), 2,5 µL de tampão, 0,3 µL (1,5 U) de *Taq* polimerase (Invitrogen) e água estéril para completar 25 µL de volume final. Para os genes EF-1αF1 e *period*, cada amostra continha 5-7 µL de DNA, 4 µL de cada um dos *primers* (0,4 pmol/µL), 4 µL de tampão, 0,54 µL (1,5 U) de *Taq* polimerase (Invitrogen) e água estéril para completar 40 µL de volume final.

Os primers utilizados para o isolamento, amplificação e sequenciamento de cada gene estão descritos na TABELA 2.

As condições das PCR para amplificação foram específicas para cada gene:

COI - 1'30" a 94°C e 30 ciclos de 40" a 94°C, 40" a 46°C e 2' a 72°C; COII – 1'30" a 94°C e 30 ciclos de 40" a 94°C, 40" a 50°C e 2' a 72°C; EF-1αF1 – 2' a 94°C e 35 ciclos de 20" a 94°C, 30" a 51°C e 1' a 72°C, seguidos por 10' a 72°C de extensão final, *transformer* - 1'30" a 94°C e 30 ciclos de 40" a 94°C, 40" a 48°C e 2' a 72°C; *period* - 1'30" a 94°C e 39 ciclos de 30" a 94°C, 30" a 55°C e 1' a 72°C e um último ciclo de extensão de 30" a 94°C, 30" a 55°C e 1' a 72°C.

Tabela 2. Primers utilizados nas PCR para isolamento, amplificação e sequenciamento de DNA das espécies do complexo *D.buzzatii*.

Locus	Sequência dos primers	Referência
COI	5'CAATTTATCGCCTAACTTCAGCC3' (<i>forward</i>)	Simon <i>et al.</i> (1994)
	5'CCCGGTAAAATATAAACTTC' (<i>reverse</i>)	Simon <i>et al.</i> (1994)
COII	5'AATATGGCAGATTAGTGCA3' (<i>forward</i>)	Simon <i>et al.</i> (1994)
	5'CCACAAATTCTGAACATTGACCA3' (<i>reverse</i>)	Simon <i>et al.</i> (1994)
EF-1αF1	5'TCCGGATGGCAYGGCGAGAAYATG3' (<i>forward</i>)	Hovemann <i>et al.</i> (1988)
	5' TGTGAGCAGTGTGGCAATCCAA3' (<i>reverse</i>)	Hovemann <i>et al.</i> (1988)
<i>transformer</i>	5' TTTGTTGTTTTTTTCTAGTG 3' (<i>forward</i>)	O'Neil & Belote (1992)
	5'GCGCCATACATGGGATTAAA3' (<i>reverse</i>)	O'Neil & Belote (1992)
<i>period</i>	5' TGGGAGGGCGAGGCCAACAA 3' (<i>forward</i>)	Franco <i>et al.</i> (2010)
	5' GGCATGGGTTGGTACATCAT 3' (<i>reverse</i>)	Franco <i>et al.</i> (2010)

3.4. Eletroforese de verificação

Foi utilizada a eletroforese em gel de agarose 1% para a verificação da amplificação. Para preparação do gel, 0,7g de agarose foi diluída em 70 mL de Tampão TBE 1X (Tris-Base 0,089M; ácido bórico 0,089M; EDTA 2mM), a uma temperatura de aproximadamente 100°C. A solução foi despejada sobre uma placa molde. Após a

solidificação da agarose, a placa foi transferida para cuba de eletroforese, contendo o mesmo tampão de fusão da agarose. A eletroforese foi realizada em 50mA, por aproximadamente 1 hora. Os géis foram corados com uma solução de brometo de etídeo, diluído em Tampão 1X TBE. O DNA no gel foi visualizado com sua exposição à luz UV, em um transiluminador.

3.5. Eletroforese para recuperação das bandas de interesse (EF-1 α F1 e *period*)

O produto da PCR do gene EF-1 α F1 resultou em duas bandas depois de submetido à eletroforese de verificação em gel de agarose 1%, enquanto o produto da PCR do gene *period* resultou em uma ou mais bandas. Visando a recuperação apenas dos fragmentos de interesse, o de 500 pb para o gene EF-1 α F1 e o de 400 pb para o gene *period*, foi utilizada uma nova eletroforese em gel de agarose 1%, em uma cuba de eletroforese maior e por um intervalo de tempo mais longo, neste caso 5 horas, para permitir um maior espaçamento entre as bandas, evitando, deste modo, a contaminação no processo de recuperação da banda de interesse. Para preparação do gel, 2g de agarose foi diluída em 200 mL de Tampão TBE 1X a 100°C, aproximadamente. Os géis foram também corados com solução de brometo de etídeo, e o DNA no gel novamente visualizado com sua exposição à luz UV no transiluminador.

3.6. Purificação dos fragmentos dos genes COI, COII e *transformer*

Os produtos da PCR dos genes COI, COII e *transformer* foram purificados com o conjunto de reagentes *ExoSAP-IT* (Amersham Biosciences Part of GE Healthcare), seguindo-se as instruções do fabricante. Para cada 5 μ l do produto da PCR foram adicionados 2 μ l de *ExoSAP-IT*. Em seguida, essa mistura foi incubada por 15 minutos a 37°C e outros 15 minutos a 80°C. Esses procedimentos permitem a degradação de “primers” e dNTPs presentes na PCR. Após esses procedimentos, as amostras estavam prontas para realização da reação de sequenciamento.

3.7. Purificação dos fragmentos dos genes EF-1 α F1 e *period*

As bandas de interesse cortadas do gel de agarose 1% da eletroforese de recuperação foram eluídas e purificadas com o conjunto de reagentes “*GFX PCR DNA and Gel Band Purification Kit*” (GE Healthcare), de acordo com as recomendações do

fabricante. O DNA purificado foi, então, utilizado diretamente nas reações de sequenciamento.

3.8. Sequenciamento direto

Na reação de sequenciamento foi utilizado o conjunto de reagentes BigDye® *Terminator v3.1 Cycle Sequencing kit* (Applied Biosystems), que utiliza o protocolo de sequenciamento Sanger *et al.* (1977).

Cada amostra a ser sequenciada continha 1,5-2,0 µl de BigDye, que contém os nucleotídeos e as enzimas necessárias à reação, 5-10 µl de DNA purificado, 2 µl de um dos *primers* (*forward* ou *reverse*) e, quando necessário, água destilada até completar 10 µl.

As amostras foram submetidas à PCR nas seguintes condições: 2'' a 96°C e 25 ciclos de 10'' a 96°C, 5'' a 50°C e 4' a 60°C.

A reação de precipitação envolveu várias lavagens do DNA. Primeiramente, foram adicionados 80 µl de isopropanol 75% aos tubos de microcentrífuga contendo as reações de sequenciamento, que foram agitados gentilmente e deixados em repouso por 15 minutos à temperatura ambiente. Em seguida, cada amostra foi centrifugada a 15.000 g por 45 minutos e o sobrenadante descartado. Posteriormente, cada tubo recebeu 200 µl de etanol 70% gelado, sendo novamente agitados e centrifugados a 15.000 g por 10 minutos. O sobrenadante foi descartado e o DNA novamente lavado com etanol 70% gelado. As amostras foram mantidas em estufa a 37°C até a completa evaporação do álcool.

Antes de serem aplicadas no sequenciador, as amostras foram desnaturadas pela adição de 12 µl de *Template Suppression Reagent* (Applied Biosystems), seguido de incubação por 2 minutos a 95°C. As amostras foram sequenciadas no sequenciador automático de DNA modelo ABI Prism 377.

3.9. Análises das sequências de DNA

Sequências *forward* e *reverse* de cada táxon foram comparadas, corrigidas e editadas usando o programa Chromas Lite 2.0. Os alinhamentos múltiplos das sequências de cada grupo de dados foram gerados no programa Clustal W versão 1.8 (Thompson *et al.*, 1994). As análises descritivas das sequências e as distâncias

genéticas foram realizadas usando o programa MEGA 4.0 (Tamura et al. 2007). A presença de saturação nas substituições nucleotídicas foi verificada usando o programa DAMBE 4.2.13 (Xia & Xie, 2001). A congruência entre os dados de cada partição gênica foi estimada pelo teste de homogeneidade por partição (PHT), que essencialmente é o teste de ILD (*Incongruence Length Difference test*) de Farris et al. (1995). Utilizou-se método de busca heurística, sendo feitas 1.000 replicações.

3.10. Reconstrução filogenética

As análises filogenéticas foram realizadas através dos critérios de Máxima Parcimônia (MP), Máxima Verossimilhança (ML) e Inferência Bayesiana (BI), as duas primeiras utilizando o programa PAUP*4.0b10 (Swofford, 2002), e a última com o auxílio do programa MrBayes v.3.1.2 (Huelsenbeck & Ronquist 2001).

Na MP as árvores iniciais foram obtidas por *stepwise addition* com adição aleatória de sequência, restando-se 10 árvores após cada réplica. Foram executadas buscas heurísticas para obtenção das melhores árvores por *branch-swapping*, utilizando o algoritmo *tree-bisection-reconnection* (TBR), com 10.000 réplicas e restando-se 1 árvore por réplica. Os caracteres foram tratados como não ordenados e com pesos iguais (“Parcimônia de Fitch”, Fitch 1971). A sustentação dos ramos foi estimada usando a análise de *bootstrap* (Felsenstein 1985), calculada em 1000 réplicas por meio de busca heurística, utilizando-se adição aleatória de táxon e *branch-swapping* por TBR.

Na ML as árvores iniciais também foram obtidas por *stepwise addition* com adição aleatória de sequência, restando-se uma árvore após cada réplica. A obtenção das melhores árvores foi feita por *branch-swapping*, utilizando o algoritmo *tree-bisection-reconnection* (TBR), com 1.000 réplicas e restando-se 1 árvore por réplica. A sustentação dos ramos foi estimada usando a análise de *bootstrap* (BP), calculada em 100 réplicas através de busca heurística utilizando-se adição simples de táxon e *branch-swapping* por SPR.

Na BI as buscas consistiram em quatro corridas independentes, cada uma com quatro cadeias simultâneas (i.e., uma “cold chain” e três “heated chains”). As cadeias de Markov foram iniciadas com uma árvore aleatória, o número inicial de gerações foi de 1.000.000, amostrada a cada 1.000 gerações. Quando completado o número total de gerações, era monitorado visualmente se o programa havia atingido a convergência (i.e., *average standart deviation of split frequencies* < 0.01), em caso negativo, um número

maior de gerações era aplicado (500.000 – 1.000.000). Após descartar 25% das árvores, correspondentes ao “burnin”, uma árvore de consenso era calculada a partir das “cold chains” de cada uma das quatro corridas. As Probabilidades Posteriores (PP) foram calculadas durante a análise para a sustentação dos ramos.

Os modelos de substituição nucleotídica empregados nas análises de ML e BI foram determinados utilizando-se o programa Modeltest 3.7 (Posada & Crandal, 1998).

Para cada gene foram feitas análises filogenéticas utilizando os três critérios de otimização citados acima. Foram também realizadas análises por BI combinando os dados dos genes mitocondriais (COI e COII), dos genes nucleares (EF-1 α F1, *transformer* e *period*) e de todos os genes.

3.11. Testes de neutralidade baseados em dn e ds

Foram realizados testes para detecção de desvios em relação ao esperado pela teoria neutra da evolução (Kimura, 1983) para os genes *transformer* e *period*, a partir da comparação de taxas de substituições sinônimas (ds) e não sinônimas (dn). Para cada espécie, as taxas de dn e ds foram comparadas de acordo com o método Nei-Gojobori (Nei e Gojobori, 1986). As diferenças entre essas taxas foram testadas com o Teste-Z de seleção, com um nível de significância de 5%, em que o valor de p é a probabilidade de rejeitar a hipótese nula de neutralidade (H_0 : $dn=ds$, Nei e Kumar, 2000). O Teste-Z de seleção foi realizado no programa MEGA 4.0 (Tamura et al. 2007).

Sinais de seleção natural em aminoácidos específicos foram investigados no servidor SELECTON (disponível em <http://selecton.bioinfo.tau.ac.il>), com o qual foram calculadas as razões de dn/ds (ω) para cada códon, com uma metodologia Bayesiana. A árvore obtida por ML de cada gene foi utilizada na análise para detectar seleção em ramos específicos da filogenia. Para detectar seleção direcional, os resultados obtidos através do modelo evolutivo M8 (Yang et al. 2000), que permite a identificação de seleção direcional, neutra ou purificadora, foram comparados com os modelos M8a (Swanson et al. 2003) e M7 (Yang et al. 2000), os quais permitem que ω atinja apenas valores entre -1 e 0, fornecendo assim um modelo nulo para detecção de seleção direcional. A comparação entre os resultados do modelo M8, com aqueles obtidos pelos modelos M8a e M7, foi realizada através do teste da razão de máxima verossimilhança (*Likelihood Ratio Test*: LRT), como sugerido por Yang et al. (2000) e Stern et al. (2007).

3.12. Tempos de divergência

Verificou-se, inicialmente, para quais partições gênicas as taxas de substituição nucleotídica não variavam de forma estatisticamente significativa ao longo dos ramos reconstruídos a partir da matriz de sequências, isto é, quais dos genes estudados estão evoluindo de acordo com uma taxa evolutiva global (hipótese do relógio molecular). Para isso, foram realizados para cada partição gênica estudada, com o auxílio do programa PAUP*4.0b10, testes de razão de verossimilhança (LRT) entre os modelos de relógio molecular forçado e relaxado.

O gene que apresentou taxas de substituição homogêneas entre os ramos foi, então, utilizado nos cálculos de tempo de divergência por MPL (*Tree-based mean path length method*; Bremer & Gustafsson, 1997; Britton *et al.* 2002), com o auxílio do programa PATHd8 (Britton *et al.* 2006). O programa utiliza uma árvore com tamanho de ramos proporcional ao número de substituições, obtida, por exemplo, por ML ou BI. O método baseia-se na estimativa das idades dos nós a partir das médias dos comprimentos de ramos dos próprios nós até os táxons terminais, corrigindo violações do relógio molecular sugeridas pelos nós de referência - nós em que foram definidos pontos de calibração (o programa exige a definição de ao menos um nó com idade fixa, além de permitir a definição de idades máximas ou mínimas).

4. RESULTADOS

4.1. Citocromo oxidase I

Foram obtidas sequências de 600 pb do gene COI para as espécies do complexo *D. buzzatii* (Apêndice 1A). A primeira base das sequências corresponde à base de posição 1.559 do DNA mitocondrial de *Drosophila yakuba* (número de acesso do *GenBank* NC_001322), e o primeiro códon a partir dessa base equivale ao vigésimo nono aminoácido do gene em questão. Em relação à proporção média de bases nitrogenadas, essas sequências nucleotídicas apresentaram 28% de A, 15,17% de C, 17,1% de G e 39,73% de T. Resultados que eram esperados, pois genes mitocondriais de insetos, em geral, possuem uma maior proporção de A e T (Simon *et al.* 1994).

A análise de saturação (Figura 8) revelou que as transições tornam-se saturadas a partir de, aproximadamente, 6% de divergência, indicando a ocorrência de substituições múltiplas e homoplasias nessa classe de mutações, enquanto as transversões permanecem informativas.

Figura 8 – Taxas de transições (S) e transversões (V) representadas graficamente contra a distância genética entre as sequências parciais do gene COI para as espécies do complexo *D. buzzatii*, calculada com base no modelo de substituição nucleotídica de GTR (Rodriguez *et al.* 1990).

Foram consideradas neste trabalho as seguintes categorias de sustentação de ramos: *bootstrap* (BS) > 88% e probabilidade posterior (PP) > 91% como valores

robustos, segundo Zander (2004) estes valores representam os valores mínimos requeridos para um intervalo de confiança de 95% de um nó, supondo uma distribuição binomial das probabilidades posteriores e das frequências de *bootstrap*. Outros intervalos de valores de suporte definidos arbitrariamente são moderado (76% - 87 % para BS e 85% – 90 % para PP), fraco (63% – 75 % para BS e 75% – 84 % para PP) e não significativo (< 63% para BS e < 75% para PP).

4.1.1. Máxima parcimônia

Para as análises filogenéticas por MP foram considerados 176 caracteres variáveis (incluindo os grupos externos), sendo 125 parcimoniosamente informativos.

A busca heurística, com todos os caracteres informativos pesados igualmente, resultou em 97 filogenias igualmente parcimoniosas, sendo que o cladograma consenso (Figura 9) apresentou 392 passos, índice de consistência (CI) = 0,564 e índice de retenção (RI) = 0,674.

Os principais resultados da análise filogenética por MP das sequências de COI foram:

- O monofiletismo do cluster *D. buzzatii* (BP = 99%);
- Os clusters *D. stalker*i e *D. martensis* não apresentaram monofilia recíproca e, portanto, as relações inter-cluster não foram definidas;
- *D. martensis* (cluster *D. martensis*) e *D. stalker*i (cluster *D. stalker*i) ocuparam de forma politômica a posição filogenética mais basal dentro do complexo *D. buzzatii*;
- *D. starmer*i e *D. venezolana* foram agrupadas em um clado que representou a linhagem evolutiva mais intimamente relacionada ao cluster *D. buzzatii*, porém, com valor não significativo de *bootstrap* (BP < 50%);
- O cluster *D. buzzatii* foi dividido em duas linhagens evolutivas principais, a mais basal representada pelo clado formado pelas espécies *D. buzzatii* e *D. koepferae* (BP = 70%), e a segunda representada pelo clado que compreende as demais espécies, agrupadas em uma politomia (BP = 100%).

Figura 9 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene COI, segundo máxima parcimônia. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos do consenso estrito.

A fim de explorar um pouco mais os dados do gene COI por meio da MP, foram realizadas mais duas análises filogenéticas, testando outros dois *schemes* de pesagem para cada tipo de mutação: a primeira atribuindo peso 0,25 às transições e 1 às transversões, e a segunda com as transições omitidas (peso zero).

Os cladogramas consenso obtidos nessas duas análises adicionais foram congruentes, entretanto, a análise com as transições desconsideradas (peso zero) foi menos informativa, resultando em um maior número de politomias. A topologia consenso da primeira análise ($S = 0,25$ e $V = 1$) é apresentada na Figura 10.

As filogenias obtidas com pesagem diferencial foram bastante contrastantes em relação à obtida com os caracteres informativos pesados igualmente, as principais incongruências foram:

- *D. richardsoni* foi posicionada como sendo a espécie mais basal dentro do complexo *D. buzzatii*, no entanto, com baixo suporte estatístico (BP = 58%);

- *D. satakeri*, *D. starmeri* e *D. venezolana* foram agrupadas em um mesmo clado, porém, com valor de *bootstrap* não significativo;
- *D. martensis* foi alocada como espécie irmã do cluster *D. buzzatii*, também com baixo suporte estatístico;
- Em relação ao cluster *D. buzzatii*, *D. koepferae* não foi agrupada em um mesmo ramo com *D. buzzatii* (BP = 65%);
- *D. antonietae* foi posicionada fora do clado formado pelas espécies *D. serido*, *D. gouveai*, *D. borborema* e *D. seriema*.

Figura 10 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene COI, utilizando pesagem diferencial para as transições e transversões ($S = 0,25$ e $V = 1$), segundo máxima parcimônia. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos do consenso estrito.

4.1.2. Máxima verossimilhança

O modelo GTR + I + G (Rodriguez *et al.* 1990) foi selecionado, utilizando o critério Akaike (AIC), como o modelo de substituição nucleotídica que melhor explica a evolução das sequências do gene COI.

A árvore ótima obtida por ML (Figura 11) teve como valor de verossimilhança $-\ln L = 2626.15834$. Os principais resultados da análise das sequências de COI por ML foram:

- O monofiletismo do complexo *D. buzzatii* (BP = 55%);
- O clado formado por *D. stalker* e as espécies irmãs, *D. starmeri* e *D. venezolana*, representaram a linhagem evolutiva mais basal do complexo *D. buzzatii*, entretanto, com valor de *bootstrap* não significativo (BP < 50%);
- *D. koepferae* foi mantida fora do *D. serido* “sibling set”, formando com *D. buzzatii* a linhagem evolutiva mais basal do cluster *D. buzzatii* (BP = 77%);
- *D. antonietae* foi posicionada como espécie irmã da linhagem evolutiva representada pelo clado AB (Manfrin *et al.* 2001), *D. serido*, *D. gouveai*, *D. borborema* e *D. seriema* (BP = 82%).

Figura 11 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene COI, segundo máxima verossimilhança. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos da árvore ótima. A escala indica número de substituições/sítio.

4.1.3. Inferência bayesiana

Os principais resultados da análise filogenética por BI das sequências do gene COI (Figura 12) foram:

- O monofiletismo do complexo *D. buzzatii* e do cluster *D. buzzatii*, ambos com alto suporte probabilístico (PP = 98% e PP = 100%, respectivamente);
- A divisão do cluster *D. buzzatii* em três linhagens evolutivas principais, a mais basal representada pelo clado formado por *D. buzzatii* e *D. koepferae*, a segunda pelo clado formado pela espécie *D. antonietae*, e a terceira linhagem evolutiva agrupando, numa politomia, as espécies *D. gouveai*, *D. borborema*, *D. seriema* e *D. serido*.

Figura 12 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene COI, segundo inferência bayesiana. Valores das probabilidades posteriores bayesianas indicadas acima dos ramos. A escala indica número de substituições/sítio.

4.2. Citocromo oxidase II

As sequências parciais obtidas para o gene COII foram de 555 pb (Apêndice 1B), com início na primeira base do gene em questão de *Drosophila melanogaster* (número de acesso do *GenBank* GQ229520.2). Os fragmentos apresentaram um predomínio de A (33,18%) e T (42,09%) sobre C (13,79%) e G (10,94%).

Assim como para o gene COI, a análise de saturação indicou saturação apenas na curva de transições, a partir de 9% de divergência (Figura 13), aproximadamente.

Figura 13 – Taxas de transições (S) e transversões (V) representadas graficamente contra a distância genética entre as sequências parciais do gene COII para as espécies do complexo *D. buzzatii*, calculada com base no modelo de substituição nucleotídica GTR (Rodríguez *et al.* 1990).

4.2.1. Máxima parcimônia

Para a análise filogenética por MP foram considerados 156 caracteres variáveis (incluindo os grupos externos), sendo 104 parcimoniosamente informativos.

A busca heurística, com todos os caracteres informativos pesados igualmente, resultou em nove filogenias igualmente parcimoniosas, sendo que o cladograma consenso apresentou 313 passos, CI = 0,625 e RI = 0,707 (Figura 14).

Os principais resultados da análise filogenética por MP das sequências de COII foram:

- O monofiletismo do complexo *D. buzzatii* (BP = 84%), que foi dividido em dois grandes clados, um agrupando os clusters *D. stalker*i e *D. martensis* (BP = 55%) e outro agrupando as espécies do cluster *D. buzzatii* (BP = 97%);
- O monofiletismo do cluster *D. martensis*, entretanto, com valor não significativo de *bootstrap* (BP < 50%), que ocupou uma posição filogenética derivada em relação ao cluster *D. stalker*i, que constituiu um grupo parafilético;
- A linhagem evolutiva representada pela espécie *D. koepferae* foi posicionada como sendo a mais basal dentro do cluster *D. buzzatti*, que foi dividido em mais outras duas linhagens evolutivas, uma representada pelo clado que compreendeu as espécies *D. buzzatii*, *D. antonietae* (BP = 63%), e outra pelo clado que agrupou de forma politômica as espécies *D. serido*, *D. gouveai*, *D. borborema*, *D. seriema* (BP = 86%).

Figura 14 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene COII, segundo máxima parcimônia. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos do consenso estrito.

4.2.2. Máxima verossimilhança

O modelo TIM + I + G foi selecionado, utilizando o critério Akaike (AIC), como o modelo de substituição nucleotídica que melhor explica a evolução das sequências do gene COII.

A árvore ótima obtida por ML (Figura 15) teve como valor de verossimilhança $-\ln L = 2244.79775$. Os mesmos agrupamentos observados na análise por MP foram obtidos na análise por ML das sequências de COII, diferindo apenas nos valores de suporte estatístico.

Figura 15 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene COII, segundo máxima verossimilhança. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos da árvore ótima. A escala indica número de substituições/sítio.

4.2.3. Inferência bayesiana

Os principais resultados da análise filogenética por BI das sequências do gene COII (Figura 16) foram:

- O monofiletismo do complexo *D. buzzatii* (PP = 99%), do cluster *D. buzzatii* (PP = 100%) e do ramo que compreende os clusters *D. stalker* e *D. martensis* (BP = 95%), cujas espécies, neste caso, foram agrupadas de forma politômica;
- Contrastando com as duas topologias anteriores, o ramo representado pela espécie *D. antonietae* não foi agrupado em um mesmo clado com os representantes da espécie *D. buzzatii*;
- O cluster *D. buzzatii* foi dividido em quatro linhagens evolutivas, a mais basal representada pela espécie *D. koepferae* (PP = 98%), a segunda pela espécie *D. buzzatii* (PP = 100%), a terceira pelo clado que compreende a espécie *D. antonietae* e o representante H6 da espécie *D. gouveai* (PP = 100%), e a mais derivada agrupando *D. serido*, *D. borborema*, *D. seriema* e o representante J78 da espécie *D. gouveai* (PP = 96%).

Figura 16 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene COII, segundo inferência bayesiana. Valores das probabilidades posteriores bayesianas indicadas acima dos ramos. A escala indica número de substituições/sítio.

4.3. EF-1 α F1

As sequências parciais obtidas para o gene EF-1 α F1 foram de 448 pb (Apêndice 1C), a primeira base das sequências corresponde à base de posição 2.664 do gene EF-1 α F1 de *Drosophila melanogaster* (Hovemann *et al.* 1988) (número de acesso do GenBank X06869.1). Toda a porção do fragmento sequenciado localiza-se dentro do segundo éxon do gene em questão. Em relação à proporção média de bases nitrogenadas, as sequências nucleotídicas apresentaram 22,13% de A, 29,88% de C, 26,4% de G e 21,58% de T.

A análise de saturação não sugeriu a presença de saturação nas substituições nucleotídicas, indicando ausência de substituições múltiplas (Figura 17), este resultado é obtido, normalmente, em análises de genes conservados (Simon *et al.* 1994).

Nas análises filogenéticas com o gene EF-1 α F1, apenas as espécies *D. hydei* e *D. mojavensis* foram utilizadas como grupos externos, devido à grande divergência genética de *D. virilis*. Esta precaução foi tomada para evitar um viés nas topologias condicionado por uma heterogeneidade excessiva dos ramos, isto é, evitar a formação de falsos agrupamentos devido à atração de ramos excessivamente longos ou curtos, condição denominada *long-branch attraction artefact* ou *short-branch attraction artefact* (Philippe *et al.* 2005). Estudos recentes evidenciam que não só a MP é sensível à atração de ramos longos, mas também a BI (Kolaczkowski & Thornton, 2009).

Figura 17 – Taxas de transições (S) e transversões (V) representadas graficamente contra a distância genética entre as sequências parciais do gene EF-1 α F1 para as espécies do complexo *D. buzzatii*, calculada com base no modelo de substituição nucleotídica de Tamura-Nei (Tamura & Nei, 1993).

4.3.1. Máxima parcimônia

Para a análise filogenética por MP foram considerados 35 caracteres variáveis (incluindo os grupos externos), sendo 11 parcimoniosamente informativos.

A busca heurística, com todos os caracteres informativos pesados igualmente, resultou em 370 filogenias igualmente parcimoniosas, sendo que o cladograma consenso apresentou 40 passos, CI = 0,95 e RI = 0,951 (Figura 18).

Análise filogenética por MP das sequências do gene EF-1 α F1 foi pouco informativa, não definindo relações intra e inter-cluster. Resultados obtidos:

- O monofiletismo do complexo *D. buzzatii* com alto suporte estatístico (BP = 70%);
- As espécies do complexo *D. buzzatii* foram agrupadas em uma politomia, destacando-se a manutenção da espécie *D. starmeri* como espécie irmã de *D. venezolana* (BP = 88%).

Figura 18 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene EF-1 α F1, segundo máxima parcimônia. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos do consenso estrito.

4.3.2. Máxima verossimilhança

O modelo TrN + I (Tamura e Nei, 1993) foi selecionado, utilizando o critério Akaike (AIC), como o modelo de substituição nucleotídica que melhor explica a evolução das sequências do gene EF-1 α F1.

A árvore ótima obtida por ML (Figura 19) teve como valor de verossimilhança $-\ln L = 881.09694$. Sua topologia foi similar à obtida por MP e, portanto, foi pouco esclarecedora. Principais resultados:

- O monofiletismo do complexo *D. buzzatii*, neste caso com suporte estatístico inferior (BP = 64%);
- O posicionamento de *D. starmeri* como espécie irmã de *D. venezolana* (BP = 88%).
- Apesar de não definir níveis hierárquicos, o cluster *D. buzzatii* foi sugerido como monofilético, entretanto, com baixo suporte estatístico (BP = 55%).

Figura 19 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene EF-1 α F1, segundo máxima verossimilhança. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos da árvore ótima. A escala indica número de substituições/sítio.

4.3.3. Inferência bayesiana

A topologia obtida por BI a partir da análise das sequências do gene EF-1 α F1 (Figura 20) foi congruente com as topologias obtidas pelos critérios de otimização anteriores.

Figura 20 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene EF-1 α F1, segundo inferência bayesiana. Valores das probabilidades posteriores bayesianas indicadas acima dos ramos. A escala indica número de substituições/sítio.

4.4. Transformer

As sequências parciais obtidas para o gene *transformer* foram de 409 pb (Apêndice 1D), sendo a primeira base correspondente ao 427º nucleotídeo do gene *tra* de *D. virilis* (O’Neil & Belote, 1992, número de acesso X66528.1). O trecho sequenciado está inteiramente dentro do segundo éxon. A proporção média de bases nitrogenadas das sequências nucleotídicas foi 25,19% de A, 29,74% de C, 26,63% de G e 18,44% de T.

A análise de saturação revelou que, na maioria dos casos, tanto as transições quanto as transversões são informativas, detectando-se um início de saturação apenas na curva de transições, quando a divergência das sequências ultrapassou 22%, aproximadamente (Figura 21).

Nas análises filogenéticas com o gene *transformer*, apenas as espécies *D. hydei* e *D. mojavensis* foram utilizadas como grupos externos, pelo mesmo motivo exposto para o gene EF-1 α F1. Dificuldades na obtenção de sequências para a espécie *D. richardsoni*, impediram sua utilização nas análises filogenéticas com este gene.

Figura 21 – Taxas de transições (S) e transversões (V) representadas graficamente contra a distância genética entre as sequências parciais do gene *transformer* para as espécies do complexo *D. buzzatii*, calculada com base no modelo de substituição nucleotídica de Tamura-Nei (Tamura & Nei, 1993), que incorpora as características do modelo K81uf + G, não disponível na base de dados do programa DAMBE v. 4.2.13 (Xie & Xia, 2001).

4.4.1. Máxima parcimônia

Para a análise filogenética por MP foram considerados 168 caracteres variáveis (incluindo os grupos externos), sendo 109 parcimoniosamente informativos.

A busca heurística, com todos os caracteres informativos pesados igualmente, resultou em 18 filogenias igualmente parcimoniosas, sendo que o cladograma consenso apresentou 272 passos, CI = 0,757 e RI = 0,812 (Figura 22).

Os principais resultados da análise filogenética por MP das sequências do gene *transformer* foram:

- O monofiletismo do complexo *D. buzzatii* (BP = 100%) e dos clusters *D. martensis* (BP = 80%) e *D. buzzatii* (BP = 90%). A ausência da espécie *D. richardsoni* impossibilita pontuar considerações a respeito do monofiletismo do cluster *D. stalker*, cuja espécie representante utilizada na análise, *D. stalker*, ocupou a posição filogenética mais basal dentro do complexo *D. buzzatii*;
- O cluster *D. martensis* foi posicionado de forma intimamente relacionada ao cluster *D. buzzatii*, no entanto, com suporte estatístico não significativo (BP = 51%);
- *Drosophila starmeri* e *D. venezolana* foram alocadas novamente como espécies irmãs, e *D. martensis* foi posicionada como espécie mais basal dentro cluster *D. martensis*;
- O cluster *D. buzzatii* foi dividido em duas linhagens evolutivas, uma representada pela espécie *D. buzzatii* (BP = 100%) e outra, com baixo suporte estatístico (BP = 50%), pelas demais espécies do cluster (*D. serido* “sibling set”): *D. koepferae*, *D. antonietae*, *D. serido*, *D. gouveai*, *D. borborema* e *D. seriema*, agrupadas de forma politômica.

Figura 22 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene *transfronter*, segundo máxima parcimônia. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos do consenso estrito.

4.4.2. Máxima verossimilhança

O modelo K81uf + G (K81 *unequal base frequencies*) foi selecionado, utilizando o critério Akaike (AIC), como o modelo de substituição nucleotídica que melhor explica a evolução das sequências do gene *transformer*.

A árvore ótima obtida por ML (Figura 23) teve como valor de verossimilhança $-\ln L = 1645.55657$. Os mesmos agrupamentos observados na análise anterior foram obtidos na análise por ML das sequências do gene *transformer*, diferindo apenas nos valores de suporte estatístico.

Figura 23 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene *transformer*, segundo máxima verossimilhança. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos da árvore ótima. A escala indica número de substituições/sítio.

4.4.3. Inferência bayesiana

A topologia obtida por BI a partir da análise das sequências do gene *transformer* (Figura 24) foi congruente com as topologias obtidas pelos critérios de otimização

anteriores, entretanto, apresentou menor resolução em sua base, não esclarecendo as relações inter-cluster.

Figura 24 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene *transfomer*, segundo inferência bayesiana. Valores das probabilidades posteriores bayesianas indicadas acima dos ramos. A escala indica número de substituições/sítio.

4.5. *Period*

As sequências parciais obtidas para o gene *period* foram de 443 pb (Apêndice 1E). Os fragmentos correspondem às regiões conservadas C4 e C5, intercaladas pelas regiões não conservadas N3 e N4, dentro do domínio CCID do gene nuclear *period* (Figura 7). A primeira base das sequências corresponde ao 2353º nucleotídeo do gene *period* de *D. mojavensis* (número de acesso XM_002010950.1). Em relação à proporção média de bases nitrogenadas, essas sequências nucleotídicas apresentaram 22,13% de A, 29,88% de C, 26,4% de G e 21,58% de T.

A análise de saturação revelou que, na maioria dos casos, tanto as transições quanto as transversões são informativas, detectando-se um início de saturação apenas na

curva de transições, quando a divergência das sequências ultrapassou 14% (Figura 25), aproximadamente.

Nas análises filogenéticas com o gene *period*, apenas as espécies *D. hydei* e *D. mojavensis* foram utilizadas como grupos externos, pelo mesmo motivo exposto para os dois genes anteriores. Dificuldades na obtenção de sequências para as espécies *D. richardsoni* e *D. starmeri*, impediram a utilização das mesmas nas análises filogenéticas com o gene em questão.

Figura 25 – Taxas de transições (S) e transições (V) representadas graficamente contra a distância genética entre as sequências parciais do gene *transformer* para as espécies do complexo *D. buzzatii*, calculada com base no modelo de substituição nucleotídica de Tamura-Nei (Tamura & Nei, 1993), que incorpora as características do modelo HKY + G (Hasegawa *et al.* 1985), não disponível na base de dados do programa DAMBE v. 4.2.13 (Xie & Xia, 2001).

4.5.1. Máxima parcimônia

Para a análise filogenética por MP foram considerados 122 caracteres variáveis (incluindo os grupos externos), sendo 68 parcimoniosamente informativos.

A busca heurística, com todos os caracteres informativos pesados igualmente, resultou em 20 filogenias igualmente parcimoniosas, sendo que o cladograma consenso apresentou 198 passos, CI = 0,737 e RI = 0,708 (Figura 26).

A análise filogenética por MP das sequências do gene *period* foi pouco informativa, permitindo poucas inferências filogenéticas, das quais se pode destacar:

- O monofiletismo do complexo *D. buzzatii* (BP = 83%);

- A divisão do complexo *D. buzzatii* em duas linhagens evolutivas, a mais basal representada pelo clado que reuniu as duas espécies analisadas do cluster *D. martensis*, *D. martensis* e *D. venezolana* (BP = 52%), e a segunda representada pelo clado politômico que agrupou a única espécie do cluster *D. stalker*i utilizada, *D. stalker*i, e todas as espécies do cluster *D. buzzatii* (BP = 63%);
- Algumas relações de proximidade filogenética foram estabelecidas dentro do clado politômico. As espécies *D. gouveai*, *D. borborema* e *D. seriema* foram agrupadas em um mesmo ramo (BP = 93%), compartilhando um ancestral comum, e os representantes das quatro localidades utilizadas da espécie *D. serido* também foram reunidos em um mesmo ramo, porém, com valor não significativo de *bootstrap* (BP < 50%).

Figura 26 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene *period*, segundo máxima parcimônia. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos do consenso estrito.

4.5.2. Máxima verossimilhança

O modelo HKY + G (Hasegawa *et al.* 1985) foi selecionado, utilizando o critério Akaike (AIC), como o modelo de substituição nucleotídica que melhor explica a evolução das sequências do gene *period*.

A árvore ótima obtida por ML (Figura 27) teve como valor de verossimilhança $-\ln L = 1638.96702$. Os principais resultados da análise das sequências do gene *period* por ML foram:

- O monofiletismo do complexo *D. buzzatii* e dos clusters *D. martensis* e *D. buzzatii*, nos três casos com valores não significativos de *bootstrap*. Novamente a ausência da espécie *D. richardsoni* impossibilita pontuar considerações a respeito do monofiletismo do cluster *D. stalker*i;
- A linhagem evolutiva representada pelo clado que reuniu *D. martensis* e *D. venezolana* foi sugerida como a mais basal dentro do complexo *D. buzzatii*;
- Assim como na análise anterior, contrastando com os dados de inversões cromossômicas, foi sugerida maior proximidade filogenética entre os clusters *D. stalker*i e *D. buzzatii*, uma vez que *D. stalker*i foi posicionada como espécie irmã do ramo que reúne as espécies do cluster *D. buzzatii*;
- O cluster *D. buzzatii* foi dividido em uma tricotomia, dois clados com altos suportes estatísticos (BP = 94% e BP = 84%), e um com valor não significativo de *bootstrap* (BP = 53%);
- No primeiro clado (BP = 53%), *D. koepferae* foi alocada como espécie irmã de *D. buzzatii*, e o ramo que agrupou os quatro representantes da espécie *D. serido* foi sugerido como grupo irmão do ramo que reuniu as duas espécies anteriores;
- Os dois representantes da espécie *D. antonietae* foram agrupados no segundo clado, com alto suporte estatístico (BP = 94%);
- *D. gouveai*, *D. borborema* e *D. seriema* foram reunidas no terceiro clado (BP = 84%), com *D. gouveai* ocupando uma posição basal e *D. borborema* e *D. seriema* apresentando monofilia recíproca, entretanto, com baixo suporte estatístico.

Figura 27 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene *period*, segundo máxima verossimilhança. Valores de *bootstrap* em porcentagem indicados acima dos ramos com suporte igual ou superior a 50%, concordantes com os ramos da árvore ótima. A escala indica número de substituições/sítio.

4.5.3. Inferência bayesiana

Os mesmos agrupamentos observados na análise por ML foram obtidos na análise por BI das sequências do gene *period* (Figura 28), diferindo apenas nos valores de suporte dos ramos.

Figura 28 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise de sequências parciais do gene *period*, segundo inferência bayesiana. Valores das probabilidades posteriores bayesianas indicadas acima dos ramos. A escala indica número de substituições/sítio.

4.6. Teste de congruência

Os resultados dos testes de ILD (*Incongruence Length Difference test*) (Farris *et al.*, 1995) ou PHT (*Partition Homogeneity Test*) para os cinco grupos de dados combinados dois a dois estão apresentados na Tabela 3. Valores de $p < 0,05$ indicam a presença de dados incongruentes entre as partições analisadas.

Tabela 3 – Valores de p resultantes do teste de ILD entre os grupos de dados utilizados neste estudo, combinados dois a dois.

	COI	COII	EF-1 α F1	<i>transformer</i>	<i>period</i>
COI	--	0,001	0,116	0,005	0,001
COII		--	0,159	0,001	0,001
EF-1 α F1			--	0,01	0,013
<i>transformer</i>				--	0,017
<i>period</i>					--

Um novo teste ILD foi realizado entre os dados dos genes COI e COII, porém, com as transições excluídas da análise. Não foi sugerida incongruência significativa entre as partições mitocondriais ($p = 0,124$) neste segundo teste. Foram feitos também testes de homogeneidade entre as sequências parciais dos três genes nucleares (EF-1 α F1, *transformer* e *period*) e entre todos os genes. Foi detectada heterogeneidade significativa entre as partições nucleares ($p = 0,005$) e entre todas as partições gênicas ($p = 0,001$), inclusive com as transições dos genes mitocondriais eliminadas da análise ($p = 0,005$).

4.7. Análises combinadas

4.7.1. mtDNA

A combinação das sequências parciais dos genes COI e COII resultou em uma matriz de 1.155 caracteres. A análise por BI dos dados mitocondriais concatenados (Figura 29) teve como principais resultados:

- O monofiletismo do complexo *D. buzzatii* e do cluster *D. buzzatii*, com valores máximos de suporte probabilístico;

- Assim como nas análises filogenéticas individuais do gene COI, os clusters *D. stalker* e *D. martensis* não apresentaram monofilia recíproca e, portanto, as relações entre os clusters não foram resolvidas;
- Em concordância com as análises individuais do gene COII, a linhagem evolutiva representada pela espécie *D. koepferae* ocupou a posição filogenética mais basal dentro do cluster *D. buzzatii*, que foi dividido em mais três linhagens evolutivas, a segunda representada pela espécie *D. buzzatii*, a terceira pela espécie *D. antonietae* e a última pelo ramo que agrupou as espécies do clado AB: *D. serido*, *D. gouveai*, *D. borborema* e *D. seriema*, com *D. borborema* ocupando a posição mais basal dentro deste clado.

Figura 29 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise combinada de sequências parciais dos genes mitocondriais COI e COII, segundo inferência bayesiana. Valores das probabilidades posteriores bayesianas indicados acima dos ramos. A escala indica número de substituições/sítio.

4.7.2. nDNA

A combinação das sequências parciais dos genes EF-1 α F1, *transformer* e *period* resultou em uma matriz de 1.300 caracteres. A análise por BI dos dados nucleares concatenados (Figura 30) teve como principais resultados:

- O monofiletismo do complexo *D. buzzatii* (PP = 99%) e dos clusters *D. martensis* (PP = 83%) e *D. buzzatii* (PP = 71);
- Assim como nas análises individuais do gene *period*, foi sugerida maior proximidade filogenética entre os clusters *D. stalker*i e *D. buzzatii*;
- O cluster *D. buzzatii* foi dividido em três linhagens evolutivas, a mais basal representada pela espécie *D. buzzatii* (PP = 100%), a segunda pelo clado formado pelas espécies *D. koepferae* e *D. serido* (PP = 100%), e a terceira pelo clado que reuniu as espécies *D. antonietae*, *D. gouveai*, *D. borborema* e *D. seriema* (PP = 93%);
- No terceiro clado, formado pelas espécies *D. antonietae*, *D. gouveai*, *D. borborema* e *D. seriema*, *D. antonietae* ocupou a posição mais basal, *D. gouveai* foi alocada de forma intermediária, e as espécies *D. borborema* e *D. seriema* foram as mais derivadas filogeneticamente.

Figura 30 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise combinada de sequências parciais dos genes nucleares EF-1 α F1, *transformer* e *period*, segundo inferência bayesiana. Valores das probabilidades posteriores bayesianas indicados acima dos ramos. A escala indica número de substituições/sítio.

4.7.3. mtDNA + nDNA

A combinação das sequências parciais de todos os genes estudados no presente trabalho, COI, COII, EF-1 α F1, *transformer* e *period*, resultou em uma matriz de 2.455 caracteres. A análise por BI dos dados mitocondriais e nucleares concatenados (Figura 31) teve como principais resultados:

- O monofiletismo do complexo *D. buzzatii* e do cluster *D. buzzatii*, com valores máximos de suporte probabilístico;
- *D. richardsoni* ocupou a posição mais basal dentro do complexo *D. buzzatii*;
- A topologia apresentou uma pobre resolução em sua base, agrupando *D. martensis*, *D. starkeri* e as espécies irmãs *D. starkeri* e *D. venezolana* em uma politomia;

- As relações de parentesco entre as espécies do cluster *D. buzzatii* foram resolvidas em todos os casos com monofilia recíproca;
- *D. buzzatii* ocupou a posição mais basal dentro do cluster *D. buzzatii*, estando proximamente relacionada à espécie *D. koepferae*. *D. antonietae* foi posicionada de forma intermediária às espécies *D. koepferae* e *D. serido*, sendo a última alocada como táxon irmão do ramo formado pelas espécies *D. gouveai*, *D. borborema* e *D. seriema*, com *D. gouveai* ocupando a posição mais basal deste ramo e *D. borborema* e *D. seriema* sugeridas como espécies irmãs;
- No ramo formado pelos indivíduos da espécie *D. serido*, os representantes das localidades ao norte da distribuição da espécie (N45 – Mucugê-BA e 1431 – Milagres-BA) ocuparam posições filogenéticas mais basais, e os representantes das localidades sul (N20 – Arraial do Cabo-RJ e N49 - Bertioga-SP), posições derivadas.

Figura 31 – Relações filogenéticas entre as espécies do complexo *D. buzzatii* inferidas a partir da análise combinada de sequências parciais dos genes COI, COII, EF-1 α F1, *transformer* e *period*, segundo inferência bayesiana. Valores das probabilidades posteriores bayesianas indicados acima dos ramos. A escala indica número de substituições/sítio.

4.8. Testes de neutralidade

4.8.1. *Transfromer*

O teste Z para detecção de seleção natural, baseado no método desenvolvido por Nei e Goboiori (Nei & Goboiori, 1986), aplicado para as sequências parciais do gene *transfromer*, sugeriu que estas estão sob seleção purificadora (Tabela 4). Para confirmar esse resultado, a hipótese nula de evolução neutra foi testada também com uma metodologia de análise Bayesiana, que permite a identificação de seleção natural em aminoácidos específicos do gene (Stern *et al.* 2007). Utilizando o modelo M8 (Yang *et al.* 2000), foi detectada razão entre dn e ds (ω) maior do que um, indicativo de seleção direcional, para os seguintes aminoácidos: 7, 40, 50, 71, 76, 79, 85, 86, 89, 106, 107, 129, 130 e 131 (Tabela 5). No entanto, o limite mínimo do intervalo de confiança de ω foi menor do que um para todos eles (Tabela 5), em não concordância com a hipótese de seleção direcional (Stern *et al.* 2007). Além disso, quando os resultados do modelo M8 foram comparados com os resultados baseados nos modelos M8a (Swanson *et al.* 2003) e M7 (Yang *et al.* 2000), pelo teste da razão de máxima verossimilhança, o sinal de seleção positiva nos aminoácidos citados anteriormente não foi estatisticamente significativo. Todos os aminoácidos restantes apresentaram ω e o limite inferior do intervalo de confiança menores que um (Tabela 5), com grande parte deles com todo intervalo de confiança não incluindo um, o que pode ser interpretado como um indicativo de seleção purificadora, estando de acordo com os resultados do teste Z (Tabela 4).

Tabela 4 – Resultados do Teste-Z de seleção natural, baseado no método de Nei e Goboiori (Nei & Goboiori, 1986), aplicado às sequências parciais do gene *transfromer* das espécies do complexo *buzzatii*.

	Teste-Z de seleção (H0: dn=ds)	dn-ds	Significância
Fragmento total			
	Teste de Neutralidade (HA: dN≠dS)	-5.73	p < 0.0001 (rejeito H0)
	Seleção Positiva (HA: dN > dS)	-5.83	p = 1.0000 (aceito H0)
	Seleção Purificadora (HA: dN < dS)	-5.76	p < 0.0001 (rejeito H0)

Tabela 5 – Valores da razão dn/ds (ω) para cada códon do fragmento do gene *transformer* analisado. AA = amino ácido. IC = Intervalo de Confiança de ω .

Posição	AA	ω	IC	Posição	AA	ω	IC
1	T	0.16	(0,0016;0,54)	44	R	0.13	(0,0016;0,54)
2	R	0.13	(0,0016;0,54)	45	E	0.49	(0,038;1,5)
3	R	0.62	(0,038;1,5)	46	R	0.13	(0,0016;0,54)
4	R	0.44	(0,038;1,5)	47	P	0.14	(0,0016;0,54)
5	A	0.54	(0,038;1,5)	48	R	0.49	(0,038;1,5)
6	R	0.14	(0,0016;0,54)	49	R	0.13	(0,0016;0,54)
7	P	1.1	(0,13;1,5)	50	A	1.3	(0,32;1,5)
8	K	0.14	(0,0016;0,54)	51	R	0.13	(0,0016;0,54)
9	D	0.5	(0,038;1,5)	52	S	0.14	(0,0016;0,54)
10	K	0.22	(0,0016;1,5)	53	R	0.14	(0,0016;0,54)
11	E	0.21	(0,0016;1,5)	54	S	0.16	(0,0016;0,54)
12	K	0.14	(0,0016;0,54)	55	Y	0.14	(0,0016;0,54)
13	V	0.18	(0,0016;1,5)	56	S	0.16	(0,0016;0,54)
14	P	0.14	(0,0016;0,54)	57	V	0.62	(0,038;1,5)
15	Y	0.15	(0,0016;0,54)	58	E	0.13	(0,0016;0,54)
16	F	0.17	(0,0016;1,5)	59	R	0.14	(0,0016;0,54)
17	A	0.16	(0,0016;0,54)	60	N	0.12	(0,0016;0,54)
18	D	0.12	(0,0016;0,54)	61	C	0.46	(0,038;1,5)
19	E	0.14	(0,0016;0,54)	62	C	0.94	(0,13;1,5)
20	A	0.52	(0,038;1,5)	63	H	0.92	(0,13;1,5)
21	R	0.13	(0,0016;0,54)	64	R	0.18	(0,0016;1,5)
22	E	0.13	(0,0016;0,54)	65	R	0.13	(0,0016;0,54)
23	R	0.14	(0,0016;0,54)	66	R	0.14	(0,0016;0,54)
24	D	0.12	(0,0016;0,54)	67	R	0.16	(0,0016;0,54)
25	R	0.18	(0,0016;1,5)	68	S	0.15	(0,0016;0,54)
26	V	0.18	(0,0016;1,5)	69	R	0.13	(0,0016;0,54)
27	R	0.2	(0,0016;1,5)	70	S	0.11	(0,0016;0,54)
28	N	0.12	(0,0016;0,54)	71	Y	1.4	(0,54;1,5)
29	L	0.21	(0,0016;1,5)	72	E	0.13	(0,0016;0,54)
30	R	0.17	(0,0016;1,5)	73	R	0.43	(0,038;1,5)
31	L	0.5	(0,038;1,5)	74	R	0.13	(0,0016;0,54)
32	R	0.13	(0,0016;0,54)	75	H	0.12	(0,0016;0,54)
33	K	0.13	(0,0016;0,54)	76	D	1.3	(0,32;1,5)
34	T	0.14	(0,0016;0,54)	77	S	0.76	(0,076;1,5)
35	S	0.17	(0,0016;1,5)	78	R	0.14	(0,0016;0,54)
36	T	0.51	(0,038;1,5)	79	Y	1.2	(0,21;1,5)
37	T	0.5	(0,038;1,5)	80	K	0.13	(0,0016;0,54)
38	R	0.14	(0,0016;0,54)	81	S	0.97	(0,076;1,5)
39	P	0.14	(0,0016;0,54)	82	A	0.52	(0,038;1,5)
40	S	1.5	(0,54;1,5)	83	T	1	(0,13;1,5)
41	T	0.14	(0,0016;0,54)	84	T	0.98	(0,13;1,5)
42	P	0.54	(0,038;1,5)	85	A	1.3	(0,32;1,5)
43	P	0.58	(0,038;1,5)	86	A	1.3	(0,32;1,5)

Tabela 5 – Continuação.

Posição	AA	ω	IC	Posição	AA	ω	IC
87	T	0.99	(0,13;1,5)	110	T	0.16	(0,0016;0,54)
88	K	0.14	(0,0016;0,54)	111	P	0.14	(0,0016;0,54)
89	T	1.4	(0,32;1,5)	112	R	0.14	(0,0016;0,54)
90	R	0.97	(0,13;1,5)	113	I	0.14	(0,0016;0,54)
91	R	0.17	(0,0016;1,5)	114	I	0.14	(0,0016;0,54)
92	R	0.13	(0,0016;0,54)	115	T	0.15	(0,0016;0,54)
93	L	0.5	(0,038;1,5)	116	V	0.18	(0,0016;1,5)
94	S	0.12	(0,0016;0,54)	117	P	0.16	(0,0016;0,54)
95	R	0.44	(0,038;1,5)	118	V	0.19	(0,0016;1,5)
96	S	0.12	(0,0016;0,54)	119	P	0.14	(0,0016;0,54)
97	R	0.14	(0,0016;0,54)	120	V	0.19	(0,0016;1,5)
98	N	0.8	(0,13;1,5)	121	P	0.14	(0,0016;0,54)
99	R	0.16	(0,0016;1,5)	122	A	0.14	(0,0016;0,54)
100	S	0.35	(0,038;1,5)	123	A	0.16	(0,0016;0,54)
101	R	0.16	(0,0016;1,5)	124	D	0.12	(0,0016;0,54)
102	S	0.42	(0,038;1,5)	125	Y	0.14	(0,0016;0,54)
103	R	0.19	(0,0016;1,5)	126	P	0.16	(0,0016;0,54)
104	S	0.64	(0,038;1,5)	127	Y	0.15	(0,0016;0,54)
105	R	0.27	(0,0016;1,5)	128	A	0.16	(0,0016;0,54)
106	S	1.2	(0,21;1,5)	129	P	1.5	(0,54;1,5)
107	R	1.2	(0,21;1,5)	130	L	1.5	(0,54;1,5)
108	S	0.14	(0,0016;0,54)	131	V	1.1	(0,13;1,5)
109	R	0.18	(0,0016;1,5)				

4.8.2. *Period*

O teste Z aplicado para as sequências parciais do gene *period* também indicou evolução sob seleção purificadora (Tabela 6). Para confirmar o resultado, assim como para o gene *transformer*, a hipótese nula de evolução neutra foi testada pela metodologia de análise Bayesiana (Stern *et al.* 2007). Utilizando o modelo M8 (Yang *et al.* 2000), não foi detectada razão entre dn e ds (ω) maior do que um para nenhum dos aminoácidos (Tabela 7). Os intervalos de confiança compreenderam valores inferiores a um, exceto para dois aminoácidos, 113 e 115. A comparação dos resultados do modelo M8 com os resultados baseados nos modelos M8a (Swanson *et al.* 2003) e M7 (Yang *et al.* 2000), pelo teste de razão de máxima verossimilhança, também não evidenciou sinal de seleção positiva estatisticamente significativo, em concordância com os resultados obtidos por Franco *et al.* (2010) para o gene *period*.

Tabela 6 – Resultados do Teste-Z de seleção natural, baseado no método de Nei e Goboiori (Nei & Goboiori, 1986), aplicado às sequências parciais do gene *period* das espécies do complexo buzzatii.

	Teste-Z de seleção (H0: dn=ds)	dn-ds	Significância
Fragmento total			
	Teste de Neutralidade (HA: dN≠dS)	-11.25	p < 0.0001 (rejeito H0)
	Seleção Positiva (HA: dN > dS)	-10.85	p = 1.0000 (aceito H0)
	Seleção Purificadora (HA: dN < dS)	-10.72	p < 0.0001 (rejeito H0)

Tabela 7 – Valores da razão dn/ds (ω) para cada códon do fragmento do gene *period* analisado. AA = amino ácido. IC = Intervalo de Confiança de ω .

Posição	AA	ω	IC	Posição	AA	ω	IC
1	V	0.036	(1,2e-06;0,36)	34	N	0.17	(0,0076;0,36)
2	G	0.051	(1,2e-06;0,36)	35	L	0.044	(1,2e-06;0,36)
3	V	0.039	(1,2e-06;0,36)	36	N	0.043	(1,2e-06;0,36)
4	G	0.055	(1,2e-06;0,36)	37	C	0.038	(1,2e-06;0,36)
5	V	0.044	(1,2e-06;0,36)	38	A	0.028	(1,2e-06;0,15)
6	G	0.046	(1,2e-06;0,36)	39	Q	0.038	(1,2e-06;0,36)
7	V	0.035	(1,2e-06;0,36)	40	N	0.03	(1,2e-06;0,15)
8	G	0.31	(0,062;0,36)	41	I	0.028	(1,2e-06;0,15)
9	V	0.037	(1,2e-06;0,36)	42	N	0.03	(1,2e-06;0,15)
10	G	0.3	(0,062;0,36)	43	L	0.039	(1,2e-06;0,36)
11	V	0.039	(1,2e-06;0,36)	44	W	0.046	(1,2e-06;0,36)
12	G	0.3	(0,062;0,36)	45	P	0.033	(1,2e-06;0,36)
13	V	0.039	(1,2e-06;0,36)	46	P	0.03	(1,2e-06;0,15)
14	P	0.036	(1,2e-06;0,36)	47	F	0.032	(1,2e-06;0,15)
15	D	0.035	(1,2e-06;0,36)	48	S	0.033	(1,2e-06;0,36)
16	T	0.052	(1,2e-06;0,36)	49	L	0.036	(1,2e-06;0,36)
17	L	0.034	(1,2e-06;0,36)	50	G	0.033	(1,2e-06;0,36)
18	S	0.027	(1,2e-06;0,15)	51	I	0.2	(0,0076;0,36)
19	K	0.035	(1,2e-06;0,36)	52	T	0.028	(1,2e-06;0,15)
20	W	0.046	(1,2e-06;0,36)	53	T	0.17	(0,0076;0,36)
21	Q	0.048	(1,2e-06;0,36)	54	P	0.029	(1,2e-06;0,15)
22	T	0.16	(0,0076;0,36)	55	S	0.031	(1,2e-06;0,15)
23	P	0.034	(1,2e-06;0,36)	56	A	0.23	(0,0076;0,36)
24	N	0.039	(1,2e-06;0,36)	57	H	0.031	(1,2e-06;0,15)
25	A	0.03	(1,2e-06;0,15)	58	S	0.17	(0,0076;0,36)
26	S	0.16	(0,0076;0,36)	59	S	0.032	(1,2e-06;0,15)
27	S	0.27	(0,062;0,36)	60	H	0.032	(1,2e-06;0,36)
28	N	0.21	(0,0076;0,36)	61	T	0.029	(1,2e-06;0,15)
29	Q	0.038	(1,2e-06;0,36)	62	A	0.03	(1,2e-06;0,15)
30	Y	0.18	(0,0076;0,36)	63	V	0.031	(1,2e-06;0,15)
31	V	0.032	(1,2e-06;0,15)	64	A	0.047	(1,2e-06;0,36)
32	A	0.29	(0,062;0,36)	65	Q	0.038	(1,2e-06;0,36)
33	N	0.26	(0,062;0,36)	66	R	0.034	(1,2e-06;0,36)

Tabela 7 – Continuação.

Posição	AA	ω	IC	Posição	AA	ω	IC
67	S	0.032	(1,2e-06;0,15)	108	S	0.032	(1,2e-06;0,15)
68	F	0.033	(1,2e-06;0,36)	109	T	0.027	(1,2e-06;0,15)
69	S	0.032	(1,2e-06;0,15)	110	S	0.032	(1,2e-06;0,15)
70	P	0.029	(1,2e-06;0,15)	111	Q	0.043	(1,2e-06;0,36)
71	Q	0.038	(1,2e-06;0,36)	112	Q	0.038	(1,2e-06;0,36)
72	H	0.032	(1,2e-06;0,36)	113	G	0.9	(0,15;3,1)
73	S	0.047	(1,2e-06;0,36)	114	A	0.27	(0,062;0,36)
74	L	0.029	(1,2e-06;0,15)	115	A	0.64	(0,062;3,1)
75	F	0.034	(1,2e-06;0,36)	116	T	0.27	(0,062;0,36)
76	P	0.038	(1,2e-06;0,36)	117	T	0.03	(1,2e-06;0,15)
77	A	0.026	(1,2e-06;0,15)	118	A	0.042	(1,2e-06;0,36)
78	F	0.032	(1,2e-06;0,15)	119	M	0.027	(1,2e-06;0,15)
79	Y	0.05	(1,2e-06;0,36)	120	P	0.051	(1,2e-06;0,36)
80	Y	0.049	(1,2e-06;0,36)	121	L	0.21	(0,0076;0,36)
81	I	0.044	(1,2e-06;0,36)	122	Q	0.038	(1,2e-06;0,36)
82	P	0.032	(1,2e-06;0,36)	123	Y	0.047	(1,2e-06;0,36)
83	A	0.029	(1,2e-06;0,15)	124	M	0.027	(1,2e-06;0,15)
84	P	0.032	(1,2e-06;0,15)	125	A	0.026	(1,2e-06;0,15)
85	L	0.039	(1,2e-06;0,36)	126	G	0.035	(1,2e-06;0,36)
86	A	0.026	(1,2e-06;0,15)	127	V	0.045	(1,2e-06;0,36)
87	N	0.043	(1,2e-06;0,36)	128	M	0.027	(1,2e-06;0,15)
88	A	0.036	(1,2e-06;0,36)	129	Y	0.037	(1,2e-06;0,36)
89	T	0.029	(1,2e-06;0,15)	130	P	0.046	(1,2e-06;0,36)
90	A	0.26	(0,062;0,36)	131	H	0.046	(1,2e-06;0,36)
91	A	0.033	(1,2e-06;0,36)	132	P	0.033	(1,2e-06;0,36)
92	A	0.029	(1,2e-06;0,15)	133	S	0.039	(1,2e-06;0,36)
93	P	0.16	(0,0076;0,36)	134	L	0.029	(1,2e-06;0,15)
94	S	0.19	(0,0076;0,36)	135	F	0.032	(1,2e-06;0,15)
95	V	0.029	(1,2e-06;0,15)	136	Y	0.053	(1,2e-06;0,36)
96	S	0.22	(0,0076;0,36)	137	T	0.029	(1,2e-06;0,15)
97	P	0.033	(1,2e-06;0,36)	138	H	0.046	(1,2e-06;0,36)
98	Q	0.038	(1,2e-06;0,36)	139	P	0.034	(1,2e-06;0,36)
99	R	0.035	(1,2e-06;0,36)	140	A	0.038	(1,2e-06;0,36)
100	S	0.045	(1,2e-06;0,36)	141	A	0.049	(1,2e-06;0,36)
101	H	0.031	(1,2e-06;0,15)	142	A	0.043	(1,2e-06;0,36)
102	K	0.17	(0,0076;0,36)	143	A	0.03	(1,2e-06;0,15)
103	T	0.33	(0,15;0,36)	144	A	0.033	(1,2e-06;0,36)
104	C	0.038	(1,2e-06;0,36)	145	T	0.049	(1,2e-06;0,36)
105	D	0.22	(0,0076;0,36)	146	A	0.026	(1,2e-06;0,15)
106	Q	0.038	(1,2e-06;0,36)	147	M	0.027	(1,2e-06;0,15)
107	P	0.032	(1,2e-06;0,36)				

4.9. Tempos de divergência

Os testes de razão de verossimilhança (LRT) revelaram que, para o complexo *D. buzzatii*, apenas as taxas de substituição do gene COI são estatisticamente homogêneas ao longo dos ramos e, portanto, a aplicação da hipótese de relógio molecular é apropriada. As demais partições gênicas estudadas apresentaram taxas evolutivas variando de forma estatisticamente significativa entre os ramos.

Para as estimativas de tempo de divergência com o programa PATHd8 e as sequências do gene COI, foi obtida uma nova árvore por ML (Figura 32), utilizando adicionalmente sequências parciais retiradas do *GeneBank* das espécies *D. melanogaster* (número de acesso GQ229519.1), *D. yakuba* (número de acesso NC_001322), *Scaptodrosophila pattersoni* (número de acesso EU493687.1) e *S. latifasciaeformis* (número de acesso EU493684.1). Foi definida como idade fixa a divergência entre *D. melanogaster* e *D. yakuba*, estimada em 6,1 milhões de anos (Russo *et al.* 1995; Nagaraja *et al.* 2004, Franco, 2009).

Figura 32 – Árvore obtida segundo máxima verossimilhança. Nova análise do gene COI utilizando adicionalmente sequências retiradas do *GeneBank* das espécies *D. melanogaster*, *D. yakuba*, *Scaptodrosophila pattersoni* e *S. latifasciaeformis*. Ponto de calibração indicado acima do nó entre *D. melanogaster* e *D. yakuba*. A escala refere-se ao número de substituições/sítio.

Os cálculos de tempo de divergência realizados a partir da análise de sequências parciais do gene COI, baseada na hipótese de relógio molecular, indicam que o tempo para o ancestral comum mais recente das espécies do cluster *D. buzzatii* é de 3,98 milhões de anos, com um desvio de $\pm 1,12$ milhão de anos. Os cálculos ainda sugerem que as espécies *D. buzzatii* e *D. koepferae* divergiram há cerca de 2,61 milhões de anos, com um desvio de $\pm 0,87$ milhão de anos; o tempo de divergência entre *D. antonietae* e o ancestral comum das espécies do clado AB é cerca de 1,27 milhão de anos, com desvio de $\pm 0,55$ milhão de anos; e o tempo para o ancestral comum mais recente das espécies do clado AB é de 0,76 milhão de anos, desvio de $\pm 0,34$ milhão de anos.

5. DISCUSSÃO

5.1. Análise filogenética do complexo *Drosophila buzzatii*

Análises individuais

Em todas as hipóteses filogenéticas, elaboradas a partir das análises individuais dos genes COI, COII, EF-1 α F1, *transformer* e *period*, foi sugerido o monofiletismo do complexo *D. buzzatii*, em concordância com estudos citogenéticos e moleculares prévios (Ruiz & Wasserman, 1993; Spicer, 1995; Rodriguez-Trelles *et al.* 2000; Durando *et al.* 2000). As hipóteses filogenéticas obtidas a partir das análises individuais dos genes COI, COII e *transformer* também confirmaram o monofiletismo do cluster *D. buzzatii*, com suportes estatísticos e probabilísticos robustos, em concordância com estudos moleculares anteriores com o grupo (Manfrin *et al.* 2001, Franco *et al.* 2010).

As relações de parentesco entre os clusters *D. stalker*i, *D. martensis* e *D. buzzatii* não foram totalmente esclarecidas, pois a base das topologias obtidas nas análises individuais foram bastante incongruentes, não sendo observada monofilia recíproca entre os clusters *D. stalker*i e *D. martensis* na maioria das hipóteses filogenéticas. Existem pelo menos duas alternativas, não mutuamente exclusivas, para explicar este fato: manutenção de polimorfismo ancestral ou *lineage sorting* incompleto e hibridação introgressiva (Machado *et al.* 2002; Funk e Omland, 2003; Hey, 2006). Entretanto, considerando a distribuição geográfica atual das espécies desses dois clusters, é pouco provável que os resultados observados estejam associados a eventos de hibridação introgressiva.

De acordo com Hillis & Huelsenbeck (1992), sequências de DNA ou de outro tipo de dado molecular, quando comparadas entre organismos, podem conter sinal filogenético ou estar totalmente randomizadas em relação à história filogenética. Com isso, a pobre resolução na base das topologias obtidas nas análises individuais com o gene mitocondrial COI pode ainda estar associada aos maiores níveis de saturação nas substituições de suas sequências, em especial nas transições, como foi possível observar a partir do gráfico da análise de saturação (Figuras 8) e dos índices de consistência (CI) e de retenção (RI) nas análises por MP com este gene. A problemática de se trabalhar com genes com altas taxas de substituição e, consequentemente, maior ocorrência de substituições múltiplas e homoplasias, fica evidente quando são comparadas as hipóteses filogenéticas obtidas a partir das análises do gene COI por MP com todos os caracteres informativos pesados igualmente e com pesagem diferencial entre transições

e transversões. No primeiro caso, as espécies *D. stalker*i e *D. martensis* ocuparam a posição mais basal do complexo de forma politômica, e o clado formado pelas espécies *D. starmeri* e *D. venezolana* foi posicionado como clado irmão do cluster *D. buzzatii*, em ambos os casos com suporte estatístico não significativo (Figura 9). Já na topologia obtida na segunda análise, atribuindo menor peso às transições (saturadas), *D. richardsoni* foi posicionada como a espécie mais basal do complexo, em concordância com hipóteses filogenéticas elaboradas a partir da análise combinada dos genes COI, COII e COIII (Spicer, 1995) e a partir da análise do gene *Xanthine dehidrogenase* (Rodriguez-Trelles *et al.* 2000). Além disso, *D. martensis* foi alocada como espécie irmã do cluster *D. buzzatii*, embora com suporte estatístico não significativo (Figura 10). Essa alternância significativa no posicionamento das espécies dos clusters *D. stalker*i e *D. martensis* entre as duas análises abordadas acima é um indício de que o sinal filogenético do gene COI para a base do complexo *D. buzzatii* está, possivelmente, comprometido pela saturação de suas transições.

Os genes mais informativos para as relações de parentesco inter-cluster foram os genes COII e *transformer*, além de apresentarem os maiores valores de verossimilhança nas topologias ótimas obtidas por ML, depois do gene COI. Entretanto, suas topologias foram bastante incongruentes.

Nas análises por MP e ML do gene COII, os clusters *D. stalker*i e *D. martensis* foram agrupados em um mesmo clado, porém, com suportes estatísticos não significativos. O cluster *D. stalker*i ocupou uma posição mais basal, constituindo um grupo parafilético dentro do clado, enquanto o cluster *D. martensis* ocupou uma posição mais derivada, constituindo um grupo monofilético. Desta forma, estes clusters seriam filogeneticamente próximos e constituiriam uma linhagem evolutiva coesa, representando o grupo irmão do cluster *D. buzzatii* (Figuras 14 e 15), em desacordo com os dados de inversões cromossômicas, que aproximam os clusters *D. martensis* e *D. buzzatii* (Wasserman, 1982; Ruiz & Wasserman, 1993).

Nas topologias obtidas a partir da análise do gene *transformer* por MP e ML, o cluster *D. stalker*i, representado pela espécie *D. stalker*i, ocupou a posição mais basal do complexo *D. buzzatii*, com suporte estatístico máximo (Figuras 22 e 23). O cluster *D. martensis*, por sua vez, foi sugerido como monofilético, sendo posicionado como grupo irmão do cluster *D. buzzatii*, no entanto, com valores não significativos de bootstrap. A ausência da espécie *D. richardsoni* nas análises individuais com o gene *tra* impede que sejam pontuadas considerações a respeito de qual seria a espécie mais basal

do cluster *D. stalker*i e se os resultados deste gene estão em completa concordância com hipóteses filogenéticas de estudos moleculares (Spicer, 1995; Rodriguez-Trelles *et al.* 2000) e citogenéticos (Wasserman, 1982; Ruiz & Wasserman, 1993) anteriores.

O gene conservado EF-1 α F1 não definiu níveis hierárquicos entre os clusters, e a ausência de sequências do gene *period* para as espécies *D. richardsoni* e *D. starmeri* dificultou a comparação das análises individuais deste gene com as dos demais em nível inter-cluster.

Desta forma, não foi possível esclarecer de forma conclusiva, a partir das análises individuais dos genes COI, COII, EF-1 α F1, *transformer* e *period*, as relações de parentesco entre os três clusters que constituem o complexo *D. buzzatii*, uma vez que todas as possibilidades de proximidade filogenética entre os mesmos foram verificadas: clusters *D. martenis*/*D. buzzatii* – gene *transformer* (Figura 23); clusters *D. stalker*i/*D. buzzatii* – gene *period* (Figura 27); clusters *D. stalker*i/*D. martenis* – gene COII (Figura 15).

Ruiz & Wasserman (1993) postulam que o ancestral do complexo *D. buzzatii* teria habitado a região hoje ocupada pelo cluster *D. martenis*. A partir desta linhagem primitiva teria ocorrido uma dispersão inicial de migrantes para as ilhas do Caribe, os quais deram origem ao cluster *D. stalker*i. Posteriormente, a linhagem continental expandiu sua distribuição para o sul da América do Sul, dando origem ao cluster *D. buzzatii*, enquanto ao norte da América do Sul ela continuou diferenciando-se, dando origem ao cluster *D. martensis*. De acordo com este cenário, o cluster *D. stalker*i deveria constituir um grupo monofilético, mas os resultados obtidos a partir de marcadores moleculares mitocondriais (Spicer, 1995, presente trabalho) e nucleares (Rodriguez-Trelles *et al.* 2000) indicam que, na verdade, o cluster *D. stalker*i se trata de um grupo parafilético. Ao menos dois eventos de colonização independentes seriam necessários para dar origem às espécies *D. richardsoni* e *D. stalker*i (Rodriguez-Trelles *et al.* 2000).

Considerando que o grupo *repleta* seja um grupo neotropical originado em regiões desérticas do México (Patterson & Stone, 1952; Ward & Stone, 1952), diante da distribuição geográfica atual das espécies do complexo *D. buzzatti*, outra hipótese plausível é a de diversificação no sentido norte-sul, isto é, o ancestral do complexo *D. buzzatii* teria habitado regiões desérticas do sul da América do Norte, dando origem às espécies do cluster *D. staker*i a partir de eventos de dispersão independentes para

Flórida e Caribe, e, posteriormente, com o avanço da linhagem primitiva para o norte e sul da América do Sul, teriam surgido os clusters *D. martensis* e *D. buzzatii*.

Apesar das relações de parentesco permanecerem em aberto em nível inter-cluster, as análises filogenéticas individuais possibilitaram várias inferências em nível intra-cluster:

Drosophila starmeri e *D. venezolana* foram sugeridas como espécies irmãs, com valores robustos de *bootstrap* ou probabilidades posteriores, em todas as análises filogenéticas em que sequências de ambas foram incluídas, ficando clara a relação de proximidade evolutiva entre as mesmas, em concordância com estudos moleculares anteriores (Spicer, 1995; Rodriguez-Trelles *et al.* 2000; Durando *et al.* 2000). Já o compartilhamento de um ancestral comum entre a linhagem evolutiva representada pelas duas espécies anteriores e *D. martensis* não ficou tão evidente, pois em apenas uma das topologias em que *D. martensis*, *D. starmeri* e *D. venezolana* foram agrupadas em um ramo monofilético (Figuras 14, 15, 22 e 23), com alto suporte estatístico (Figura 22). Entretanto, a distribuição geográfica atual destas três espécies (Vilela, 1983) e o compartilhamento das inversões cromossômicas fixas $3r^2$, $3w$ e $5d^2$ (Ruiz & Wasserman, 1993), somados aos resultados apresentados no presente trabalho e em estudos moleculares anteriores com o grupo (Spicer, 1995; Rodriguez-Trelles *et al.* 2000), podem ser considerados fortes indícios de que as espécies do cluster *D. martensis* aqui estudadas constituem uma linhagem evolutiva coesa.

Para o cluster *D. buzzatii*, foi possível observar nas topologias obtidas a partir das análises individuais, que os dados dos genes mitocondriais, em especial do gene COII, contrastam com os dados do gene *transformer* e de estudos baseados em outros marcadores nucleares (Rodriguez-Trelles *et al.* 2000; Franco *et al.* 2008b) e morfológicos (Moraes *et al.* 2004), que sugerem *D. buzzatii* como espécie mais basal do cluster e *D. koepferae* como pertencente ao “*D. serido* sibling set” (Tidon e Sene, 2001), que inclui todas as espécies do cluster, com exceção de *D. buzzatii*.

Estes padrões contrastantes poderiam ser explicados pelo compartilhamento de polimorfismos entre *D. koepferae* e *D. buzzatii*, esclarecendo, em parte, porque em alguns estudos *D. koepferae* é classificada como espécie irmã de *D. buzzatii* (Durando *et al.* 2000; Manfrin *et al.* 2001) e em outros não (Rodriguez-Trelles *et al.* 2000). Contudo, o tempo de divergência entre *D. buzzatii* e *D. koepferae* estimado no presente trabalho, cerca de 2,61 milhões de anos, com um desvio de $\pm 0,87$ milhão de anos, e em outras datações moleculares, girando em torno de quatro milhões de anos atrás segundo

Gomes & Hasson (2003) e Manfrin & Sene (2006), seria tempo suficiente para perda da maioria dos polimorfismos (Clark, 1997), sobretudo em genes que evoluem sob influencia de seleção purificadora, que reduz o tempo de coalescência em relação aos genes neutros (Nachman, 2006).

Portanto, a similaridade encontrada entre sequências de *D. buzzatii* e *D. koepferae* para o gene COI e o padrão filogenético observado para o gene COII não seriam devido à manutenção de polimorfismo ancestral, mas, possivelmente, resultado de eventos de hibridação introgressiva. Apoiando esta hipótese, existe o fato dessas espécies serem simpátricas em muitas localidades dentro do Domínio do Chaco, e estudos empíricos que indicam que as mesmas são capazes de cruzar em laboratório (Machado *et al.* 2006; Soto *et al.* 2008). Adicionalmente, estudos prévios baseados em dados moleculares dos genes nucleares também sugerem eventos de hibridação introgressiva envolvendo as espécies *D. buzzatii* e *D. koepferae* (Gomez & Hasson, 2003; Piccinali *et al.* 2004; Franco *et al.* 2010).

Para esclarecer se há relação entre os possíveis eventos de hibridação introgressiva relacionados aos padrões observados para os genes COI e COII, análises filogenéticas com um maior número de indivíduos de diferentes populações de *D. buzzatii* e *D. koepferae* são necessárias para que seja possível uma averiguação confiável da existência de uma ou mais linhagens evolutivas para cada gene, em ambas as espécies, assim como em Franco *et al.* (2010) para o gene *period*. Apesar da progênie de machos originada pelo cruzamento interespecífico entre *D. buzzatii* e *D. koepferae* ser estéril (Soto *et al.* 2008), eventos de hibridação introgressiva bidirecional ou unidirecional seriam possíveis mediante retrocruzamentos entre fêmeas híbridas F1 férteis com machos de *D. koepferae* e/ou *D. buzzatii*.

Hibridação introgressiva entre espécies de animais filogeneticamente próximas parece ser um fenômeno mais comum do que se pensava anteriormente (Arnold, 1997; Seehausen, 2004). Diversos exemplos foram relatados na literatura entre espécies do gênero *Drosophila* e em Diptera de forma geral (Wang & Hey, 1996; Manfrin *et al.* 2001; Gomez e Hasson, 2003; Piccinali *et al.* 2004 Franco *et al.* 2006; Mazzoni *et al.* 2006; Franco *et al.* 2010; Kokudai *et al.* 2011).

Outros resultados também sugerem eventos de hibridação introgressiva entre espécies do cluter *D. buzzatii*. Manfrin *et al.* (2001) encontraram haplótipos mitocondriais (COI) característicos de *D. antonietae* em indivíduos de *D. gouveai* provenientes da localidade de Analândia-SP, próxima ao limite nordeste da distribuição

de *D. antonietae*. Diante deste fato, foi proposta a hipótese de contato secundário entre *D. gouveai* e *D. antonietae* na região, seguido de hibridação introgressiva (Manfrin *et al.* 2001), mesmo que restrita aos genes mitocondriais, como sugerido por Franco *et al.* (2006b) e confirmado pela análise de haplótipos do gene nuclear *period* (Franco *et al.* 2010). Segundo Arnold (1992), em eventos de hibridação, mtDNA tende a introgridir mais rapidamente que marcadores nucleares.

Nos possíveis casos de hibridação entre *D. buzzatii* e *D. koepferae* mencionados anteriormente, o processo de introgressão não teria sido restrito aos genes mitocondriais, uma vez que Franco *et al.* (2010) detectou haplótipos do gene nuclear *period* característicos de *D. buzzatii* em uma das duas linhagens evolutivas existentes em *D. koepferae* (Figura 4D). As hipóteses filogenéticas obtidas por ML e BI com gene *period* (Figuras 27 e 28) reforçam essa hipótese, pois *D. buzzatii* e *D. koepferae* foram agrupadas em um ramo monofilético, entretanto, com valores de *bootstrap* e probabilidades posteriores não significativos.

Além disso, as análises individuais com o gene *period*, em desacordo com os dados de DNA mitocondrial obtidos no presente estudo e em Manfrin *et al.* (2001), não corroboraram o clado AB como uma linhagem evolutiva coesa, já que *D. serido* não formou um grupo monofilético com *D. gouveai*, *D. seriema* e *D. borborema* (Figuras 27 e 28), estando mais relacionada filogeneticamente às espécies *D. buzzatii* e *D. koepferae*, assim como o verificado em análises populacionais com o gene *period* por Franco *et al.* (2010). Esta incongruência pode estar relacionada a uma perturbação do sinal filogenético, uma vez que os testes de neutralidade realizados no presente trabalho e por Franco *et al.* (2010) evidenciaram a presença de seleção purificadora atuando como principal força evolutiva sobre os fragmentos analisados do gene *period*.

Considerando o compartilhamento da inversão $2x^7$ entre *D. serido* e *D. antonietae*, a incongruência entre os dados mitocondriais e do gene *period* abordada acima e a pobre resolução das relações de parentesco entre as espécies do “*D. serido* sibling set” (Tidon-Sklorz & Sene, 2001) nas topologias do gene *transformer* (Figuras 21, 22 e 23), o posicionamento filogenético de *D. serido* não pôde ser elucidado de forma conclusiva a partir das análises individuais.

A dificuldade em esclarecer as relações de parentesco desta espécie com as demais do cluster *D. buzzatii*, pode ainda estar relacionada ao padrão politípico apresentado por *D. serido*, como foi verificado em relação aos cromossomos metafásicos (Baimai *et al.* 1983), inversões polimórficas (Ruiz & Wasserman, 1993;

Ruiz *et al.* 2000), morfologia do edeago (Franco *et al.* 2008a) e haplótipos de DNA mitocondrial (de Brito *et al.* 2002a; Morales, 2005).

A hipótese de monofiletismo entre as espécies *D. gouveai*, *D. seriema* e *D. borborema*, baseada em dados de inversões cromossômicas (Ruiz & Wasserman, 1993), do gene mitocondrial COI (Manfrin *et al.* 2001; presente trabalho) e sequências de DNA satélite (Kuhn & Sene, 2005), foi corroborada pelas análises dos genes COII (Figuras 14, 15 e 16) e *period* (Figuras 26, 27 e 28), ficando evidente a relação de proximidade entre estas três espécies, cujas relações de parentesco foram resolvidas com alto suporte estatístico e probabilístico. *Drosophila gouveai* ocupou uma posição mais basal e *D. seriema* e *D. borborema* foram sugeridas como espécies irmãs, em concordância com Moraes *et al.* (2009) e Franco *et al.* (2010).

Finalmente, para o cluster *D. stalker*, devido à pobre resolução das topologias obtidas nas análises com o gene EF-1 α F1 e a ausência da espécie *D. richardsoni* nas análises com os genes *transformer* e *period*, as inferências das relações intra-cluster ficaram restritas às análises dos genes mitocondriais COI e COII. Como foi possível observar, apesar do suporte estatístico não significativo, as análises por MP e ML do gene COII (Figuras 14 e 15) e a análise por MP com pesagem diferencial de caracteres do gene COI (Figura 10) corroboraram o posicionamento de *D. richardsoni* como a espécie mais basal do cluster parafilético *D. stalker*, em concordância com os dados do gene *Xdh* (Rodriguez-Trelles *et al.* 2000), e contrastando com a hipótese de monofilia do cluster *D. stalker* e posicionamento da espécie *D. stalker* como mais basal do cluster, baseada em dados de inversões cromossômicas (Ruiz & Wasserman, 1993). Portanto, novas análises com os genes *transformer*, *period*, ou com outros marcadores nucleares, utilizando ambas as espécies do cluster *D. stalker*, serão necessárias para o esclarecimento dessa aparente incongruência entre os dados de marcadores moleculares e citogenéticos.

Análises combinadas

Os testes de ILD (Farris *et al.*, 1995) não sugeriram heterogeneidade significativa apenas entre os dados dos genes COI e EF-1 α F1, COII e EF-1 α F1 e entre os genes mitocondriais, quando desconsideradas as transições. Foi possível observar a partir dos valores de *p* dos grupos de dados combinados dois a dois (Tabela 3), que não há um

gene em especial a ser considerado como fonte principal da alta incongruência verificada entre os dados.

Os testes de homogeneidade foram realizados em caráter exploratório e não restritivo, pois estudos prévios mostram que mesmo diante de uma provável incongruência, resultados interessantes podem ser obtidos em análises combinadas de grupos de dados aparentemente conflitantes (Remsen & DeSalle, 1998; Remsen & O'Grady, 2002; Driskell *et al.* 2004; Gatesy *et al.* 2004). Além disso, os valores de *p* apresentam baixa correlação com a melhora na resolução filogenética resultante da concatenação (Barker & Lutzoni, 2002). Dessa forma, mesmo diante de heterogeneidade estatisticamente significativa, a concatenação e análise simultânea dos dados dos genes nucleares (EF-1 α F1, *transformer* e *period*) e de todos os genes (COI, COII, EF-1 α F1, *transformer* e *period*) foram realizadas.

A análise simultânea dos dados mitocondriais (Figura 29) representou uma somatória das topologias dos genes COI e COII, e em relação à principal incongruência entre ambos, o posicionamento de *D. koepferae*, houve a manutenção desta como a linhagem mais basal dentro do cluster. Já a resolução da base do cladograma foi condicionada pelos dados do gene COI, sendo mantido o posicionamento das espécies dos clusters *D. stalker*i e *D. martensis* assim como verificado nas análises individuais com este gene.

Por sua vez, a análise combinada com os genes nucleares (Figura 30) também representou uma somatória dos agrupamentos sugeridos em suas análises individuais, sendo mantido o nonofiletismo do cluster *D. martenis*, com forte apoio probabilístico, e a divisão do cluster *D. buzzatii* em duas linhagens evolutivas principais, *D. buzzatii* e “*D. serido* sibling set”, condicionados pelos dados do gene *transformer*, além do posicionamento de *D. stalker*i como espécie irmã do cluster *D. buzzatii* e a não indicação de uma linhagem evolutiva coesa entre *D. serido* e as demais espécies do clado AB, condicionados pelos dados do gene *period*.

A base da topologia obtida a partir da análise combinada com todos os genes, COI, COII, EF-1 α F1, *transformer* e *period* (Figura 31), refletiu a incongruência observada entre as análises individuais dos mesmos, não definindo níveis hierárquicos entre a maioria das espécies dos clusters *D. stalker*i e *D. martensis*, agrupando-as de forma polítômica. Contudo, a topologia apresentou uma alta resolução para o cluster *D. buzzatii*, definindo com suporte probabilístico robusto as relações de parentesco entre

suas espécies, em todos os casos com monofilia recíproca. Devido ao aumento da probabilidade do sinal filogenético se impor sobre o ruído, com o aumento do número de caracteres envolvidos na análise (De Queiroz *et al.* 1995, De Queiroz & Gatesy, 2007), e, conseqüentemente, o aumento da probabilidade de que as relações filogenéticas sugeridas sejam reais, considera-se que a análise combinada final resolveu de forma conclusiva as relações de parentesco entre as espécies do cluster *D. buzzatii*, uma vez que a hipótese filogenética mais provável para o cluster, diante de estudos morfológicos (Silva & Sene, 1991; Tidon-Sklorz & Sene, 1995a; Tidon-Sklorz & Sene, 2001, Moraes *et al.* 2004, Franco *et al.* 2006a) citogenéticos (Ruiz & Wasserman, 1993) e moleculares (Manfrin *et al.* 2001; Manfrin & Sene, 2006; Franco *et al.* 2010) prévios, foi corroborada com alto suporte probabilístico:

Drosophila buzzatii ocupou a posição mais basal dentro do cluster *D. buzzatii*, estando proximamente relacionada à espécie *D. koepferae*. *D. antonietae* foi sugerida como intermediária às espécies *D. koepferae* e *D. serido*, sendo a última alocada como táxon irmão do ramo formado pelas espécies *D. gouveai*, *D. borborema* e *D. seriema*, com *D. gouveai* ocupando a posição mais basal e *D. borborema* e *D. seriema* sugeridas como espécies irmãs. O ramo formado pelos indivíduos da espécie *D. serido* indicou uma relação de ancestralidade dos representantes das localidades de interior, ao norte da distribuição da espécie (N45 – Mucugê-BA e 1431 – Milagres-BA), em relação aos representantes das localidades litorâneas ao sul (N20 – Arraial do Cabo-RJ e N49 – Bertiooga-SP), em concordância com a hipótese do nordeste do Brasil, mais precisamente o interior do estado da Bahia, como centro de dispersão da espécie *D. serido* (de Brito *et al.* 2002a), e a divisão da espécie entre as populações do nordeste, que apresentam as inversões polimórficas 2a⁸, 2b⁸, 2c⁸ e 2d⁸, e as populações litorâneas, que apresentam as inversões 2y⁹, 2x⁸ e 2w⁸ (Tosi & Sene, 1989; Ruiz *et al.* 2000).

5.2. Os genes *transformer* e *period*

Neste trabalho, foi observado que os genes *transformer* e *period* são marcadores informativos no estudo interespecífico com espécies do complexo *D. buzzatii*.

Estimativas de diversidade nucleotídica com o gene *period* entre regiões conservadas e não conservadas sugeriram que tais regiões estão evoluindo em diferentes taxas para as espécies do cluster *D. buzzatii* (Franco, 2010). No presente trabalho, por

meio dos testes de razão de verossimilhança (LRT), taxas estatisticamente heterogêneas foram também verificadas ao longo dos ramos das árvores geradas com o gene *period*.

Apesar da variabilidade significativa, o alinhamento das regiões conservadas C4 e C5 do gene *period* de *D. buzzatii* e *D. mojavensis*, pertencentes ao subgênero *Drosophila*, com as mesmas regiões do *period* descritas para *D. melanogaster*, pertencente ao subgênero *Sophophora* (Franco *et al.* 2010), revelou um alto grau de conservação nessas regiões, a despeito do longo tempo de divergência, cerca de 40 milhões de anos atrás entre estes subgêneros (Markow & O'Grady, 2007). Esse alto grau de conservação entre espécies de *Drosophila*, e mesmo entre outros insetos, tais como *Ceratitis capitata* (Mazzotta *et al.* 2005), indica um papel funcional para as regiões C4 e C5. De fato, essas regiões estão incluídas no domínio CCID do gene *period*, que possui um importante papel na regulação do relógio circadiano: a atividade transcricional inibitória do heterodímero CLOCK/CYCLE de *Drosophila* (Chang & Reppert, 2003), que se liga aos promotores dos genes *period* e *timeless*, estimulando a transcrição desses genes (Tauber & Kyriacou, 2008).

Observou-se a partir dos testes de neutralidade aqui realizados que o fragmento do gene *period* analisado está sob ação de seleção purificadora (Tabelas 6 e 7), em concordância com esse importante papel na maquinaria do relógio circadiano (Franco *et al.* 2010).

No entanto, a indicação de seleção purificadora como principal força evolutiva e o indicativo de hibridação introgressiva envolvendo *D. buzzatii* e *D. koepferae* (Franco *et al.* 2010; presente trabalho), refutam a idéia de que o gene *period*, ou no mínimo o fragmento analisado, seja considerado um gene da especiação nas espécies do cluster *D. buzzatii*, pois tais genes geralmente evoluem sob ação de seleção direcional (Ting *et al.* 2000), e devem colaborar para aquisição de isolamento reprodutivo e especiação mesmo quando a divergência ocorre com fluxo gênico (Wu, 2001; Hey, 2006). Vários genes candidatos para o isolamento reprodutivo, e que satisfazem a premissa de uma taxa acelerada de substituições não sinônimas, têm sido descritos em espécies do gênero *Drosophila*, como o gene *Odysseus*, que contém um domínio *Homeobox* e que evoluiu sob forte influência de seleção direcional desde a divergência das espécies irmãs *D. simulans* e *D. mauritiana* (Ting *et al.* 1998; Ting *et al.* 2000; Sun *et al.* 2004). Outros genes relacionados ao isolamento reprodutivo pós-zigótico, tais como os genes *hybrid male rescue* e *nucleoporin 96–98*, também apresentam sinais de evolução Darwiniana

em *D. simulans* e *D. melanogaster* (Presgraves *et al.* 2003; Barbash *et al.* 2004; Orr *et al.* 2007).

Os testes de neutralidade com o gene *transformer* também evidenciaram seleção purificadora atuando como principal força evolutiva sobre os fragmentos analisados (Tabelas 4 e 5). Entretanto, comparativamente ao gene *period*, o gene *transformer* apresentou um número considerável de aminoácidos em que foi detectada razão entre d_n e d_s (ω) maior do que um. Apesar de não estar completamente de acordo com a hipótese de seleção direcional (Stern *et al.* 2007), devido aos intervalos de confiança não compreenderem intervalos inteiramente maiores que um (Tabela 5), o indicativo de uma possível seleção direcional atuando sobre as sequências deste gene, no complexo *D. buzzatii*, é bem mais relevante que no caso do gene anterior. Hipótese bastante plausível diante das altas taxas evolutivas verificadas para o gene *transformer* em *Drosophila* (Kulathinal *et al.* 2003), como pode ser observado a partir das sequências alinhadas no Apêndice 1D.

Os resultados obtidos para o gene *transformer*, portanto, contrastam com estudos dos padrões de polimorfismos em populações de *Drosophila americana* (McAllister & McVean, 2000) e com espécies do complexo *D. melanogaster* (Kulathinal *et al.* 2003), os quais sugerem deriva genética como responsável pela evolução rápida do gene *tra* e, assim, evolução neutra para o locus.

5.3. Diversificação e distribuição geográfica do cluster *Drosophila buzzatii*

O planeta Terra passou por vários períodos glaciais durante sua história. Na era Paleozóica, por exemplo, existem registros de dois expressivos eventos de glaciação, no início do Cambriano (570 milhões de anos atrás) e no final do Permiano (250 milhões de anos atrás) (Neto e Nery, 2005). No entanto, no início do período Quaternário (Gelasiano: 2,5 a 1,8 milhões de anos atrás) houve um crescimento expressivo na frequência dos períodos glaciais, sendo registradas pelo menos 16 glaciações na época do Pleistoceno (2,5 a 0,0117 milhões de anos atrás), com duração média de 100 mil anos, intercaladas por períodos de climas mais quentes (ou interglaciais) (Salgado-Labouriau, 1994; Suguio *et al.* 2005). Esses ciclos glaciais foram responsáveis por mudanças climáticas que levaram a alteração nos biomas durante cada ciclo (Hewitt, 2000; Colbeck *et al.* 2008) e, conseqüentemente, podem ter tido impacto profundo no padrão da distribuição geográfica de espécies atuais.

Análises fitossociológicas indicam que a distribuição geográfica das Florestas Tropicais Sazonalmente Secas (FTSS), conjunto de formações vegetais representado pelo Domínio da Caatinga, áreas na bacia do Paraná-Paraguai, vales Andinos e Caribe, foi mais ampla durante os períodos frios e secos do Pleistoceno, formando o chamado arco Pleistocênico (Prado & Gibbs, 1993; Pennington *et al.* 2000). Alteração neste conjunto de formações e nos demais domínios na América do Sul, em decorrência de mudanças climáticas globais, deve ter sido acompanhada de mudanças demográficas em sua fauna associada.

Portanto, as flutuações climáticas pleistocênicas podem ter contribuído para a origem de uma parcela da diversidade neotropical (Rull, 2008), influenciando na diferenciação populacional de espécies associadas à vegetação xerófita nos períodos interglaciais (Wuster *et al.* 2005, Almeida *et al.* 2007), e à vegetação mesófila nos períodos glaciais (Carnaval & Bates, 2007).

No hemisfério norte, durante os períodos glaciais, as geleiras avançaram sobre os continentes, extinguindo grande parte das populações de espécies não adaptadas a condições glaciais, e criando refúgios com condições adequadas à manutenção das populações sobreviventes. Por esse motivo, um padrão comumente encontrado em estudos filogeográficos neste hemisfério é o isolamento geográfico durante as fases glaciais, com expansão populacional nos períodos interglaciais subsequentes (Hewitt, 2000; Alexandrino *et al.* 2002; Kotlik *et al.* 2004). No hemisfério sul, não houve aumento significativo das geleiras sobre os continentes durante os períodos glaciais, com exceção de algumas partes da Patagônia (Zeisset & Beebe, 2008), mas as consideráveis alterações de temperatura e umidade podem ter contribuído para a alteração de biomas neotropicais durante cada ciclo (Ab'Saber, 1977; Hewitt, 2000).

Essas alterações foram consideradas importantes para explicar a origem da biodiversidade na região Neotropical (Haffer, 1969; Vanzoline e Williams, 1970; Prance, 1973), embora o efeito causal das flutuações climáticas do Pleistoceno sobre as taxas de especiação nessa região não seja consenso (Burnham & Graham, 1999; Behling, 2002; Pennington *et al.* 2004; Ledru *et al.* 2005; Bush & de Oliveira, 2006). Estimativas recentes de tempo de divergência, a partir de análises de relógio molecular, indicam que um número pequeno de táxons se diversificou neste período (Costa, 2003; Moritz *et al.* 2000; Pennington *et al.* 2000, Pennington *et al.* 2005; Almeida *et al.* 2007).

As estimativas obtidas, considerando a hipótese de relógio molecular com o gene COI, indicam que uma parcela da diversificação do cluster *D. buzzatii* está,

possivelmente, relacionada ao cenário exposto acima, ou seja, ao avanço e retração da vegetação xerófita na América do Sul, nas flutuações climáticas pleistocênicas, assim como sugerido por Franco (2009). Entretanto, as divisões iniciais e o processo de divergência a partir da espécie ancestral, que deu origem às duas espécies mais basais do cluster *D. buzzatii*, *D. buzzatii* e *D. koepferae*, estão, provavelmente, associados a eventos do período anterior, o Neógeno, mais precisamente o Plioceno (5,33 a 2,58 milhões de anos atrás).

O tempo para o ancestral comum mais recente das espécies do cluster *D. buzzatii* é de 3,98 milhões de anos, com um desvio de $\pm 1,12$ milhão de anos, e o início da divergência entre as espécies *D. buzzatii* e *D. koepferae* é estimado em cerca de 2,61 milhões de anos, com um desvio de $\pm 0,87$ milhão de anos. Na hipótese filogenética do gene COI para o cluster *D. buzzatii* (Figura 31), *D. buzzatii* e *D. koepferae* são sugeridas como monofiléticas, diferentemente do proposto a partir da hipótese filogenética final do presente trabalho (Figura 30). Desta forma, conciliando as estimativas de tempo à hipótese filogenética final, deve-se considerar que os dois eventos independentes que deram origem à *D. buzzatii* e *D. koepferae*, a partir da linhagem ancestral, ocorreram no período de 3,98 a 2,61 milhões de anos atrás, aproximadamente.

O início da principal fase de elevação dos Andes ocorreu no final do Mioceno e início do Plioceno (Gregory-Wodzicki, 2000), e evidências fósseis indicam a presença de savanas nas regiões proto-andinas de Argentina e Bolívia (Webb, 1978). Considerado a hipótese de que *D. buzzatii* tenha se originado no interior do Chaco argentino (Carson & Wasserman, 1965; Carson, 1965; Vilela *et al.* 1980; Fontdevila, 1989), além da distribuição atual de *D. koepferae*, é presumível que a linhagem ancestral do cluster *D. buzzatii* tenha se expandido para o sul da América do Sul através de áreas proto-andinas de FTSS, alcançando, assim, as regiões hoje ocupadas pelo Domínio do Chaco. Com o início da fase tectônica mais abrupta de soerguimento dos Andes no Plioceno, eventos geológicos condicionaram mudanças na distribuição das FTSS a leste América do Sul e nas áreas de savanas de Argentina e Bolívia. Tais eventos de remodelagem do relevo na região, associados a alterações na temperatura, umidade e, por consequência, na própria distribuição dos Domínios ali presentes, podem estar relacionados aos eventos de especiação de *D. buzzatii* e *D. koepferae*, no início da diversificação do cluster *D. buzzatii*.

Em outros estudos (Zanella, 1899; Silva, 1995; Colli, 2005; Almeida *et al.* 2007; Aguiar & Melo, 2008) também são relatados achados que permitem hipotetizar um

quadro histórico associado a eventos de vicariância anteriores aos eventos climáticos do Quaternário. Silva (1995) reconhece a subsidência do Chaco e áreas próximas e o soerguimento do Planalto Central brasileiro como o evento determinante para a vicariância de grupos de aves que hoje se encontram com disposição disjunta na América do Sul, com espécies no Cerrado e/ou Cadeia do Espinhaço, e em áreas de vegetação aberta mais ao sul, como Chaco, Pampa, Patagônia e setores subandinos (Carvalho & Almeida, 2010).

Atualmente a espécie *D. buzzatii* apresenta uma ampla distribuição na América do Sul, sendo encontrada em regiões áridas ou semi-áridas de Argentina, Bolívia, Paraguai e Brasil (Figueiredo & Sene, 1992). O processo de expansão da espécie ao longo da diagonal de vegetação aberta do continente, chegando até o Domínio da Caatinga, no nordeste brasileiro, é um tema controverso.

De acordo com Ab'Saber (1977), os Domínios da Caatinga e do Chaco foram conectados durante os períodos glaciais, pela expansão desses dois domínios em detrimento de áreas de Cerrado do Brasil central. No entanto, análises fitossociológicas abrangendo áreas de vegetação aberta, da região Neotropical, indicam que a Caatinga e o Chaco não são historicamente conectados, uma vez que o Chaco possui composição florística associada às vegetações temperadas secas do extremo sul dos Andes, enquanto a Caatinga possui conexão florística com áreas da região da bacia do Paraná-Paraguai, da costa Caribenha e com pequenos isolados nos vales secos dos Andes na Bolívia, Peru e Equador (Prado & Gibbs, 1993; Pennington *et al.* 2000).

A hipótese do Chaco argentino como centro de origem de *D. buzzatii* na América do Sul foi proposta com base em estudos populacionais de densidade demográfica e análises dos padrões de polimorfismo das inversões cromossômicas (Wasserman, 1962; Carson & Wasserman, 1965; Vilela *et al.*, 1980; Fontdevila, 1989; Figueiredo & Sene, 1992; Tidon-Sklorz *et al.* 1994), nos quais foi verificado um declínio na densidade das populações naturais e nos níveis de polimorfismo, conforme se avança do Chaco, para o sul do Brasil, região central e nordeste.

Entretanto, estudos biogeográficos recentes com *D. buzzatii* (de Brito *et al.* 2002b; Santos, 2011) mostraram que, para o gene COI, o maior índice de diversidade nucleotídica é encontrado entre as populações da Caatinga, quase duas vezes superior ao índice observado entre as populações do sul do Brasil, a segunda região mais polimórfica. As populações do Chaco, por sua vez, apresentaram os menores níveis de polimorfismo, representando cerca de 15% da variabilidade observada entre as

populações brasileiras. Estes resultados não são condizentes com um processo de expansão de *D. buzzatii* a partir do Chaco, pois em um evento de expansão, em geral, apenas alguns dos haplótipos presentes na distribuição original são encontrados na área mais recentemente ocupada (Cann *et al.* 1987; Templeton, 1998). Os dados do gene COI, portanto, sugerem uma presença mais ancestral da espécie *D. buzzatti* no nordeste brasileiro, hipótese reforçada pela existência de alelos aloenzimáticos exclusivos em suas populações (Barker *et al.* 1985).

Análises biogeográficas com novos marcadores moleculares poderão ajudar a esclarecer este contraste entre os dados de inversões cromossômicas e do gene COI, e caso a hipótese sugerida pelos dados do segundo marcador seja confirmada, um cenário provável seria a expansão da linhagem ancestral do cluster *D. buzzatii* não só rumo ao sul da América do Sul através das áreas de FTSS, nas regiões proto-andinas, mas também pelo norte do Brasil, antes da formação da bacia Amazônica no Plioceno (Hooghiemstra & van der Hammen, 1998), que promoveu o aumento das chuvas e da extensão das florestas tropicais, causando uma disjunção das áreas secas.

Apesar de *D. buzzatii* e *D. koepferae* terem surgido ancestralmente ao longo do processo de diversificação do cluster *D. buzzatii*, a diferença no padrão de distribuição entre as duas espécies é contrastante (Figura 2). Este fato pode estar relacionado às características ecológicas inerentes de cada espécie, as quais possibilitaram ou não eventos de expansão e colonização de novas áreas no passado, em períodos propícios de clima frio e seco (glaciais), resultando em uma ampla distribuição na América do sul de *D. buzzatii* - que recentemente colonizou Europa, África e Austrália, por ação antrópica (Fontdevila *et al.* 1981; Barker *et al.* 1985) - e uma distribuição restrita de *D. koepferae*.

Ambas as espécies ovopositam e se desenvolvem em tecidos em decomposição de várias espécies do gênero *Opuntia* e de cactos colunares dos gêneros *Cereus* e *Trichocereus* (Fontdevila *et al.* 1988; Hasson *et al.* 1992; Fanara *et al.* 1999). Embora não sejam diferentemente atraídas por *Opuntia* (*O. sulphurea*) ou cactos hospedeiros colunares (*T. terscheckii*) (Fanara *et al.* 1999), a composição da comunidade de moscas que emerge de substratos naturais em ambos os tipos de cactos difere significativamente (Fanara *et al.* 1999). Na natureza ou em laboratório, *D. buzzatii* ovoposita preferencialmente em cactos do gênero *Opuntia*, enquanto *D. koepferae* tem como sítio de ovoposição preferencial a espécie *T. terscheckii* (Fanara & Hasson 2001).

Como *D. buzzatii* e *D. koepferae* apresentaram a mesma viabilidade, tempo de desenvolvimento e tamanho de corpo em cultura média preparada com material

hospedeiro de *O. sulphurea* ou *T. terschekii* (Fanara *et al.* 1999), há duas alternativas para explicar a diferença entre a capacidade e os padrões encontrados na natureza: preferência por sítios de ovoposição distintos e/ou competição interespecífica (Fanara *et al.* 1999, Fanara & Hasson, 2001).

Estudos com populações de *D. buzzatii* da Austrália indicaram a existência de variabilidade na preferência por sítios de ovoposição com composições da micro-flora (leveduras e bactérias) diferentes, e que tal característica possui traços hereditários (Barker 1992; Barker *et al.* 1994; Barker & Starmer 1999). Fanara & Hasson (2001) confirmaram a existência de linhagens com padrões de ovoposição distintos para diferentes cactos hospedeiros (*O. sulphurea*, *O. quimilo* e *T. terschekii*) em *D. buzzatii*, o que não foi verificado para *D. koepferae*. Esta diferença de variabilidade genética, relacionada à diversificação de comportamentos ecológicos, foi, provavelmente, o fator mais relevante para a capacidade de colonização e sucesso evolutivo contrastantes entre as duas espécies, e evoluiu, possivelmente, em resposta à distribuição, desenvolvimento e bioquímica de seus respectivos hospedeiros (Thompson & Pellmyr, 1991).

Passando para a parcela de diversificação do cluster *D. buzzatii* restrita ao território brasileiro, as estimativas feitas com o gene COI indicam cerca de 1,27 milhão de anos, com desvio de $\pm 0,55$ milhão de anos, como tempo de divergência entre *D. antonietae* e a linhagem ancestral das espécies do clado AB, e 760 mil anos, com desvio de ± 340 mil anos, para o ancestral comum mais recente destas últimas espécies. Muito próximo do tempo estimado por Franco (2009) para o ancestral comum mais recente do clado AB (cerca de 840 mil anos), por meio de análises com o COI, baseadas na teoria da coalescência.

Portanto, os eventos de especiação que deram origem a *D. antonietae*, *D. serido*, *D. gouveai*, *D. borborema* e *D. seriema* estariam situados inteiramente no período Quaternário, principalmente no Pleistoceno.

A análise da diversidade haplotípica do DNA mitocondrial de *D. antonietae* indica uma expansão de área das populações do interior do estado do Rio Grande do Sul em direção a costa Atlântica (de Brito *et al.* 2002a) e colonização desta região para o norte (Morales *et al.* 2004), seguindo as populações de cactos *Cereus hildmaniannus* através das dunas do estado do Rio Grande do Sul e Santa Catarina (Morales, 2005). Com isso, é provável que a linhagem ancestral que deu origem à *D. koepferae* no Chaco, tenha se expandido para o interior do Brasil pelo sul, dando origem à espécie *D. antonietae*.

Não é possível afirmar, no entanto, que o isolamento da linhagem ancestral no Chaco, dando origem à *D. koepferae*, e na bacia do Paraná-Paraguai, dando origem a *D. antonietae*, está associado a eventos de retração da vegetação xerófita nos períodos interglaciais, uma vez que se observa uma repetição deste padrão biogeográfico para vários grupos de animais e plantas distribuídos ao longo da região temperada no sul da América do Sul (Morrone, 1993; Morrone *et al.* 1994; Amorin & Pires, 1996; Morrone, 2000), o que levou à proposição de um cladograma de área por Morrone (1993) e Morrone *et al.* (1994), que consideraram eventos de vicariância para explicar o padrão de distribuição e endemismo entre grupos irmãos presentes em regiões subandinas, no Chaco e bacia do Paraná-Paraguai.

As áreas de ocorrência atual das espécies do clado AB, com diversos pontos de sobreposição (Figura 2), tornam difícil o entendimento do contexto geográfico dos processos de especiação que deram origem a estas espécies, e podem ter ocorrido tanto em alopatria, com posterior contato secundário entre as linhagens, quanto na presença de áreas em simpatria, mediante diferenciação ecológica (Losos & Glor, 2003; Nosil, 2008).

Especiação alopátrica foi considerada o principal processo de diversificação e origem de diversidade no gênero *Drosophila* (Orr *et al.* 2007). No entanto, tem crescido a quantidade de dados sugerindo cenários em que a diversificação ocorreu na presença de fluxo gênico entre espécies desse gênero (Machado *et al.* 2007), bem como em outros grupos de insetos (Nosil, 2008).

No caso da linhagem que originou as espécies *D. borborema* e *D. seriema* é possível que tenha ocorrido a influência de componentes ecológicos, relacionados à adaptação a diferentes cactos hospedeiros e a ambiente de campos rupestres (Franco, 2009). Tal inferência está baseada nos seguintes fatos:

1- *D. gouveai*, espécie irmã da linhagem *D. borborema/D. seriema*, está associada ao cacto *Pilosocereus machrisis*, que ocorre em afloramentos rochosos areníticos do Brasil central, na maior parte de sua distribuição geográfica (Tidon-Sklorz & Sene, 2001; Moraes & Sene, 2007), sendo assim, talvez a adaptação a campos rupestres seja uma condição pleiomórfica do grupo.

2- A espécie *D. seriema* é restrita a ambientes de campos rupestres, acima de 1000m, da Serra do Espinhaço e Chapada da Diamantina.

3- *D. borborema*, embora não ocorra exclusivamente em campos rupestres, ocorre nesses ambientes na Chapada da Diamantina e na Serra de Grão Mogol (MG),

localizada ao norte da Serra do Espinhaço. Na Serra de Grão Mogol, inclusive, foi descrita a emergência de moscas pertencentes às espécies *D. borborema* e *D. seriema* a partir do mesmo cacto (Tidon-Sklorz, 1992), indicando que ambas espécies compartilham características metabólicas.

Nesse contexto, a diferenciação da linhagem *D. seriema/D. borborema* pode ter ocorrido através de adaptação aos cactos de campos rupestres, que promoveram um isolamento micro-geográfico relacionado a altitude, não sendo descartados contatos recorrentes com troca de material genético entre a linhagem *D. seriema/D. borborema*, com linhagens ancestrais do clado AB a oeste da Cadeia do Espinhaço, que se diferenciaram em *D. gouveai*, e a leste da Cadeia do Espinhaço, que originaram *D. serido* (Franco, 2009).

Considerando a distribuição geográfica atual da espécie *D. serido* (Figura 2), sua estruturação geográfica (Manfrin *et al.* 2001; de Brito *et al.* 2002a; Morales, 2005; Franco, 2010; presente trabalho) e seu caráter politípico (Baimai *et al.* 1983, Tosi & Sene, 1989), foram propostas duas explicações para sua ocorrência no nordeste e ao longo da costa atlântica. Em uma delas, as populações teriam expandido suas áreas de ocorrência do nordeste em direção ao sul pelo litoral e, devido a eventos vicariantes (expansão e retração da vegetação xerófita), teriam se tornado populações isoladas e diferenciadas (de Brito *et al.* 2002a; Morales, 2005). A outra possibilidade considera eventos independentes de colonização da costa a partir de populações do interior, especialmente da Serra do Espinhaço e Serra da Mantiqueira, através dos vales de rios que formam um corredor de dispersão, conectando as populações do interior do Brasil à costa atlântica (de Brito *et al.* 2002a; Morales, 2005). Além disso, foi proposta a ocorrência de migrações bidirecionais entre as populações do litoral sul da Bahia, Rio de Janeiro e São Paulo, que estariam estruturadas de acordo com o modelo de isolamento por distância (Morales, 2005; Franco, 2009). A distribuição destas populações politípicas de *D. serido* ao longo da costa coincide com a distribuição geográfica de outras espécies de *Drosophila*, tal como *D. meridionalis* (Baimai *et al.* 1983), e também coincide com áreas de endemismo de outros insetos e grupos de vertebrados (Amorin & Pires, 1996). Associações de grupos independentes a áreas zoogeográficas sugerem eventos históricos de isolamento geográfico determinando o padrão de diferenciação destas populações.

Ainda em relação à linhagem *D. borborema/D. seriema*, os dados apresentados por Franco (2009) indicam que *D. borborema* pode ter tido uma origem peripátrica a partir

de populações isoladas de *D. seriema*. A especiação peripátrica é um tipo particular de especiação alopátrica, que envolve o estabelecimento de uma nova população a partir de um pequeno número de indivíduos, ou seja, o efeito do fundador (Mayr, 1963; Coyne & Orr, 2004; Templeton, 2008b). Entre os fatos que apóiam tal hipótese pode-se citar: *D. seriema* apresenta um nível de variabilidade genética maior do que *D. borborema*, tanto para o gene mitocondrial COI, quanto para o gene nuclear *period*; o tamanho populacional histórico da espécie *D. borborema* é o menor dentre as espécies do clado AB; o tempo para o ancestral comum mais recente estimado (por análises de coalescência) para *D. borborema* é cerca de quatro vezes menor do que o estimado para *D. seriema*; os haplótipos (COI e *period*) de *D. borborema* são derivados de haplótipos de *D. seriema*.

Os pontos discutidos anteriormente, portanto, evidenciam o padrão complexo de divergência apresentado pelo cluster *D. buzzatii* (Sene *et al.* 1988). Além de fatores intrínsecos à reconstrução filogenética de espécies filogeneticamente próximas (Machado & Hey, 2003), eventos recorrentes de hibridação introgressiva entre as espécies do cluster (Manfrin *et al.* 2001; de Brito *et al.* 2002a; Gomes & Hasson, 2003; Piccinalli *et al.* 2004; Manfrin & Sene, 2006; Franco *et al.* 2010; presente trabalho) dificultam ainda mais o entendimento dos processos pelos quais o grupo passou ao longo de sua diversificação, que iniciou-se no final do Neógeno, no período Plioceno, provavelmente em decorrência de eventos de vicariância associados à elevação dos Andes, sendo também influenciada pelo avanço e retração da vegetação xerófita nas flutuações climáticas do Pleistoceno, no Quaternário.

6. CONCLUSÕES

- As espécies do complexo *D. buzzatii* formam um grupo monofilético.
- As espécies do cluster *D. buzzatii* formam um grupo monofilético.
- As espécies *D. martensis*, *D. starmeri* e *D. venezolana* constituem uma linhagem evolutiva coesa.
- As espécies do cluster *D. stalker*i formam um grupo parafilético.
- As relações de parentesco entre as espécies do cluster *D. buzzatii* são expressas pelo seguinte cladograma:

- Há uma relação de ancestralidade das populações do nordeste de *D. serido* em relação às populações litorâneas.
- O gene *transformer* está evoluindo sob ação de seleção purificadora para as espécies do complexo *D. buzzatii*, entretanto, há indícios estatisticamente não significativos de uma possível seleção direcional atuando como principal força evolutiva no locus.
- O fragmento do gene *period* analisado também está evoluindo sob ação de seleção purificadora para as espécies do complexo *D. buzzatii*, em concordância com seu importante papel na maquinaria do relógio circadiano.
- O cluster *D. buzzatii* apresenta um complexo padrão de divergência, e os processos pelos quais o grupo passou ao longo de sua diversificação estão relacionados tanto ao período do Plioceno, no Neógeno, quanto ao período do Pleistoceno, no Quaternário.

7. REFERÊNCIAS BIBLIOGRÁFICA

- Ab' Saber, A.N. 1977. Espaços ocupados pela expansão dos climas secos da América do Sul, por ocasião dos períodos glaciais quaternários. *Paleoclimas*, 3, 1-19.
- Aguiar, A. J. C. & Melo, G. A. R. 2008. Biogeografia das áreas abertas da América do Sul com ênfase em abelhas da tribo Tapinotaspidini. In: VIII Encontro Sobre Abelhas. Ribeirão Preto. *Anais do VIII Encontro Sobre Abelhas*, 322-326.
- Alexandrino, J., Arntzen, J. & Ferrand, N. 2002. Nested clade analysis and the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Heredity*, 88, 66-74.
- Almeida, F. C., Bonvicino, C. R. & Cordeiro-Estrela, P. 2007. Phylogeny and temporal diversification of *Calomys* (Rodentia, Sigmodontinae): implications for the biogeography of an endemic genus of the open/dry biomes of South America. *Molecular Phylogenetics and Evolution*, 42, 449-466.
- Amorim, D.S. 1997. Elementos básicos de sistemática filogenética. 2ª Edição. Holos Editora & Sociedade Brasileira de Entomologia, Ribeirão Preto, 276pp.
- Amorim, D. S. & Pires, M. 1996. Neotropical biogeography and a method for maximum biodiversity estimation. *Biodiversity in Brazil, a first approach*. São Paulo, CNPq, 326, 183-219.
- Arnold, M. L. 1992. Natural hybridization as an evolutionary process. *Annual Review of Ecology and Systematics*, 23, 237-261.
- Arnold, M. L. 1997. *Natural hybridization and evolution*: Oxford University Press, USA.
- Avice, J. 2000. *Phylogeography: the history and formation of species*: Harvard University Press. 447p.
- Avice, J. C. 2009. Phylogeography: retrospect and prospect. *Journal of Biogeography*, 36, 3-15.
- Avice, J. C. & Wollenberg, K. 1997. Phylogenetics and the origin of species. *Proceedings of the National Academy of Sciences (USA)*, 94, 7748-7755.
- Bächli G. 2010. Taxo-Dros: the database on taxonomy of *Drosophilidae*, version April 2010. <http://taxodros.unizh.ch>.
- Baimai, V., Sene, F. & Pereira, M. 1983. Heterochromatin and karyotypic differentiation of some neotropical cactus-breeding species of the *Drosophila repleta* species group. *Genetica*, 60, 81-92.

- Ballard, J. W. O. & Whitlock, M. C. 2004. The incomplete natural history of mitochondria. *Molecular Ecology*, 13, 729-744.
- Barbash, D. A., Awadalla, P. & Tarone, A. M. 2004. Functional divergence caused by ancient positive selection of a *Drosophila* hybrid incompatibility locus. *PLoS Biology*, 2, 839-848.
- Barker, F. K. & Lutzoni, F. M. 2002. The utility of the incongruence length difference test. *Systematic Biology*, 51, 625-637.
- Barker, J. 1992. Genetic variation in cactophilic *Drosophila* for oviposition on natural yeast substrates. *Evolution*, 46, 1070-1083.
- Barker, J., de M. Sene, F., East, P. & Pereira, M. 1985. Allozyme and chromosomal polymorphism of *Drosophila buzzatii* in Brazil and Argentina. *Genetica*, 67, 161-170.
- Barker, J. S. F., Starmer, W. T. & Fogleman, J. C. 1994. Genotype-specific habitat selection for oviposition sites in the cactophilic specie *Drosophila buzzatii*. *Heredity*, 72: 384-395.
- Barker, J. & Starmer, W. 1999. Environmental effects and the genetics of oviposition site preference for natural yeast substrates in *Drosophila buzzatii*. *Hereditas*, 130, 145-175.
- Barker, J. S. F., de M. Sene, F., East, P. & Pereira, M. A. Q. R. 1985. Allozyme and chromosomal polymorphism of *Drosophila buzzatii* in Brazil and Argentina. *Genetica*, 67, 161-170.
- Barr, N. B., Cui, L. & McPherson, B. A. 2005. Molecular systematics of nuclear gene period in genus *Anastrepha* (Tephritidae). *Annals of the Entomological Society of America*, 98, 173-180.
- Barrett, M., Donoghue, M. J. & Sober, E. 1991. Against consensus. *Systematic Zoology*, 40, 486-493.
- Bauzer, L., Souza, N., Ward, R., Kyriacou, C. & Peixoto, A. 2002. The period gene and genetic differentiation between three Brazilian populations of *Lutzomyia longipalpis*. *Insect Molecular Biology*, 11, 315-323.
- Beaver, L., Gvakharia, B., Vollintine, T., Hege, D., Stanewsky, R. & Giebultowicz, J. 2002. Loss of circadian clock function decreases reproductive fitness in males of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 2134.
- Beaver, L., Rush, B., Gvakharia, B. & Giebultowicz, J. 2003. Noncircadian regulation

- and function of clock genes period and timeless in oogenesis of *Drosophila melanogaster*. *Journal of biological rhythms*, 18, 463.
- Beaver, L. M. & Giebultowicz, J. M. 2004. Regulation of copulation duration by period and timeless in *Drosophila melanogaster*. *Current Biology*, 14, 1492-1497.
- Behling, H. 2002. South and southeast Brazilian grasslands during Late Quaternary times: a synthesis. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 177, 19-27.
- Bizzo, N. M. V. & Sene, F. M. 1982. Studies on the natural populations of *Drosophila* from Peruíbe (SP), Brazil (Diptera; Drosophilidae). *Revista Brasileira de Biologia*, 42, 539-544.
- Bremer, K. & Gustafsson, M. H. G. 1997. East Gondwana ancestry of the sunflower alliance of families. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 9188.
- Britton, T. 2005. Estimating divergence times in phylogenetic trees without a molecular clock. *Systematic Biology*, 54, 500.
- Britton, T., Anderson, C., Jaquet, D., Lundqvist, S. & Bremer, K. 2006. PATHd8 - a new method for estimating divergence times in large phylogenetic trees without a molecular clock. Available from the authors (www.math.su.se/PATHd8).
- Brower, A. V. Z. & DeSalle, R. 1994. Practical and theoretical considerations for choice of a DNA sequence region in insect molecular systematics, with a short review of published studies using nuclear gene regions. *Annals of the Entomological Society of America*, 87, 702-716.
- Bull, J., Huelsenbeck, J. P., Cunningham, C. W., Swofford, D. L. & Waddell, P. J. 1993. Partitioning and combining data in phylogenetic analysis. *Systematic Biology*, 42, 384.
- Burnham, R. J. & Graham, A. 1999. The history of neotropical vegetation: new developments and status. *Annals of the Missouri Botanical Garden*, 86, 546-589.
- Bush, M. B. & Oliveira, P. E. 2006. The rise and fall of the Refugial Hypothesis of Amazonian speciation: a paleoecological perspective. *Biota Neotropica*, 6, 0-0.
- Cahill, G. M. 2002. Clock mechanisms in zebrafish. *Cell and tissue research*, 309, 27-34.
- Cann, R. L., Stoneking, M. & Wilson, A. 1987. Mitochondrial DNA and human evolution. *Nature*, 325, 1200.
- Carnap, R. 1950. Logical foundations of probability. Univ. Chicago Press, Chicago.

- Carnaval, A. C. & Bates, J. M. 2007. Amphibian DNA shows marked genetic structure and tracks Pleistocene climate change in northeastern Brazil. *Evolution*, 61, 2942-2957.
- Carson, H. L. 1965. *Chromosomal morphism in Drosophila*. In *The Genetics of Colonizing Species* H.G. Barker and G.L. Stebbins (eds) Academic Press New York.
- Carson, H. L. & Wasserman, M. 1965. A widespread chromosomal polymorphism in a widespread species, *Drosophila buzzatii*. *The American Naturalist*, 99, 111-115.
- Carvalho, C. J. B. & Almeida, E. A. B. 2010. *Biogeografia da América do Sul: padrões e processos*. 1ª Edição. Editora Roca.
- Chang, D. C. & Reppert, S. M. 2003. A novel C-terminal domain of *Drosophila* PERIOD inhibits dCLOCK: CYCLE-mediated transcription. *Current Biology*, 13, 758-762.
- Cho, S., Mitchell, A., Regier, J., Mitter, C., Poole, R. W., Friedlander, T. P., et al. 1995. A highly conserved nuclear gene for low-level phylogenetics: elongation factor-1 alpha recovers morphology-based tree for heliothine moths. *Molecular Biology and Evolution*, 12, 650.
- Citri, Y., Colot, H. V., Jacquier, A. C., Yu, Q., Hall, J. C., Baltimore, D., et al. 1987. A family of unusually spliced biologically active transcripts encoded by a *Drosophila* clock gene. *Nature*, 326, 42-47.
- Clark, A. G. 1997. Neutral behavior of shared polymorphism. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 7730.
- Colbeck, G. J., Gibbs, H. L., Marra, P. P., Hobson, K. & Webster, M. S. 2008. Phylogeography of a widespread north american migratory songbird (*Setophaga ruticilla*). *Journal of Heredity*, 99, 453.
- Colli, G., Scariot, A., Souza-Silva, J. & Felfili, J. 2005. As origens e a diversificação da herpetofauna do Cerrado. *Cerrado: ecologia, biodiversidade e conservação*, 247-264.
- Colot, H. V., Hall, J. C. & Rosbash, M. 1988. Interspecific comparison of the period gene of *Drosophila* reveals large blocks of non-conserved coding DNA. *The EMBO Journal*, 7, 3929.
- Costa, L. P. 2003. The historical bridge between the Amazon and the Atlantic Forest of Brazil: a study of molecular phylogeography with small mammals. *Journal of Biogeography*, 30, 71-86.

- Costa, R., Peixoto, A. A., Barbuji, G. & Kyriacou, C. P. 1992. A latitudinal cline in a *Drosophila* clock gene. *Proceedings: Biological Sciences*, 250, 43-49.
- Coyne, J. A. 1992. Genetics and speciation. *Nature*, 355, 511-515.
- Coyne, J. A., Crittenden, A. P. & Mah, K. 1994. Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. *Science*, 265, 1461.
- Coyne, J. A. & Orr, H. A. 2004. *Speciation*: Sinauer Associates.
- Coyne, J. A. & Oyama, R. 1995. Localization of pheromonal sexual dimorphism in *Drosophila melanogaster* and its effect on sexual isolation. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 9505.
- Crews, S. T., Thomas, J. B. & Goodman, C. S. 1988. The *Drosophila* single-minded gene encodes a nuclear protein with sequence similarity to the per gene product. *Cell*, 52, 143-151.
- Cunningham, C. W. 1997. Can three incongruence tests predict when data should be combined? *Molecular Biology and Evolution*, 14, 733.
- Danforth, B. N. & Ji, S. 1998. Elongation factor-1 alpha occurs as two copies in bees: implications for phylogenetic analysis of EF-1 alpha sequences in insects. *Molecular Biology and Evolution*, 15, 225.
- Darlu, P. & Lecointre, G. 2002. When does the incongruence length difference test fail? *Molecular Biology and Evolution*, 19, 432.
- de Brito, R. A., Manfrin, M. H. & Sene, F. M. 2002a Nested cladistic analysis of Brazilian populations of *Drosophila serido*. *Molecular Phylogenetics and Evolution*, 22, 131-143.
- de Brito, R. A., Manfrin, M. H. & Sene, F. M. 2002b Mitochondrial DNA phylogeography of Brazilian populations of *Drosophila buzzatii*. *Genetics and Molecular Biology*, 25, 161-171.
- de Queiroz, A., Donoghue, M. J. & Kim, J. 1995. Separate versus combined analysis of phylogenetic evidence. *Annual Review of Ecology and Systematics*, 26, 657-681.
- de Queiroz, A. & Gatesy, J. 2007. The supermatrix approach to systematics. *Trends in Ecology & Evolution*, 22, 34-41.
- Driskell, A. C., Ané, C., Burleigh, J. G., McMahon, M. M., O'Meara, B. C. & Sanderson, M. J. 2004. Prospects for building the tree of life from large sequence databases. *Science*, 306, 1172.
- Duda Jr, T.F. & Lee, T. 2009. Isolation and population divergence of a widespread Indo-West Pacific marine gastropod at Easter Island. *Mar. Biol.* 156: 1193–1202.

- Durando, C. M., Baker, R. H., Etges, W. J., Heed, W. B., Wasserman, M. & DeSalle, R. 2000. Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. *Molecular Phylogenetics and Evolution*, 16, 296-307.
- Etges, W. J. 1993. Genetics of host-cactus response and life-history evolution among ancestral and derived populations of cactophilic *Drosophila mojavensis*. *Evolution*, 47, 750-767.
- Etges, W. J. 2002. Divergence in mate choice systems: does evolution play by rules? *Genetica*, 116, 151-166.
- Etges, W. J. & Ahrens, M. A. 2001. Premating isolation is determined by larval-rearing substrates in cactophilic *Drosophila mojavensis*. V. Deep geographic variation in epicuticular hydrocarbons among isolated populations. *American Naturalist*, 158, 585-598.
- Ewing, A. W. 1983. Functional aspects of *Drosophila* courtship. *Biological Reviews*, 58, 275-292.
- Fanara, J., Fontdevila, A. & Hasson, E. 1999. Oviposition preference, viability, developmental time and body size in the cactophilic sibling species *Drosophila koepferae* and *D. buzzatii* in association to their natural hosts. *Evol Ecol*, 13, 173-190.
- Fanara, J. J. & Hasson, E. 2001. Oviposition acceptance and fecundity schedule in the cactophilic sibling species *Drosophila buzzatii* and *D. koepferae* on their natural hosts. *Evolution*, 55, 2615-2619.
- Farris, J. S., Källersjö, M., Kluge, A. G. & Bult, C. 1995. Constructing a significance test for incongruence. *Systematic Biology*, 44, 570.
- Farris, J. S., Källersjö, M., Kluge, A. G. & Bult, C. 1994. Testing significance of incongruence. *Cladistics*, 10, 315-319.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39, 783-791.
- Ferveur, J. F., Cobb, M., Boukella, H. & Jallon, J. M. 1996. World-wide variation in *Drosophila melanogaster* sex pheromone: behavioural effects, genetic bases and potential evolutionary consequences. *Genetica*, 97, 73-80.
- Figueiredo, V. L. C. & Sene, F. M. 1992. Chromosome variability in Brazilian populations of *Drosophila buzzatii* (Diptera, Drosophilidae). *Rev. bras. biol*, 52, 555-561.

- Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology*, 20, 406-416.
- Fontdevila, A. 1989. Founder effects in colonizing populations: the case of *Drosophila buzzatii*. *Evolutionary biology of transient unstable populations*, 74-95.
- Fontdevila, A., Pla, C., Hasson, E., Wasserman, M., Sanchez, A., Navieira, H., et al. 1988. *Drosophila koepferae*: a new member of the *Drosophila serido* (Diptera: Drosophilidae) superspecies taxon. *Annals of the Entomological Society of America*, 81, 380-385.
- Fontdevila, A., Ruiz, A., Alonso, G. & Ocaña, J. 1981. Evolutionary history of *Drosophila buzzatii*. I. Natural chromosomal polymorphism in colonized populations of the Old World. *Evolution*, 148-157.
- Ford, M. J., Yoon, C. K. & Aquadro, C. F. 1994. Molecular evolution of the period gene in *Drosophila athabasca*. *Molecular Biology and Evolution*, 11, 169.
- Franco, F.F. 2009. História evolutiva do “cluster” *Drosophila buzzatii* (grupo *D. repleta*): eventos históricos e diversificação de espécies no Brasil. Tese de Doutorado. Faculdade de Medicina de Ribeirão Preto, USP.
- Franco, F. F., Silva-Bernardi, E., Sene, F., Hasson, E. & Manfrin, M. 2010. Intra- and interspecific divergence in the nuclear sequences of the clock gene period in species of the *Drosophila buzzatii* cluster. *Journal of Zoological Systematics and Evolutionary Research*, 48, 322-331.
- Franco, F. F., Kuhn, G. C., Sene, F. M. & Manfrin, M. H. 2006b Conservation of pBuM-2 satellite DNA sequences among geographically isolated *Drosophila gouveai* populations from Brazil. *Genetica*, 128, 287-295.
- Franco, F. F., Prado, P. R. R., Sene, F. M., Costa, L. F. & Manfrin, M. H. 2006a Aedeagus morphology as a discriminant marker in two closely related cactophilic species of *Drosophila* (Diptera; Drosophilidae) in South America. *Anais da Academia Brasileira de Ciências*, 78, 203-212.
- Franco, F. F., Sene, F. M. & Manfrin, M. H. 2008b. Molecular characterization of SSS139, a new satellite DNA family in sibling species of the *Drosophila buzzatii* cluster. *Genetics and Molecular Biology*, 31, 155-159.
- Franco, F. F., Soto, I. M., Sene, F. M. & Manfrin, M. H. 2008a. Phenotypic variation of male external genitalia of *Drosophila serido* Vilela & Sene (Diptera, Drosophilidae). *Neotropical Entomology*, 37.
- Funk, D. J. & Omland, K. E. 2003. Species-level parphyly and polyphyly: frequency,

- causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, 397-423.
- Gatesy, J., Baker, R. H. & Hayashi, C. 2004. Inconsistencies in arguments for the supertree approach: supermatrices versus supertrees of *Crocodylia*. *Systematic Biology*, 53, 342-355.
- Gilbert, S.F. 2003. *Biologia do desenvolvimento*. 5ª Edição. FUNPEC Editora.
- Gleason, J. M. & Powell, J. R. 1997. Interspecific and intraspecific comparisons of the period locus in the *Drosophila willistoni* sibling species. *Molecular Biology and Evolution*, 14, 741.
- Gómez, G. A. & Hasson, E. 2003. Transpecific polymorphisms in an inversion linked esterase locus in *Drosophila buzzatii*. *Molecular Biology and Evolution*, 20, 410.
- Gregory-Wodzicki, K. M. 2000. Uplift history of the Central and Northern Andes: a review. *Geological Society of America Bulletin*, 112, 1091.
- Haffer, J. 1969. Speciation in Amazonian forest birds. *Science*, 165, 131-137.
- Hasegawa, M., Kishino, H. & Yano, T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of molecular evolution*, 22, 160-174.
- Hasson, E., Naveira, H. & Fontdevila, A. 1992. The breeding sites of Argentinian cactophilic species of the *Drosophila mulleri* complex (subgenus *Drosophila -repleta* group). *Revista Chilena de Historia Natural*, 65, 319-326.
- Hasson, E., Rodriguez, C., Fanara, J. J., Naveira, H., Reig, O. A. & Fontdevila, A. 1995. The evolutionary history of *Drosophila buzzatti*. XXVI. Macrogeographic patterns of inversion polymorphism in New World populations. *Journal of Evolutionary Biology*, 8, 369-384.
- Hebert, P. D. N., Ratnasingham, S. & de Waard, J. R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270, S96.
- Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907-913.
- Hey, J. 2006. Recent advances in assessing gene flow between diverging populations and species. *Current opinion in genetics & development*, 16, 592-596.
- Higgie, M. & Blows, M. W. 2007. Are traits that experience reinforcement also under sexual selection? *American Naturalist*, 170, 409-420.
- Hillis, D. & Huelsenbeck, J. 1992. Signal, noise, and reliability in molecular

- phylogenetic analyses. *Journal of Heredity*, 83, 189.
- Hilton, H. & Hey, J. 1996. DNA sequence variation at the period locus reveals the history of species and speciation events in the *Drosophila virilis* group. *Genetics*, 144, 1015.
- Hooghiemstra, H. & van der Hammen, T. 1998. Neogene and Quaternary development of the neotropical rain forest: the forest refugia hypothesis, and a literature overview. *Earth-Science Reviews*, 44, 147-183.
- Hovemann, B., Richter, S., Walldorf, U. & Cziepluch, C. 1988. Two genes encode related cytoplasmic elongation factors 1 {alpha}(EF-1 {alpha}) in *Drosophila melanogaster* with continuous and stage specific expression. *Nucleic acids research*, 16, 3175.
- Huelsenbeck, J. P., Bull, J. & Cunningham, C. W. 1996. Combining data in phylogenetic analysis. *Trends in Ecology & Evolution*, 11, 152-158.
- Huelsenbeck, J. P. & Ronquist, F. 2001 MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17, 754-755.
- Hurtado, L., Erez, T., Castrezana, S. & Markow, T. 2004. Contrasting population genetic patterns and evolutionary histories among sympatric Sonoran Desert cactophilic *Drosophila*. *Molecular Ecology*, 13, 1365-1375.
- Jallon, J. M. & David, J. R. 1987. Variation in cuticular hydrocarbons among the eight species of the *Drosophila melanogaster* subgroup. *Evolution*, 41, 294-302.
- Kimura, M. 1983. The neutral theory of Molecular Evolution. Cambridge University Press, Cambridge, Massachusetts.
- Kliman, R. M. & Hey, J. 1993. DNA sequence variation at the period locus within and among species of the *Drosophila melanogaster* complex. *Genetics*, 133, 375.
- Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (*Boidae*, *Serpentes*). *Systematic Zoology*, 38, 7-25.
- Knowles, L. 2008. Why does a method that fails continue to be used? *Evolution*, 62, 2713-2717.
- Kolaczowski, B. & Thornton, J. W. 2009. Long-branch attraction bias and inconsistency in Bayesian phylogenetics. *PLoS One*, 4, e7891.
- Kokudai, C.S., Manfrin, M.H., Sene, F.M. 2011. No prelo. Sympatry and asymmetric introgression between the cactophilic species *Drosophila serido* and *Drosophila antonietae* (Diptera/Drosophilidae). *Annals of the Entomological Society of America*.

- Konopka, R. J. & Benzer, S. 1971. Clock mutants of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, 68, 2112.
- Konopka, R. J., Kyriacou, C. P. & Hall, J. C. 1996. Mosaic Analysis in the *Drosophila* Cns of circadian and courtship-Song rhythms affected by a period clock mutation: Short Communication. *Journal of Neurogenetics*, 11, 117-139.
- Kotlik, P., Bogutskaya, N. & Ekmekci, F. 2004. Circum Black Sea phylogeography of *Barbus* freshwater fishes: divergence in the Pontic glacial refugium. *Molecular Ecology*, 13, 87-95.
- Kuhn, G. & Sene, F. M. 2005. Evolutionary turnover of two pBuM satellite DNA subfamilies in the *Drosophila buzzatii* species cluster (*repleta* group): from alpha to alpha/beta arrays. *Gene*, 349, 77-85.
- Kulathinal, R. J., Skwarek, L., Morton, R. A. & Singh, R. S. 2003. Rapid evolution of the sex-determining gene, transformer: structural diversity and rate heterogeneity among sibling species of *Drosophila*. *Molecular Biology and Evolution*, 20, 441.
- Kyriacou, C. & Hall, J. C. 1980. Circadian rhythm mutations in *Drosophila melanogaster* affect short-term fluctuations in the male's courtship song. *Proceedings of the National Academy of Sciences of the United States of America*, 77, 6729.
- Kyriacou, C., Peixoto, A. & Costa, R. 2007. A cline in the *Drosophila melanogaster* period gene in Australia: neither down nor under. *Journal of Evolutionary Biology*, 20, 1649-1651.
- Kyriacou, C., Sawyer, L.A., Piccin, A., Couchman, M.E. & Chalmers, D. 1996. Evolution and population biology of the period gene. *Semin. Cell Dev. Biol.* 7: 803-810.
- Ledru, M. P., Rousseau, D. D. & Cruz, F. 2005. Paleoclimate changes during the last 100,000 yr from a record in the Brazilian Atlantic rainforest region and interhemispheric comparison. *Quaternary Research*, 64, 444-450.
- Leigh, J. W., Susko, E., Baumgartner, M. & Roger, A. J. 2008. Testing congruence in phylogenomic analysis. *Systematic Biology*, 57, 104.
- Li, W.H. 1997. Molecular Evolution. Sinauer Associates, Inc. Sunderland, Mass.
- Losos, J. B. & Glor, R. E. 2003. Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology & Evolution*, 18, 220-227.
- Machado, C. A. & Hey, J. 2003. The causes of phylogenetic conflict in a classic

- Drosophila* species group. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270, 1193.
- Machado, C. A., Matzkin, L. M., Reed, L. K. & Markow, T. A. 2007. Multilocus nuclear sequences reveal intra- and interspecific relationships among chromosomally polymorphic species of cactophilic *Drosophila*. *Molecular Ecology*, 16, 3009-3024.
- Machado, L. P. B., Castro, J. P. & Madi-Ravazzi, L. 2002. Evaluation of the courtship and of the hybrid male sterility among *Drosophila buzzatii* cluster species (Diptera, Drosophilidae). *Brazilian Journal of Biology*, 62, 601-608.
- Machado, L. P. B., Madi-Ravazzi, L. & Tadei, W. J. 2006 Reproductive relationships and degree of synapsis in the polytene chromosomes of the *Drosophila buzzatii* species cluster. *Brazilian Journal of Biology*, 66, 279-293.
- Madi-Ravazzi, L., Bicudo, H. E. M. C. & Manzato, J. A. 1997. Reproductive compatibility and chromosome pairing in the *Drosophila buzzatii* complex. *Cytobios*, 89, 21-30.
- Manfrin, M. H., De Brito, R. A. & Sene, F. M. 2001. Systematics and evolution of the *Drosophila buzzatii* (Diptera: Drosophilidae) cluster using mtDNA. *Annals of the Entomological Society of America*, 94, 333-346.
- Manfrin, M. H. & Sene, F. M. 2006. Cactophilic *Drosophila* in South America: A model for evolutionary studies. *Genetica*, 126, 57-75.
- Marin, I., Ruiz, A., Pla, C. & Fontdevila, A. 1993. Reproductive relationships among ten species of the *Drosophila repleta* group from South America and the West Indies. *Evolution*, 47, 1616-1624.
- Markow, T. A. & O'Grady, P. M. 2007. *Drosophila* biology in the genomic age. *Genetics*, 177, 1269.
- Matzkin, L. M. 2008. The molecular basis of host adaptation in cactophilic *Drosophila*: molecular evolution of a glutathione S-transferase gene (GstD1) in *Drosophila mojavensis*. *Genetics*, 178, 1073.
- Mayr, E. 1963. Animal species and evolution. Cambridge, MA: Harvard University Press.
- Mazzoni, C. J., Souza, N. A., Andrade-Coelho, C., Kyriacou, C. P. & Peixoto, A. A. 2006. Molecular polymorphism, differentiation and introgression in the period gene between *Lutzomyia intermedia* and *Lutzomyia whitmani*. *BMC Evolutionary Biology*, 6, 85.

- Mazzotta, G. M., Sandrelli, F., Zordan, M. A., Mason, M., Benna, C., Cisotto, P., et al. 2005. The clock gene period in the medfly *Ceratitis capitata*. *Genetics Research*, 86, 13-30.
- McAllister, B. F. & McVean, G. A. T. 2000. Neutral evolution of the sex-determining gene transformer in *Drosophila*. *Genetics*, 154, 1711.
- McClung, C. R. 2006. Plant circadian rhythms. *The Plant Cell Online*, 18, 792.
- Miyamoto, M. M. & Fitch, W. M. 1995. Testing species phylogenies and phylogenetic methods with congruence. *Systematic Biology*, 44, 64.
- Miyatake, T., Matsumoto, A., Matsuyama, T., Ueda, H. R., Toyosato, T. & Tanimura, T. 2002. The period gene and allochronic reproductive isolation in *Bactrocera cucurbitae*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269, 2467.
- Moraes, E. & Sene, F. 2007. Microsatellite and morphometric variation in *Drosophila gouveai*: the relative importance of historical and current factors in shaping the genetic population structure. *Journal of Zoological Systematics and Evolutionary Research*, 45, 336-344.
- Moraes, E., Yotoko, K., Manfrin, M., Solferini, V. & Sene, F. 2009. Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil. *Journal of Biogeography*, 36, 2136-2147.
- Moraes, E. M., Manfrin, M. H., Laus, A. C., Rosada, R. S., Bomfin, S. C. & Sene, F. M. 2004. Wing shape heritability and morphological divergence of the sibling species *Drosophila mercatorum* and *Drosophila paranaensis*. *Heredity*, 92, 466-473.
- Morales, A.C., Manfrin, M.H., Sene, F.M. 2004. Entendendo as relações evolutivas de duas espécies do "cluster" *buzzatii*: *Drosophila serido* e *Drosophila antonietae*, através da filogeografia e análise aloenzimática. *Resumos do 50º Congresso Brasileiro de Genética*: 400.
- Morales, A.C. 2005. Análise histórico-evolutiva de populações das espécies cactofílicas *Drosophila serido* e *D. antonietae* (Diptera, Drosophilidae). Tese de Doutorado. FFCLRP, USP.
- Moritz, C., Patton, J., Schneider, C. & Smith, T. 2000. Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, 31, 533-563.
- Morrone, J.J. 1993. Cladistic and biogeographic analyses of the weevil genus

- Listroderes* Schoenherr (Coleoptera: Curculionidae). *Cladistics*, 9, 397-411.
- Morrone, J.J. 2000. What is the Chacoan subregion? *Neotropica*, 46, 51-68.
- Morrone, J.J., Juñent, S. R. & Crisci, J. V. 1994. South American Beetles. *Nat. Geogr. Res. Explor*, 10(1), 104-115.
- Nachman, M. W. 2006. Detecting selection at the molecular level. *Evolutionary genetics: concepts and case studies (eds Fox CW, Wolf JB)*, 103–118.
- Nagaraja, J. & Ranganath, H. 2004. Molecular phylogeny of the *nasuta* subgroup of *Drosophila* based on 12S rRNA, 16S rRNA and COI mitochondrial genes, RAPD and ISSR polymorphisms. *Genes & Genetic Systems*, 79, 293-299.
- Nei, M. & Gojobori, T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, 3, 418.
- Neto, J.L.S.A. & Nery, J.T. 2005. Variabilidade e mudanças climáticas no Brasil e seus impactos regionais. In: Souza CRG, Suguio K, Oliveira AMS e de Oliveira PE. 2005. Quaternário do Brasil. Ed. Holos, Ribeirão Preto.
- Nosil, P. 2008. Ernst Mayr and the integration of geographic and ecological factors in speciation. *Biological Journal of the Linnean Society*, 95, 26-46.
- O'Grady, P. M., Clark, J. B. & Kidwell, M. G. 1998. Phylogeny of the *Drosophila saltans* species group based on combined analysis of nuclear and mitochondrial DNA sequences. *Molecular Biology and Evolution*, 15, 656.
- Oliveira C. 2008. Causes for sexual isolation in the *Drosophila buzzatii* species cluster. PhD Thesis. University of Arkansas. Fayetteville.
- Oliveira, C., Manfrin, M.H., Sene, F.M., Jackson, L.L. & Etges, W.J. Em preparação. Variations on a theme: diversification of cuticular hydrocarbons in a clade of cactophilic *Drosophila*.
- O'Neil, M. T. & Belote, J. M. 1992. Interspecific comparison of the transformer gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. *Genetics*, 131, 113.
- Orr, H. A., Masly, J. & Phadnis, N. 2007. Speciation in *Drosophila*: from phenotypes to molecules. *Journal of Heredity*, 98, 103.
- Panda, S., Hogenesch, J. B. & Kay, S. A. 2002. Circadian rhythms from flies to human. *Nature*, 417, 329-335.

- Patterson, J. T. & Alexander, M. L. 1952. *Drosophila wheeleri*, a new member of the *mulleri* subgroup. *Univ. Texas Publ.*, 5203, 129-136.
- Patterson, J. T. & Stone, W. S. 1952. Evolution in the genus *Drosophila*. *New York*.
- Pennington, R., Lavin, M., Prado, D., Pendry, C. & Pell, S. 2005. Climate change and speciation in Neotropical seasonally forest plants. *Tropical Forests & Global Atmospheric Change*, 191–198.
- Pennington, R. T., Lavin, M., Prado, D. E., Pendry, C. A., Pell, S. K. & Butterworth, C. A. 2004. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359, 515.
- Pennington, R. T., Prado, D. E. & Pendry, C. A. 2000. Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography*, 27, 261-273.
- Pereira, M. A. Q. R., Vilela, C. R. & Sene, F. M. 1983. Notes on breeding and feeding sites of some species of the *repleta* group of the genus *Drosophila* (Diptera, Drosophilidae). *Ciência e Cultura* 35, 1313-1319.
- Pfeiler, E., Erez, T., Hurtado, L. A. & Markow, T. A. 2007. Genetic differentiation and demographic history in *Drosophila pachea* from the Sonoran Desert. *Hereditas*, 144, 63-74.
- Philippe, H., Zhou, Y., Brinkmann, H., Rodrigue, N. & Delsuc, F. 2005. Heterotachy and long-branch attraction in phylogenetics. *BMC Evolutionary Biology*, 5, 50.
- Piccinali, R., Aguade, M. & Hasson, E. 2004. Comparative molecular population genetics of the Xdh locus in the cactophilic sibling species *Drosophila buzzatii* and *D. koepferae*. *Molecular Biology and Evolution*, 21, 141-152.
- Piccinali, R. V., Mascord, L. J., Barker, J. S. F., Oakeshott, J. G. & Hasson, E. 2007. Molecular population genetics of the alpha-esterase 5 gene locus in original and colonized populations of *Drosophila buzzatii* and its sibling *Drosophila koepferae*. *Journal of molecular evolution*, 64, 158-170.
- Posada, D. & Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics*, 14, 817.
- Prado, D. E. & Gibbs, P. E. 1993. Patterns of species distributions in the dry seasonal forests of South America. *Annals of the Missouri Botanical Garden*, 80, 902-927.

- Prance, G.T. 1973. Phytogeographic support for the theory of Pleistocene forest refuges in the Amazon Basin, based upon evidence from distribution pattern in *Caryocaraceae*, *Chrysobalanaceae*, *Dichapetalaceae* and *Lecythidaceae*. *Acta Amazon.* 3: 5-28.
- Presgraves, D. C., Balagopalan, L., Abmayr, S. M. & Orr, H. A. 2003. Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature*, 423, 715-719.
- Regier, J. C., Grant, M. C., Mitter, C., Cook, C. P., Peigler, R. S. & Rougerie, R. 2008. Phylogenetic relationships of wild silkmoths (Lepidoptera: Saturniidae) inferred from four protein coding nuclear genes. *Systematic Entomology*, 33, 219-228.
- Remsen, J. & DeSalle, R. 1998. Character congruence of multiple data partitions and the origin of the hawaiian Drosophilidae 1. *Molecular Phylogenetics and Evolution*, 9, 225-235.
- Remsen, J. & O'Grady, P. 2002. Phylogeny of *Drosophilinae* (Diptera: Drosophilidae), with comments on combined analysis and character support. *Molecular Phylogenetics and Evolution*, 24, 249-264.
- Reppert, S. M. & Weaver, D. R. 2002. Coordination of circadian timing in mammals. *Nature*, 418, 935-941.
- Rodriguez, F., Oliver, J., Marin, A. & Medina, J. R. 1990. The general stochastic model of nucleotide substitution. *Journal of theoretical biology*, 142, 485-501.
- Rodríguez-Trelles, F., Alarcón, L. & Fontdevila, A. 2000. Molecular evolution and phylogeny of the *buzzatii* complex (*Drosophila repleta* group): a maximum-likelihood approach. *Molecular Biology and Evolution*, 17, 1112.
- Ruiz, A., Cansian, A. M., Kuhn, G. C. S., Alves, M. A. R. & Sene, F. M. 2000. The *Drosophila serido* speciation puzzle: putting new pieces together. *Genetica*, 108, 217-227.
- Ruiz, A. & Wasserman, M. 1993. Evolutionary cytogenetics of the *Drosophila buzzatii* species complex. *Heredity*, 70, 582-596.
- Rull, V. I. 2008. Speciation timing and neotropical biodiversity: the Tertiary - Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology*, 17, 2722-2729.
- Russo, C., Takezaki, N. & Nei, M. 1995. Molecular phylogeny and divergence times of Drosophilidae species. *Molecular Biology and Evolution*, 12, 391.
- Sakai, T. & Ishida, N. 2001. Circadian rhythms of female mating activity governed by

- clock genes in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 9221.
- Salgado-Labouriau, M.L. 1994. História Ecológica da Terra. Ed. Edgard Blucher, São Paulo, 307pp.
- Santos, M.H. 2011. As rotas de dispersão de *Drosophila buzzatii* na América do Sul. Tese de Doutorado. Faculdade de Medicina de Ribeirão Preto, USP.
- Sawyer, L. A., Hennessy, J. M., Peixoto, A. A., Rosato, E., Parkinson, H., Costa, R., et al. 1997. Natural variation in a *Drosophila* clock gene and temperature compensation. *Science*, 278, 2117.
- Seehausen, O. 2004. Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, 19, 198-207.
- Sehgal, A., Price, J. L., Man, B. & Young, M. W. 1994. Loss of circadian behavioral rhythms and per RNA oscillations in the *Drosophila* mutant timeless. *Science*, 263, 1603.
- Sene, F. M., Pereira, M. A. Q. R. & Vilela, C. R. 1988. Contrasting patterns of differentiation inferred from traditional genetic markers in the process of speciation. *Pacific Science*, 42, 81-88.
- Silva, A. & Sene, F. 1991. Morphological geographic variability in *Drosophila serido* (Diptera, Drosophilidae). *Rev. Bras. Entomol*, 35, 455-468.
- Silva, J. 1995. Biogeographic analysis of the South American Cerrado avifauna. *Steenstrupia*, 21, 49-67.
- Silva-Bernardi, E. C. C., Morales, A. C., Sene, F. M. & Manfrin, M. H. 2006. Phylogenetic relationships in the *Drosophila fasciola* species subgroup (Diptera, Drosophilidae) inferred from partial sequences of the mitochondrial cytochrome oxidase subunit I (COI) gene. *Genetics and Molecular Biology*, 29, 566-571.
- 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. *Annals of the Entomological Society of America*, 87, 651-701.
- Soto, E., Soto, I., Carreira, V., Fanara, J. & Hasson, E. 2008. Host-related life history traits in interspecific hybrids of cactophilic *Drosophila*. *Entomologia Experimentalis et Applicata*, 126, 18-27.
- Spicer, G. S. 1995. Phylogenetic utility of the mitochondrial cytochrome oxidase gene: molecular evolution of the *Drosophila buzzatii* species complex. *Journal of*

- molecular evolution*, 41, 749-759.
- Starmer, W., Barker, J., Phaff, H. & Fogleman, J. 1986. Adaptations of *Drosophila* and yeasts: their interactions with the volatile 2-propanol in the cactus-microorganism - *Drosophila* model system. *Australian journal of biological sciences*, 39, 69.
- Stern, A., Doron-Faigenboim, A., Erez, E., Martz, E., Bacharach, E. & Pupko, T. 2007. Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic acids research*.
- Suguio, K., Sallun, A. E. M. & Soares, E. A. A. 2007. Período Quaternário: “Quo Vadis”? *Revista Brasileira de Geociências*, 35.
- Sun, S., Ting, C. T. & Wu, C. I. 2004. The normal function of a speciation gene, *Odysseus*, and its hybrid sterility effect. *Science*, 305, 81.
- Swanson, W. J., Nielsen, R. & Yang, Q. 2003. Pervasive adaptive evolution in mammalian fertilization proteins. *Molecular Biology and Evolution*, 20, 18.
- Swofford, D. L. 2002. *PAUP*. Phylogenetic analysis using parsimony (* and other methods)*. Version 4: Sinauer Associates, Sunderland, Massachusetts.
- Tamura, K., Dudley, J., Nei, M. & Kumar, S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24, 1596.
- Tamura, K. & Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10, 512.
- Tauber, E. & Kyriacou, C. P. 2008. Genomic approaches for studying biological clocks. *Functional Ecology*, 22, 19-29.
- Tauber, E., Roe, H., Costa, R., Hennessy, J. M. & Kyriacou, C. P. 2003. Temporal mating isolation driven by a behavioral gene in *Drosophila*. *Current Biology*, 13, 140-145.
- Taylor, N. & Zappi, D. 2004. Cacti of eastern Brazil. *Kew: Royal Botanic Gardens, Kew xii, 499p.-col. illus.. ISBN, 1842460560*.
- Templeton, A. R. 2005. Haplotype trees and modern human origins. *American Journal of Physical Anthropology*, 128, 33-59.
- Templeton, A.R. 2006. Population genetics and microevolutionary Theory. Hoboken, New Jersey: John Wiley & Sons. 705 p.
- Templeton, A. R. 2008. The reality and importance of founder speciation in evolution. *Bioessays*, 30, 470-479.

- Templeton, A. R., Sing, C. F., Kessling, A. & Humphries, S. 1988. A cladistic analysis of phenotype associations with haplotypes inferred from restriction endonuclease mapping. II. The analysis of natural populations. *Genetics*, 120, 1145.
- Thompson, J. D. & Higgins, D. G. 1994. Gibson TJ CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic acids research*, 22, 4673-4680.
- Thompson, J. N. & Pellmyr, O. 1991. Evolution of oviposition behavior and host preference in *Lepidoptera*. *Annual Review of Entomology*, 36, 65-89.
- Tidon-Sklorz R. 1992. Identificação da fauna de *Drosophila* (Diptera, Drosophilidae) da Cadeia do Espinhaço, e análise morfométrica das populações da espécie politípica *Drosophila serido*. Tese de Doutorado. Faculdade de Medicina de Ribeirão Preto, USP.
- Tidon, R. & Sene, F. M. 1988. A trap that retains and keeps *Drosophila* alive. *Drosophila Information Service*, 67, 90.
- Tidon-Sklorz, R. & Sene, F. M. 1995a. *Drosophila seriema* n. sp.: New member of the *Drosophila serido* (Diptera: Drosophilidae) superspecies taxon. *Annals of the Entomological Society of America*, 88, 139-142.
- Tidon-Sklorz, R. & Sene, F. M. 1995b. Evolution of the *buzzatii* cluster (*Drosophila repleta* group) in middle South America. 9, 71-85.
- Tidon-Sklorz, R. & Sene, F. M. 2001. *Drosophila antonietae* sp. n. and *Drosophila gouveai* sp. n.: Two new species of the *Drosophila serido* superspecies taxon (Diptera, Drosophilidae). *Iheringia, série Zoologia*, 90, 141-146.
- Tidon-Sklorz, R., Vilela, C., Sene, F. M. & Pereira, M. 1994. The genus *Drosophila* (Diptera, Drosophilidae) in the Serra do Cipó, State of Minas Gerais, Brazil. *Revista Brasileira de Entomologia (Brazil)*.
- Ting, C. T., Tsaur, S. C. & Wu, C. I. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 5313-5316.
- Ting, C. T., Tsaur, S. C., Wu, M. L. & Wu, C. I. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science*, 282, 1501.
- Tosi, D. & Sene, F. M. 1989. Further studies on chromosomal variability in the complex taxon *Drosophila serido* (Diptera, Drosophilidae). *Revista Brasileira de Genética*, 12, 729-745.

- Vanzolini, P. E. & Williams, E. E. 1970. South American anoles: the geographic differentiation and evolution of the *Anolis chrysolepis* species group (Sauria, Iguanidae). *Arquivos de Zoologia*, 19, 125-298.
- Vilela, C. R. 1983. A revision of the *Drosophila repleta* species group (Diptera: Drosophilidae). *Revista Brasileira de Entomologia*, 27, 1-114.
- Vilela, C. R., Sene, F. M. & Pereira, M. 1980. On the *Drosophila* fauna of Chaco and east slopes of the Andes in Argentina. *Revista Brasileira de Biologia*, 40, 837-841.
- Wang, R. L. & Hey, J. 1996. The speciation history of *Drosophila pseudoobscura* and close relatives: inferences from DNA sequence variation at the period locus. *Genetics*, 144, 1113.
- Ward, C.L. & Stone, W.S. 1952. Studies in the *repleta* group: The *melanopalpa* subgroup. *Univ. Texas Publ.*, 5204, 119-128
- Wasserman, M. 1960. Cytological and phylogenetic relationships in the *repleta* group of the genus *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 46, 842.
- Wasserman, M. 1962. Cytological studies of the *repleta* group of the genus *Drosophila*. V. *The mulleri* subgroup. *Univ. Texas Publ*, 6205, 85-117.
- Wasserman, M. 1982. Evolution of the *repleta* group. In: M. Ashburner, H. L. Carson & J. N. Thompson (Eds), *The Genetics and Biology of Drosophila*. Academic Press, New York, pp. 61-139.
- Wasserman, M. & Wasserman, F. 1992. Inversion polymorphism in island species of *Drosophila*. In: Hecht, M. K., MacIntyre, R. and Wallace, B. (eds). *Evolutionary Biology*, 26, 357-381.
- Webb, S. D. 1978. A history of savanna vertebrates in the New World. Part II: South America and the great interchange. *Annual Review of Ecology and Systematics*, 9, 393-426.
- Wu, C. I. 2001. The genic view of the process of speciation. *Journal of Evolutionary Biology*, 14, 851-865.
- Wuester, W., Ferguson, J. E., Quijada Mascareñas, A., Pook, C. E., Da Graca Salomao, M. & Thorpe, R. S. 2005. Tracing an invasion: landbridges, refugia, and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: *Crotalus durissus*). *Molecular Ecology*, 14, 1095-1108.
- Xia, X. & Xie, Z. 2001. DAMBE: software package for data analysis in molecular

- biology and evolution. *Journal of Heredity*, 92, 371.
- Yang, Z., Nielsen, R., Goldman, N. & Pedersen, A. M. K. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, 155, 431.
- Zander, R. H. 2004. Minimal values for reliability of bootstrap and jackknife proportions, decay index, and Bayesian posterior probability. *Phyloinformatics*, 2, 1-13.
- Zanella, F. C. V. 2002. Systematics and biogeography of the bee genus *Caenonomada* Ashmead, 1899 (Hymenoptera: Apidae: Tapinotaspidini). *Studies on Neotropical Fauna and Environment*, 37, 249-261.