INSTITUTO DE BIOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA ANIMAL

HENRIQUE DA ROCHA MOREIRA

HISTÓRIA EVOLUTIVA DE ESPÉCIES DO GRUPO *FLAVOPILOSA* DE *DROSOPHILA* (DIPTERA, DROSOPHILIDAE): UMA ABORDAGEM FILOGENÉTICA, FILOGEOGRÁFICA E TAXONÔMICA

PORTO ALEGRE
2020

HENRIQUE DA ROCHA MOREIRA

HISTÓRIA EVOLUTIVA DE ESPÉCIES DO GRUPO *FLAVOPILOSA* DE *DROSOPHILA* (DIPTERA, DROSOPHILIDAE): UMA ABORDAGEM FILOGENÉTICA, FILOGEOGRÁFICA E TAXONÔMICA

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Animal, Instituto de Biociências da Universidade Federal do Rio Grande do Sul, como requisito parcial à obtenção do título de Mestre em Biologia Animal.

Área de concentração: Biologia Comparada

Orientadora: Prof.ª Dr.ª Maríndia Deprá

Coorientadora: Prof.ª Dr.ª Lizandra Jaqueline Robe

PORTO ALEGRE
2020

HENRIQUE DA ROCHA MOREIRA

HISTÓRIA EVOLUTIVA DE ESPÉCIES DO GRUPO *FLAVOPILOSA* DE *DROSOPHILA* (DIPTERA, DROSOPHILIDAE): UMA ABORDAGEM FILOGENÉTICA, FILOGEOGRÁFICA E TAXONÔMICA

Aprovada em _____ de _____ de _____.

BANCA EXAMINADORA

_____

Dr. Filipe Michels Bianchi (UFRGS)

_____

Dr.ª Loreta Brandão de Freitas (UFRGS)

_____

Dr.ª Rebeca Zanini (UFRGS)

À minha avó, Maria Égide (in memoriam), pelo imensurável incentivo em seguir nesta jornada.

AGRADECIMENTOS

À minha mãe, pelo fundamental exemplo, incentivo e suporte para seguir meus sonhos e concluir esta importante etapa.

Ao meu irmão, pela disposição em ajudar nos momentos difíceis e comemorar nas conquistas.

À Maríndia, pela orientação, pelo conhecimento compartilhado e, principalmente, confiança no meu trabalho; pela compreensão, paciência, conversas... Enfim, por propiciar um ótimo ambiente de trabalho e por fazer essa etapa ser muito mais leve.

À Lizandra, pela orientação, confiança e pelo conhecimento compartilhado desde os semestres iniciais da minha graduação, na "boa e velha" FURG; minha eterna gratidão por ter ajudado a me encontrar nas inúmeras possibilidades que a Ciência oferece.

Às (aos) colegas e amigas(os) do Laboratório de *Drosophila*, em especial à Ane, Carol, Mayara, Natasha, Rebeca e Thays pela maior convivência, muito obrigado pela ajuda em inúmeros momentos, pelas conversas, pela amizade, pela paciência e pelas sagradas cervejinhas de sexta-feira (foram poucas...). Vocês foram fundamentais nesta etapa!

Às amigas da UFSM, Fran, Maiara e Tuca, pela amizade (também de muitos anos, desde a FURG...), pelo apoio nos momentos difíceis e pelas risadas nos momentos de lazer!

À Berê, Dani e Helena, pela ajuda com os meios de cultura, pelos momentos de descontração e por diariamente prepararem nosso principal combustível: o café!

Às (aos) amigas(os) de Caxias do Sul (Carol, Dani, Gilmar, Lisi, Luísa, Paula e Sabrina) e de Rio Grande (Ana, Fabi, Ju e Manu) pelo fundamental apoio e por se fazerem presentes, apesar da distância.

Aos professores dos Programas de Pós-Graduação em Biologia Animal, e em Genética e Biologia Molecular, pelo conhecimento compartilhado.

À CAPES e CNPq, pela bolsa de Mestrado concedida e pelo financiamento da minha pesquisa.

À todas e todos que contribuíram de alguma forma para o desenvolvimento da minha dissertação... MUITO OBRIGADO!

"Somos máquinas de sobrevivência – veículos robôs programados cegamente para preservar as moléculas egoístas conhecidas como genes. Esta é uma verdade que ainda me enche de surpresa."

(Richard Dawkins)

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS

°C	Graus celsius
µl	Microlitro(s)
µM	Micromolar(es)
π	Diversidade nucleotídica
AIC	Critério de informação de Akaike (*Akaike Information Criterion*)
amd	Gene Alfa metildopa-resistente
BI	Inferência bayesiana (*Bayesian Inference*)
bp	Pares de base (*base pairs*)
COI	Gene Citocromo C oxidase subunidade I
COII	Gene Citocromo C oxidase subunidade II
Ddc	Gene dopa descarboxilase
DNA	Ácido desoxirribonucleico
E	Gene *Ebony*
EBSP	Extended Bayesian Skyline Plot
en	Gene *engrailed*
F_{ST}	Índice de fixação entre populações
GTR+G	*General Time-Reversible with gamma distribution*
GTR+I	*General Time-Reversible with invariant sites*
h	Número de haplótipos
Hd	Diversidade haplotípica
He	Heterozigosidade esperada
i.e.	Isto é (*id est*)
K2P	Kimura 2-parâmetros
kl-2	Gene *male fertility factor kl-2*
kl-3	Gene *male fertility factor kl-3*

kl-5	Gene *male fertility factor kl-5*
km	Quilômetro(s)
kya	Mil anos atrás (*thousand years ago*)
LGM	Último Máximo Glacial (*Last Glacial Maximum*)
MCMC	Monte Carlo via Cadeias de Markov (*Markov chain Monte Carlo*)
MG	Minas Gerais
min	Minuto(s)
ml	Mililitro(s)
MS	Mato Grosso do Sul
mtDNA	DNA mitocondrial
Mya	Milhões de anos atrás (*million years ago*)
nDNA	DNA nuclear
ng	Nanograma(s)
N_e	Tamanho efetivo populacional
pp	Posterior probability
PR	Paraná
RS	Rio Grande do Sul
s	Segundo(s)
S	Número de sítios polimórficos
SC	Santa Catarina
sod	Gene *Cu/Zn-superóxido dismutase*
U	Unidade(s) de enzima
UMG	Último Máximo Glacial

RESUMO

HISTÓRIA EVOLUTIVA DE ESPÉCIES DO GRUPO *FLAVOPILOSA* DE *DROSOPHILA* (DIPTERA, DROSOPHILIDAE): UMA ABORDAGEM FILOGENÉTICA, FILOGEOGRÁFICA E TAXONÔMICA

Drosophilidae apresenta uma série de organismos-modelo para diferentes áreas da Ciência, incluindo a Ecologia e a Genética. Os drosofilídeos estão mundialmente distribuídos e utilizam diversos tipos de recursos ou hospedeiros. O grupo *flavopilosa* de *Drosophila*, em especial, possui alta especialização ecológica à flores do gênero *Cestrum*, utilizando esse hospedeiro como único sítio de alimentação e reprodução. No entanto, o hospedeiro possui um padrão descontínuo de floração ao longo do ano, o que torna o grupo *flavopilosa* um ótimo modelo para o entendimento de adaptações populacionais em face aos hábitos ecológicos de uma espécie. Nesta Dissertação, estudou-se a história evolutiva do grupo *flavopilosa* através de enfoques filogenéticos, filogeográficos e taxonômicos. Assim, no Capítulo II, as relações filogenéticas entre espécies brasileiras do grupo (*D. cestri*, *D. cordeiroi*, *D. flavopilosa*, *D. incompta e D. mariaehelenae*) foram inferidas através de análises filogenéticas por Inferência Bayesiana, utilizando três genes – Citocromo oxidase subunidade I (*COI*), alfametildopa (*Amd*) e dopa descarboxilase (*Ddc*). A aplicação da técnica de DNA barcode para o grupo foi revisada mediante inclusão de novas espécies e espécimes. Além de comprovar a eficiência do DNA barcode na identificação das espécies avaliadas, este estudo também permitiu recuperar a subdivisão do grupo, alocando *D. cestri* e *D. cordeiroi* junto à *D. flavopilosa* no subgrupo *flavopilosa* e *D. incompta* e *D. mariaehelenae* no subgrupo *nesiota*. Além disso, uma análise cronofilogenética datou a origem do grupo para cerca de 10 milhões de anos atrás, demonstrando que esta ocorreu antes da formação do istmo do Panamá. No Capítulo III, as dinâmicas populacionais associadas a ecologia restrita destas espécies foram avaliadas através de um enfoque filogeográfico comparativo. Para tanto, os genes *COI* e Ebony (*E*) foram amplificados para diferentes populações de *D. cestri* e *D. incompta* provenientes de Estados das regiões Sul e Sudeste do Brasil. Os resultados obtidos para ambas espécies e genes são concordantes e revelaram alta diversidade haplotípica ou alélica, com baixa diversidade nucleotídica e testes de neutralidade negativos, sugerindo expansões populacionais recentes. As redes de haplótipos de *COI* e as árvores filogenéticas dos genes concatenados, junto com testes de F_{ST}, Mantel e BAPS sugerem, ainda, uma estruturação populacional superficial. A história demográfica inferida por *skyline plots*, no entanto, indicou que a expansão populacional de *D. incompta* provavelmente teria começado muito tempo antes (há cerca de 1,5 milhão de anos) da expansão de *D. cestri* (há cerca de 300 mil anos). Este resultado sugere que similaridades encontradas podem ser uma consequência dos padrões ecológicos compartilhados por ambas espécies e as estratégias de sobrevivência das mesmas em face da efemeridade das flores de *Cestrum*. Neste caso, as espécies do grupo *flavopilosa* parecem estar organizadas em metapopulações, as quais são constituídas por subpopulações com altos níveis de migração.

Palavras-chave: *Drosophila cestri*; *Drosophila incompta*; especialização ecológica; filogeografia; metapopulação

ABSTRACT

EVOLUTIONARY HISTORY OF SPECIES OF THE *FLAVOPILOSA* GROUP OF *DROSOPHILA* (DIPTERA, DROSOPHILIDAE): A PHYLOGENETIC, PHYLOGEOGRAPHIC AND TAXONOMIC APPROACH

Drosophilidae presents several model organisms for different fields in Science, including Ecology and Genetics. In fact, drosophilids are worldwide distributed and use several kinds of resources or hosts to live. The *flavopilosa* group of the *Drosophila* genus, particularly, presents high ecological specialization to the flowers of *Cestrum* and uses this host as a unique site for both feeding and breeding. Nevertheless, these host plants display a disjunct pattern on the flourishing seasons, which make the *flavopilosa* group an excellent model for understanding population adaptations in face to ecological specializations. In this Thesis, the evolutionary history of species belonging to the *flavopilosa* group was evaluated with a phylogenetic, phylogeographic and taxonomic approaches. Therefore, in Chapther II, the phylogenetic relationships among Brazilian species of the group (*D. cestri*, *D. cordeiroi*, *D. flavopilosa*, *D. incompta* and *D. mariaehelenae*) were inferred through Bayesian inference using three genes – Cytochrome oxidase subunit I (*COI*), alphametyldopa (*Amd*) and dopa decarboxylase (*Ddc*). The effectiveness of DNA barcode technique for the group was re-evaluated through inclusion of new species and specimens. Besides proving the efficiency of DNA barcode in the identification of the evaluated species, this study also allowed to recover the subdivision of the group, placing *D. cestri* and *D. cordeiroi* with *D. flavopilosa* in the subgroup *flavopilosa* and *D. incompta* and *D. mariaehelenae* in the subgroup *nesiota*. Furthermore, a cronophylogenetic analysis dated the origin of the group for about 10 Mya, showing that this occurred before the closure of the Isthmus of Panama. In Chapter III, the population dynamics associated with the restricted ecology of these species was evaluated through a comparative phylogeographic approach. For this, the genes *COI* and Ebony (*E*) were amplified for different populations of *D. cestri* and *D. incompta* from States of the Southern and Southeast of Brazil. The results obtained for both species and genes are congruent and revealed high haplotype or allelic diversity, with low nucleotide diversity and negative results for the neutrality tests, suggesting recent population expansions. The haplotype networks for *COI* and the phylogenetic trees for the concatenated genes, together with the F_{ST}, Mantel and BAPS tests, further suggest a shallow population structure. The demographic history inferred by skyline plots, however, indicated that the population expansion for *D. incompta* probably started a long time before (about 1.5 Mya) than that of *D. cestri* (about 300 kya). This result suggests that the similarities encountered can be a consequence of ecological patterns shared by both species and their survival strategies in face of *Cestrum* ephemerity. In this case, the species belonging to the *flavopilosa* group are apparently organized into metapopulations, which are constituted by subpopulations with high levels of migration.

Keywords: *Drosophila cestri*; *Drosophila incompta*; ecological specialization; phylogeography; metapopulation

CAPÍTULO I – INTRODUÇÃO GERAL

1. FILOGEOGRAFIA: A DISTRIBUIÇÃO ESPACIAL DE LINHAGENS GENEALÓGICAS

Em meados do século XVIII, com os trabalhos pioneiros dos naturalistas Carolus Linnaeus e, principalmente, Georges-Louis Leclerc, conde de Buffon, começou-se a introduzir a concepção de que a vida na Terra não era rígida e imutável (COX; MOORE, 2005). Lineu, por exemplo, hipotetizou uma ilha central e primordial, da qual todos organismos migraram conforme as águas retrocederam, se estabelecendo e permanecendo até os dias atuais (BRIGGS; HUMPHRIES, 2004). Contudo, apesar de ainda acreditar em um único evento de criação, o pensamento hegemônico da época começou a ser refutado apenas após a publicação da ampla obra de Buffon – na qual o naturalista discute causas e hipóteses para os diferentes agrupamentos de espécies em diferentes regiões do mundo, introduzindo o conceito de "adaptação" aos diferentes ambientes (CAPONI, 2017).

Nesse contexto, à medida que novos exploradores e naturalistas realizavam expedições ao redor do mundo, teorias como o gradiente de diversidade do equador em direção aos polos e a Lei de Buffon foram sendo corroboradas, consolidando a biogeografia como uma importante ferramenta para a compreensão da biodiversidade (COX; MOORE, 2005). A biogeografia é uma disciplina no âmbito das ciências biológicas cujos objetivos centrais são a reconstrução dos padrões de distribuição da diversidade biológica e a identificação dos processos que moldam ou moldaram os mesmos durante o tempo.

Assim, tradicionalmente divide-se a biogeografia conforme a abordagem utilizada: a biogeografia ecológica, a qual explica a distribuição dos organismos em função de fatores bióticos (ecologia ou outras características de distintos organismos) ou ambientais (abióticos) atuais; e a biogeografia histórica, a qual busca entender a distribuição dos organismos sob uma perspectiva evolutiva e histórica, centrando-se em questões geológicas de milhões de anos atrás (MONGE-NÁJERA, 2008; MORRONE; CRISCI, 1995). No entanto, de acordo com SANMARTÍN (2012), a linha que separa essas duas principais abordagens tornou-se tênue nos últimos anos, uma vez que a biogeografia histórica começou a incorporar questões ecológicas em suas análises através da modelagem de nicho ecológico e novos métodos analíticos. Ademais, segundo VITT; CALDWELL (2013), com o progresso da biologia e sistemática molecular no fim do século XX, a biogeografia histórica acabou reinventando-se e emergindo em uma nova área: a filogeografia.

O termo "filogeografia intraespecífica", tradicionalmente reduzido a "filogeografia", foi

introduzido por AVISE et al. (1987) ao detectar que muitas espécies apresentam uma história filogenética geograficamente estruturada para genes mitocondriais. AVISE (1998) classifica a filogeografia como uma subdisciplina da biogeografia, a qual enfatiza a influência de fatores históricos ou de processos recorrentes na atual distribuição das linhagens de genes, utilizando para tanto dados de genética molecular e populacional, além de dados geológicos, demográficos, etológicos e filogenéticos. De fato, o principal foco da filogeografia são estudos que integram a genealogia à biogeografia em dimensões espaciais e temporais, com ênfase em aspectos históricos ou recorrentes, capazes de explicar a atual distribuição das linhagens genealógicas (AVISE, 2009).

Desde o seu estabelecimento, os estudos filogeográficos utilizam predominantemente o DNA mitocondrial (mtDNA) como marcador genético (BEHEREGARAY, 2008). Isso acontece principalmente devido à ausência de recombinação e a sua rápida evolução, aliadas ao menor tamanho efetivo populacional (N_e) em relação ao DNA nuclear (SUNNUCKS, 2000). Ou seja, por possuir cerca de ¼ do N_e em relação à marcadores nucleares devido à herança exclusivamente materna, o mtDNA proporciona a reconstrução e acesso à eventos evolutivos recentes sem um esforço amostral tão extensivo (HURST; JIGGINS, 2005).

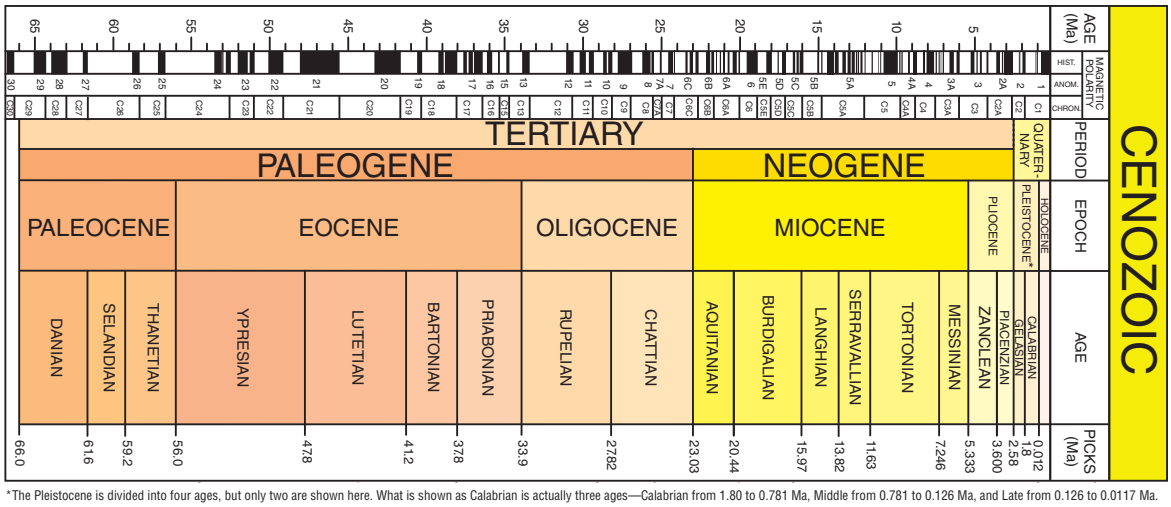
No entanto, apesar de sua fácil amplificação, o mtDNA também apresenta algumas desvantagens, como a própria uniparentalidade (a qual pode refletir apenas a história evolutiva das fêmeas); a presença de *Numts* (cópias de genes mitocondriais transpostas para o DNA genômico), as quais não são ortólogas às sequências comparadas; e a ocasional influência de endossimbiontes em artrópodes em seus processos de herança (por exemplo, o gênero de bactérias *Wolbachia*) (BENSASSON et al., 2001; BOURTZIS et al., 1996; DEN TEX et al., 2010; HURST; JIGGINS, 2005). Dessa forma, é essencial realizar uma análise concomitante e comparativa de marcadores mitocondriais e nucleares, recuperando diferentes aspectos da história evolutiva e demográfica das espécies (TEMPLETON, 2011).

Uma das principais aplicações da filogeografia é explicar os padrões atuais de distribuição da biodiversidade, tendo sido utilizada em inúmeros estudos relacionados à biologia da conservação e, particularmente, especiação (BEHEREGARAY, 2008). Além disso, métodos associados a filogeografia também facilitam o endereçamento de questões relacionadas às alterações históricas nos padrões de estruturação populacional e suas causas (HEWITT, 2000). Nesse sentido, é possível investigar uma série de questões referentes à diversidade genética e seu vínculo com processos históricos ou padrões geográficos, as quais incluem o papel de eventos evolutivos como, por exemplo, a vicariância, efeitos de *bottlenecks* e de expansões populacionais (ARBOGAST; KENAGY, 2001). Enquanto a biogeografia

histórica trabalha com eventos muito antigos na história geológica da Terra – como, por exemplo, as dinâmicas das placas tectônicas ao longo das Eras –, a filogeografia consegue acessar eventos em uma escala de tempo mais recente, frequentemente relacionados ao Período Quaternário (SANMARTÍN, 2012). Como reportado por TURCHETTO-ZOLET et al. (2013), na América do Sul, por exemplo, há um mosaico de padrões filogeográficos, no qual 57% dos tempos de divergência entre as linhagens são datados para o Quaternário. RULL (2008), por exemplo, aponta que os insetos foram impactados predominantemente pelas oscilações climáticas do Quaternário em detrimento aos eventos geológicos do Terciário.

1.1. EVENTOS PALEOGEOCLIMÁTICOS EM FILOGEOGRAFIA

Como reportado por BEHEREGARAY (2008), os tempos de divergência entre linhagens em estudos filogeográficos são frequentemente datados para os períodos Terciário e Quaternário. No entanto, apesar de ainda ser aceito, o termo Terciário vem caindo em desuso nos últimos anos e passou a ser substituído pelos períodos Paleogeno (mais antigo) e Neogeno (mais recente) (WALKER et al., 2018). Os períodos Paleogeno, Neogeno e Quaternário compõem o Cenozoico – a Era na qual atualmente nos encontramos – (Figura 1), marcada pela diversificação e domínio das angiospermas e dos mamíferos (ELIAS, 2013). De acordo com GULICK et al. (2019), estima-se que o início dessa Era tenha ocorrido após o impacto de um asteroide na península de Yucatán (evento que resultou na extinção dos dinossauros, há cerca de 66 milhões de anos).



*The Pleistocene is divided into four ages, but only two are shown here. What is shown as Calabrian is actually three ages—Calabrian from 1.80 to 0.781 Ma, Middle from 0.781 to 0.126 Ma, and Late from 0.126 to 0.0117 Ma.

Figura 1. Escala do tempo geológico da Terra, referente à Era Cenozoica. Retirado de WALKER et al. (2018).

Os períodos Paleogeno e Neogeno compreendem subdivisões do Terciário, e são

caracterizados por importantes eventos geológicos. Durante o Paleogeno (estendendo-se de 66 a 23 milhões de anos atrás), por exemplo, houve a colisão entre a Ásia e Índia, acarretando no soerguimento da cordilheira do Himalaia (SINGH, 2003); a colisão entre as placas Africana e Eurasiana, resultando na elevação dos Alpes (STECK; HUNZIKER, 1994); e a separação da América do Sul e Antártida, com a abertura da Passagem de Drake (LAWVER; GAHAGAN; COFFIN, 1992). Quanto ao período Neogeno (que se estende de 23 a 2,6 milhões de anos), destaca-se a elevação do Panamá e a consequente união das Américas do Sul e Norte (HAUG et al., 2001).

O período Quaternário corresponde ao atual período geológico e é dividido em duas épocas: o Pleistoceno e o Holoceno (WALKER et al., 2018). Enquanto o período Terciário é marcado pelos eventos geológicos, o período Quaternário é caracterizado por oscilações climáticas e períodos de glaciações intermitentes, em ciclos de aproximadamente 100.000 anos (HEWITT, 1996). As flutuações climáticas ao longo desse período tem ocasionado mudanças na distribuição espacial dos organismos, a qual varia de acordo com a topografia e a latitude dos habitats de cada espécie (HEWITT, 2000). Os principais estudos concentram-se no hemisfério Norte e apontam, por exemplo, a total cobertura das altas latitudes por mantos de gelo [como no caso do último máximo glacial (UMG), há cerca de 21.000 anos] e a consequente compressão das regiões tropicais e temperadas em direção ao equador (HEWITT, 1996, 2004; MIX; BARD; SCHNEIDER, 2001).

Na América do Sul, a principal fonte dos estudos paleoclimáticos consiste na análise e datação de material vegetal por meio dos isótopos ^{13}C e ^{12}C (SUGUIO, 1999). Durante o Quaternário, a região sudeste do Brasil experimentou uma série de oscilações climáticas rápidas, variando entre períodos frios - secos e úmidos, e com diferenças climáticas inclusive em localidades muito próximas (BEHLING; LICHTE, 1997). Há o registro, por exemplo, da substituição da atual Floresta Atlântica no sul e sudeste por campos de gramíneas com mais de 750km de extensão (BEHLING, 2002). Tais mudanças foram fundamentais para que processos evolutivos, como os eventos de gargalo de garrafa, efeito do fundador e expansões populacionais interferissem na demografia das populações e influenciassem seu potencial adaptativo (HEWITT, 2004).

2. DROSOPHILIDAE (INSECTA: DIPTERA)

Há mais de um século, diversas espécies de Drosophilidae vem sendo utilizadas como organismo modelo em inúmeras áreas do conhecimento biológico (KOHLER, 1993). Em uma

perspectiva histórica, o trabalho seminal de CASTLE et al. (1906) pode ser considerado o marco inicial para o presente *status*. Este trabalho teve suas origens ainda no início do século XX, quando da redescoberta dos trabalhos de MENDEL (1866), e visava descrever características reprodutivas e a biologia geral de *Drosophila melanogaster* Meigen, 1830 (DUNN, 1965). Castle acabou por influenciar o posterior trabalho de MORGAN (1910), o qual popularizou *Drosophila* Fallén, 1823 no meio científico ao ganhar o Prêmio Nobel de Medicina ou Fisiologia de 1933.

Além de sua aplicabilidade como organismo modelo, Drosophilidae constitui uma família de dípteros vastamente distribuídos nos mais diversos habitats e biomas, podendo ser encontrada desde as tundras até os trópicos e desde o nível do mar até grandes altitudes (THROCKMORTON, 1975). No entanto, não há um padrão geral para a sua ocorrência, uma vez que há desde espécies endêmicas a apenas uma ilha até espécies cosmopolitas (MARKOW; O'GRADY, 2006). Essa diversidade possivelmente está relacionada à capacidade de seus membros explorarem, em diferentes níveis de especialização, os mais variados recursos. Há registros de espécies, por exemplo, utilizando frutos (ATKINSON; SHORROCKS, 1977), flores (SCHMITZ; VALENTE, 2019), cactos (MANFRIN; SENE, 2006), fungos (COURTNEY; KIBOTA; SINGLETON, 1990) e até guano de morcego (TOSI et al., 1990) como substrato para algum ou vários de seus estágios de vida. De fato, as múltiplas variedades de substratos explorados, associados com a sensibilidade à evolução de nicho, parece fornecer uma ótima explicação para o sucesso evolutivo da família (ROBE; LORETO; VALENTE, 2010) e sua ampla distribuição (THROCKMORTON, 1975).

Drosophilidae compreende mais de 4 mil espécies, divididas em Steganinae e Drosophilinae, com oito e 31 gêneros, respectivamente (BÄCHLI, 2020; MARKOW; O'GRADY, 2006). Dentre esses, *Drosophila* (Drosophilinae) é um dos mais diversos, contando com 1.639 espécies formalmente descritas (BÄCHLI, 2020). Taxonomistas usualmente dividem a família em radiações (ou linhagens), representando eventos de especiação múltipla e suas subsequentes diversificações (Fig. 2) (THROCKMORTON, 1975).

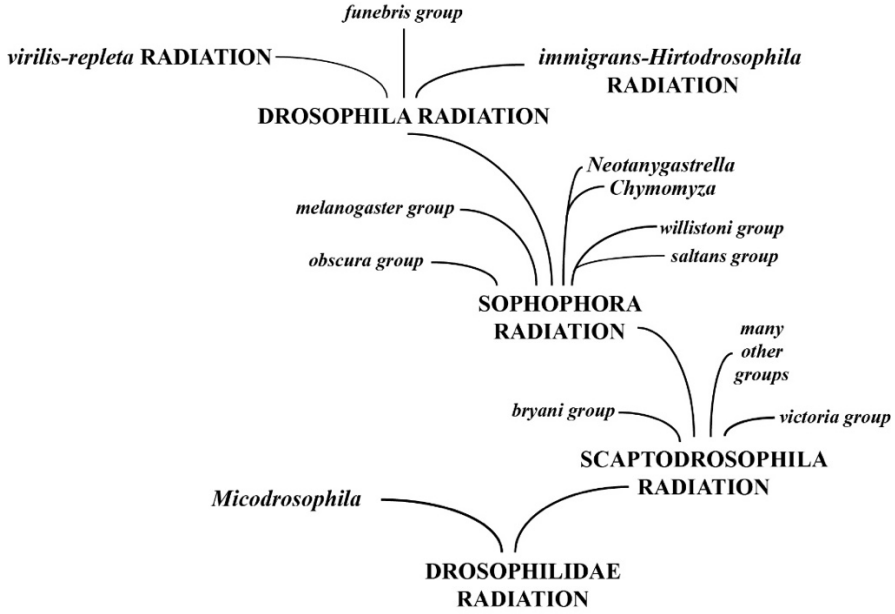


Figura 2. Representação das relações filogenéticas em Drosophilidae reconstruída com base em caracteres morfológicos por THROCKMORTON (1975). Resumido e adaptado de THROCKMORTON (1975).

Com auxílio de análises moleculares, no entanto, foi possível reclassificar as diferentes radiações de *Drosophila* ao nível de subgênero (Quadro 1) (YASSIN, 2013). Essa foi a primeira tentativa de ajuste em larga escala da taxonomia de *Drosophila* aos requerimentos de monofilia recíproca impostos pela sistemática filogenética. Cada uma destas radiações ou subgêneros é ainda subdividida em grupos e subgrupos de espécies, os quais agrupam espécies morfologicamente relacionadas (GRIMALDI, 1987, 1990).

Quadro 1. Mudanças propostas por YASSIN (2013) quanto à taxonomia das radiações no gênero *Drosophila*. Retirado de O'GRADY & DESALLE (2018).

Throckmorton (1975)	**Yassin (2013)**
Hawaiian *Drosophila*	*Idiomya* (gênero)
Radiação *virilis-repleta*	*Drosophila* (*Siphlodora*)
Radiação *immigrans-tripunctata*	*Drosophila* (*Drosophila*)
Drosophila (*Sophophora*)	*Drosophila* (*Sophophora*)
Drosophila (*Dorsilopha*)	*Drosophila* (*Dorsilopha*)

2.1. ESTUDOS FILOGEOGRÁFICOS EM DROSOPHILIDAE

Uma revisão com enfoque cienciométrico (Apêndice I) mostrou que, apesar da ampla utilização de espécies de drosofilídeos como organismo modelo na Ciência, há poucos estudos filogeográficos que utilizam Drosophilidae como espécies-alvo. Há, ainda, uma lacuna nos

táxons-alvo e nas regiões biogeográficas avaliadas – ou seja, alguns grupos de espécies são predominantemente estudados (como é o caso dos grupos *melanogaster*, *virilis* e *repleta* de *Drosophila*) enquanto outros praticamente não o são, e há um viés para estudos nas regiões Neártica e Paleártica, no qual a região Etiópica aparece como a menos explorada. Essa diversidade ainda não investigada possivelmente abriga a resolução de padrões ou questões filogeográficas que ainda são pouco compreendidas, tanto em relação às espécies quanto às regiões biogeográficas como um todo.

Para o Neotrópico, os resultados obtidos até o momento indicam, por exemplo, que diversas espécies de drosofilídeos foram severamente afetadas pelas oscilações climáticas durante o Pleistoceno (RULL, 2008). Dentre as espécies estudadas, as que apresentam dependência em climas amenos e habitats úmidos apresentam sinais de eventos de gargalo genético ou contrações populacionais durante os períodos glaciais, os quais são seguidos por expansões populacionais em períodos interglaciais [como é o caso de *D. americana* Spencer, 1983; *D. birchii* Dobzhansky & Mather, 1961; *D. maculifrons* Duda, 1927; *D. ornatifrons* Duda, 1927; and *D. serrata* Malloch, 1927 (DE RÉ et al., 2014a; GUSTANI et al., 2015; KELEMEN; MORITZ, 1999; SILLERO et al., 2014)]. Por outro lado, as espécies que usualmente habitam ambientes desérticos apresentam sinais contrastantes, com expansões datadas para períodos glaciais [como é o caso do complexo de espécies *buzzatii* – *D. buzzatii* Patterson & Wheeler, 1942 e *D. gouveai* Tidon-Sklorz & Sene, 2001 (BRITO; MANFRIN; SENE, 2002; MORAES et al., 2009)]. Em face à anterior incompreensão dos reais impactos das flutuações climáticas do Pleistoceno na biodiversidade Neotropical (RULL, 2011), esses resultados demonstram que as espécies de Drosophilidae podem ser utilizadas como modelo para o estudo de questões mais amplas, como o efeito de eventos paleogeográficos ou paleoclimáticos na fauna de diferentes regiões.

2.2. O GRUPO *FLAVOPILOSA* DE *DROSOPHILA*

Entre os diversos grupos de espécies incluídos dentro de *Drosophila*, o grupo *flavopilosa* se destaca pela notável especialização ecológica, uma vez que suas espécies apresentam uma associação íntima e restrita às flores do gênero *Cestrum* L. (Solanaceae) (BRNCIC, 1966). O grupo pertence à radiação *virilis-repleta* do subgênero *Drosophila* (BÄCHLI, 2020; ROBE et al., 2005), subgênero *Siphlodora* sensu YASSIN (2013), embora o seu posicionamento dentro deste clado seja ainda controverso (DE RÉ et al., 2017). Atualmente, o grupo é composto por um complexo de 17 espécies crípticas e endêmicas dos Neotrópicos,

divididas nos subgrupos *flavopilosa* e *nesiota* (Quadro 2), distinguidas apenas após análise da genitália interna dos machos (BÄCHLI, 2020; ROBE et al., 2013; WHEELER; TAKADA; BRNCIC, 1962).

Quadro 2. Taxonomia do grupo *flavopilosa* e divisão de espécies nos subgrupos (BÄCHLI, 2020; GUILLÍN; RAFAEL, 2018).

Espécies não alocadas em subgrupos	**Subgrupo *flavopilosa***	**Subgrupo *nesiota***
D. (Drosophila) cestri	*D. (Drosophila) acroria*	*D. (Drosophila) gentica*
D. (Drosophila) cordeiroi	*D. (Drosophila) crossoptera*	*D. (Drosophila) incompta*
D. (Drosophila) korefae	*D. (Drosophila) flavopilosa*	*D. (Drosophila) mariaehelenae*
D. (Drosophila) melina	*D. (Drosophila) hollisae*	*D. (Drosophila) nesiota*
D. (Drosophila) pseudokorefae	*D. (Drosophila) lauta*	
D. (Drosophila) sisa		
D. (Drosophila) suni		
D. (Drosophila) taxohuaycu		

As espécies pertencentes a esse grupo desenvolveram uma série de adaptações que possibilitam explorar de uma forma ampla e eficiente as flores de *Cestrum*. Entre as adaptações morfológicas, por exemplo, os membros do grupo apresentam a coloração do corpo em tons de verde/amarelo (críptica com as flores de *Cestrum*) e um pequeno a médio porte, com a presença de espinhos no órgão ovipositor das fêmeas para escarificar as flores e realizar a postura de ovos (WHEELER; TAKADA; BRNCIC, 1962).

Quanto às adaptações fisiológicas, os ovos são postos em um estágio avançado de desenvolvimento embrionário, o qual é muito importante em virtude da efemeridade das flores (AGUILAR; GALETTO, 2004; BRNCIC, 1983; WHEELER; TAKADA; BRNCIC, 1962). Além disso, a presença de toxinas em espécimes do gênero *Cestrum* é amplamente conhecida (PEARCE et al., 1992), o que deve acarretar em outras adaptações bioquímicas e moleculares que ainda não são totalmente compreendidas.

Se, por um lado, as relações ecológicas dessas espécies com seu gênero hospedeiro são razoavelmente conhecidas, as suas relações filogenéticas são predominantemente baseadas em caracteres morfológicos e não incluem todas as espécies do grupo (ROBE et al., 2013). Até o momento, de fato, há sequências de DNA disponíveis apenas para quatro das 17 espécies (CLARK et al., 2016), todas com ocorrência no Brasil. De acordo com GOTTSCHALK; HOFMANN; VALENTE (2008), no Brasil há registros de seis espécies de *Drosophila* do grupo *flavopilosa* (Figura 3): *D. cestri* Brncic (1978), *D. cordeiroi* Brncic (1978), *D. flavopilosa* Frey (1919), *D. hollisae* Vilela (1992), *D. incompta* Wheeler (1962) e *D. mariaehelenae* Vilela (1984), algumas das quais são frequentemente encontradas em simpatria e sintopia (BRNCIC,

1978; HOFMANN; NAPP, 1984).

Figura 3. Mapa da distribuição conhecida das espécies do grupo *flavopilosa* de *Drosophila* no Neotrópico. Retirado de ROBE et al. (2013). Espécies que ocorrem no Brasil estão indicadas por retângulos vermelhos.

Apesar da grande especialização ecológica, é possível notar diferenças quanto ao conjunto de espécies de *Cestrum* colonizadas pelas espécies do grupo *flavopilosa* no Brasil (Quadro 3) (SANTOS; VILELA, 2005). Neste caso, *D. flavopilosa* e *D. hollisae* apresentam maior grau de especificidade, ao serem encontradas em apenas duas espécies de *Cestrum*, enquanto *D. incompta* e *D. cordeiroi* colonizam uma ampla variedade de espécies do gênero, além de *Sessea brasiliensis* Toledo (atualmente, considerada uma espécie de *Cestrum*) (CARVALHO; SCHNOOR, 1993; SANTOS; VILELA, 2005; SOARES; VIGNOLI-SILVA; MENTZ, 2007). As espécies também diferem em termos de amplitude de distribuição: enquanto *D. flavopilosa* e *D. incompta* apresentam uma distribuição ampla, mas aparentemente disjunta, nos Neotrópicos, *D. cestri* e *D. cordeiroi* parecem apresentar uma distribuição restrita

ao sul do Brasil e norte do Uruguai (BÄCHLI, 2020).

Quadro 3. Lista de espécies do grupo *flavopilosa* de *Drosophila* encontradas no Brasil e seus respectivos hospedeiros de *Sessea* e *Cestrum*, de acordo com SANTOS & VILELA (2005).

D. cestri	***D. cordeiroi***	***D. flavopilosa***	***D. hollisae***	***D. incompta***	***D. mariahelenae***
C. calycinum	*C. calycinum*	*C. parqui*	*C. schlechtendalii*	*C. amictum*	*C. amictum*
C. corymbosum	*C. corymbosum*	*C. euanthes*	*C. sendtnerianum*	*C. calycinum*	*C. intermedium*
C. intermedium	*C. intermedium*			*C. corymbosum*	*C. nocturnum*
C. parqui	*C. nocturnum*			*C. intermedium*	*C. schlechtendalii*
C. schlechtendalii	*C. parqui*			*C. nocturnum*	*C. sendtnerianum*
S. brasiliensis	*C. schlechtendalii*			*C. parqui*	
	C. sendtnerianum			*C. schlechtendalii*	
	S. brasiliensis			*C. sendtnerianum*	
				S. brasiliensis	

Nos casos de sintopia, comportamentos ecológicos e reprodutivos distintos entre as espécies parecem ter garantido o sucesso na exploração dos recursos. Por exemplo, ao encontrar *D. cestri* e *D. incompta* associadas às mesmas flores (em *C. parqui* e *C. calycinum*), BRNCIC (1983) observou que enquanto *D. cestri* realiza a oviposição de apenas um ovo dentro de flores fechadas, *D. incompta* opta pelas flores abertas, escarificando a parte interna das pétalas e depositando até 12 ovos por flor. Além disso, segundo SEPEL et al. (2000), há uma flutuação sazonal nas populações, onde é possível verificar a predominância de eclosão de indivíduos adultos de *D. incompta* em períodos quentes e secos, enquanto *D. cestri* possui maior frequência em épocas frias e úmidas.

Os diferentes níveis de variação entre hospedeiros e entre padrões de distribuição geográfica permitem inferir que, possivelmente, existam variações no grau de compartilhamento de histórias evolutivas dentro do grupo. No entanto, dado o padrão de ocorrência tipicamente efêmero e descontínuo de *Cestrum* ao longo do ano, é provável que todas espécies do grupo *flavopilosa* tenham enfrentado uma pressão seletiva em prol de estratégias que possibilitem a sobrevivência em momentos de escassez de recursos, que são comuns para muitas das espécies hospedeiras. De fato, SANTOS; VILELA (2005), ao comparar quatro espécies de *Cestrum*, descrevem a ausência de um padrão comum de floração e persistência das flores, com picos curtos de florescimento e períodos longos de ausência ou presença dispersa (Figura 4).

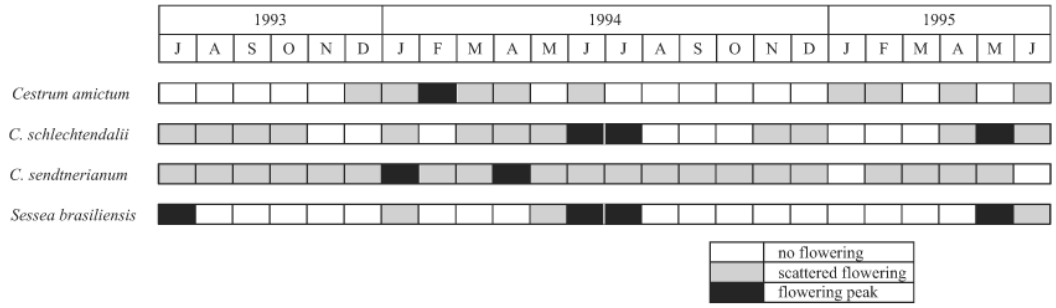


Figura 4. Fenologia comparada do florescimento de quatro espécies de *Cestrum*, monitoradas entre julho de 1993 e junho de 1995 em São Paulo, SP (SANTOS; VILELA, 2005).

Ao encontrar indivíduos com aparência mais senil em determinados períodos do ano, HOFMANN (1985) hipotetizou que alguns poucos adultos têm a capacidade de sobreviver ao longo do inverno, de modo a fundar novas populações quando o período de escassez terminar. No entanto, DE RÉ et al. (2014) consideram que os altos níveis de polimorfismo encontrados no DNA mitocondrial de *D. incompta* são prováveis consequências de migrações associadas a retrações e expansões populacionais cíclicas em busca de recurso. De fato, com base em dois genes mitocondriais e um gene nuclear, DE RÉ (2016) encontrou indícios de fluxo gênico recorrente e baixa estruturação populacional em populações brasileiras de *D. incompta*. Dessa maneira, a hipótese de trabalho da presente dissertação considera que as espécies do grupo utilizam a estratégia de migração como resposta aos períodos de limitação de recurso. Assim, nesta dissertação pretendeu-se caracterizar similaridades ou incongruências entre as histórias evolutivas de *D. incompta* e *D. cestri*, revelando padrões temporais e espaciais, bem como o impacto de eventos históricos ou processos evolutivos recorrentes, na diversificação e distribuição do grupo *flavopilosa* de *Drosophila* nos Neotrópicos.

Dada a ecologia restrita do grupo *flavopilosa* de *Drosophila*, junto à característica efêmera do recurso explorado e a proximidade filogenética junto a sobreposição geográfica parcial em face às diferenças em amplitude de distribuição, pode-se dizer que estas duas espécies constituem um excelente modelo para estudo de adaptações ecológicas ou demográficas para a especialização ecológica. Assim, torna-se possível não apenas testar padrões filogeográficos em geral, mas também elucidar eventos históricos ou recorrentes que atuam na evolução de espécies crípticas e com alta especificidade ecológica.

3. OBJETIVOS

3.1. OBJETIVO GERAL

Auxiliar na compreensão da história evolutiva e da influência da especialização ecológica na dinâmica populacional das espécies pertencentes ao grupo *flavopilosa* do gênero *Drosophila*, através de um enfoque filogenético, filogeográfico e taxonômico.

3.2. OBJETIVOS ESPECÍFICOS

- Inferir a posição filogenética de *Drosophila mariaehelenae* dentro do grupo *flavopilosa*, avaliando a monofilia do grupo e sua subdivisão taxonômica, bem como o cenário espaço-temporal associado a sua diversificação (Capítulo II);
- Determinar os padrões temporais e espaciais de distribuição da diversidade genética de duas espécies do grupo *flavopilosa*, caracterizando as causas das similaridades e/ou incongruências entre suas histórias evolutivas e avaliando a influência da especialização ecológica na dinâmica populacional de espécies de ecologia restrita (Capítulo III).

4. REFERÊNCIAS BIBLIOGRÁFICAS

AGUILAR, R.; GALETTO, L. Effects of forest fragmentation on male and female reproductive success in Cestrum parqui (Solanaceae). **Oecologia**, v. 138, n. 4, p. 513–520, mar. 2004.

ARBOGAST, B. S.; KENAGY, G. J. Comparative phylogeography as an integrative approach to historical biogeography. **Journal of Biogeography**, v. 28, n. 7, p. 819–825, 1 jul. 2001.

ATKINSON, W.; SHORROCKS, B. Breeding site specificity in the domestic species of Drosophila. **Oecologia**, v. 29, n. 3, p. 223–232, 1 set. 1977.

AVISE, J. C. et al. Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. **Annual Review of Ecology and Systematics**, v. 18, n. 1, p. 489–522, 1987.

AVISE, J. C. The history and purview of phylogeography: a personal reflection. **Molecular Ecology**, v. 7, n. 4, p. 371–379, 1 abr. 1998.

AVISE, J. C. Phylogeography: retrospect and prospect. **Journal of Biogeography**, v. 36, n. 1, p. 3–15, 1 jan. 2009.

BÄCHLI, G. **TaxoDros: The Database on Taxonomy of Drosophilidae**. Disponível em: <https://www.taxodros.uzh.ch/>. Acesso em: 7 jan. 2020.

BEHEREGARAY, L. B. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. **Molecular Ecology**, v. 17, n. 17, p. 3754–3774, set. 2008.

BEHLING, H. South and southeast Brazilian grasslands during Late Quaternary times: a synthesis. **Palaeogeography, Palaeoclimatology, Palaeoecology**, Reconstruction and Modeling of grass-dominated ecosystems. v. 177, n. 1, p. 19–27, 5 jan. 2002.

BEHLING, H.; LICHTE, M. Evidence of Dry and Cold Climatic Conditions at Glacial Times in Tropical Southeastern Brazil. **Quaternary Research**, v. 48, n. 3, p. 348–358, 1 nov. 1997.

BENSASSON, D. et al. Mitochondrial pseudogenes: evolution's misplaced witnesses. **Trends in Ecology & Evolution**, v. 16, n. 6, p. 314–321, 1 jun. 2001.

BOURTZIS, K. et al. Wolbachia Infection and Cytoplasmic Incompatibility in Drosophila Species. **Genetics**, v. 144, n. 3, p. 1063–1073, 1 nov. 1996.

BRIGGS, J. C.; HUMPHRIES, C. J. Early Classics. In: LOMOLINO, M. V.; SAX, D. F.; BROWN, J. H. (Eds.). . **Foundations of Biogeography: Classic Papers with Commentaries**. Chicago, IL: The University of Chicago Press, 2004. p. 5–266.

BRITO, R. A. DE; MANFRIN, M. H.; SENE, F. M. Mitochondrial DNA phylogeography of Brazilian populations of Drosophila buzzatii. **Genetics and Molecular Biology**, v. 25, n. 2, p. 161–171, 2002.

BRNCIC, D. Ecological and Cytogenetic Studies of Drosophila Flavopilosa, a Neotropical Species Living in Cestrum Flowers. **Evolution**, v. 20, n. 1, p. 16–29, 1 mar. 1966.

BRNCIC, D. A note on the flavopilosa group of species of Drosophila in Rio Grande do Sul, with the description of two new species (Diptera, Drosophilidae). **Revista Brasileira de Biologia**, v. 38, p. 647–651, 1978.

BRNCIC, D. The flavopilosa group of species as an example of flower-breeding species. In: ASHBURNER, M.; CARSON, H. L.; THOMPSON, J. N. (Eds.). . **The genetics and biology of Drosophila**. New York, NY: Academic Press, 1983. p. 360–377.

CAPONI, G. The discontinuity between humans and animals in Buffon's Natural history. **História, Ciências, Saúde-Manguinhos**, v. 24, n. 1, p. 59–74, jan. 2017.

CARVALHO, L. A. F.; SCHNOOR, A. Sessea Carvalho et Schnoor - nova seção para o gênero Cestrum (Solanaceae). **Rodriguésia**, v. 45–49, n. 71–75, p. 15–24, 1993.

CASTLE, W. E. et al. The Effects of Inbreeding, Cross-Breeding, and Selection upon the Fertility and Variability of Drosophila. **Proceedings of the American Academy of Arts and Sciences**, v. 41, n. 33, p. 731–786, 1906.

CLARK, K. et al. GenBank. **Nucleic Acids Research**, v. 44, n. D1, p. D67–D72, 4 jan. 2016.

COURTNEY, S. P.; KIBOTA, T. T.; SINGLETON, T. A. Ecology of Mushroom-feeding Drosophilidae. In: BEGON, M.; FITTER, A. H.; MACFADYEN, A. (Eds.). . **Advances in Ecological Research**. [s.l.] Academic Press, 1990. v. 20p. 225–274.

COX, C. B.; MOORE, P. D. **Biogeography: An Ecological and Evolutionary Approach**. 7. ed. Malden, MA: Wiley-Blackwell, 2005.

DE RÉ, F. C. et al. Brazilian populations of Drosophila maculifrons (Diptera: Drosophilidae): low diversity levels and signals of a population expansion after the Last Glacial Maximum. **Biological Journal of the Linnean Society**, v. 112, n. 1, p. 55–66, 1 maio 2014a.

DE RÉ, F. C. et al. Characterization of the complete mitochondrial genome of flower-breeding Drosophila incompta (Diptera, Drosophilidae). **Genetica**, v. 142, n. 6, p. 525–535, 1 dez. 2014b.

DE RÉ, F. C. **Evolução molecular e padrões macro e microevolutivos em Drosophila incompta (Diptera, Drosophilidae)**. Dissertation—Santa Maria, RS: Universidade Federal de Santa Maria, mar. 2016.

DE RÉ, F. C. et al. Inferring the phylogenetic position of the Drosophila flavopilosa group: Incongruence within and between mitochondrial and nuclear multilocus datasets. **Journal of Zoological Systematics and Evolutionary Research**, v. 55, n. 3, p. 208–221, 1 ago. 2017.

DEN TEX, R.-J. et al. Nuclear copies of mitochondrial genes: another problem for ancient DNA. **Genetica**, v. 138, n. 9, p. 979–984, 1 out. 2010.

DUNN, L. C. **William Ernest Castle 1867-1962**. Washington, D.C.: National Academy of Sciences, 1965.

ELIAS, S. A. Earth History. In: **Reference Module in Earth Systems and Environmental Sciences**. [s.l.] Elsevier, 2013.

GOTTSCHALK, M. S.; HOFMANN, P. R. P.; VALENTE, V. L. S. Diptera, Drosophilidae: historical occurrence in Brazil. **Check List**, v. 4, n. 4, p. 485–518, 1 nov. 2008.

GRIMALDI, D. A. Phylogenetics and taxonomy of Zygothrica (Diptera: Drosophilidae). v. 186, p. 103–268, 1987.

GRIMALDI, D. A. A phylogenetic, revised classification of genera in the Drosophilidae (Diptera). v. 197, p. 1–139, 1990.

GULICK, S. P. S. et al. The first day of the Cenozoic. **Proceedings of the National Academy of Sciences**, v. 116, n. 39, p. 19342–19351, 24 set. 2019.

GUSTANI, E. C. et al. Demographic Structure and Evolutionary History of Drosophila ornatifrons (Diptera, Drosophilidae) from Atlantic Forest of Southern Brazil. **Zoological Science**, v. 32, n. 2, p. 141–150, abr. 2015.

HAUG, G. H. et al. Role of Panama uplift on oceanic freshwater balance. **Geology**, v. 29, n. 3, p. 207–210, 1 mar. 2001.

HEWITT, G. The genetic legacy of the Quaternary ice ages. **Nature**, v. 405, n. 6789, p. 907–913, jun. 2000.

HEWITT, G. M. Some genetic consequences of ice ages, and their role in divergence and speciation. **Biological Journal of the Linnean Society**, v. 58, n. 3, p. 247–276, 1 jul. 1996.

HEWITT, G. M. Genetic consequences of climatic oscillations in the Quaternary. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 359, n. 1442, p. 183–195, 29 fev. 2004.

HOFMANN, P. R. P. Variabilidade genética em espécies de nicho ecológico restrito. **Ciência e Cultura**, v. 37, p. 579–581, 1985.

HOFMANN, P. R. P.; NAPP, M. Genetic-environmental relationships in Drosophila incompta, a species of restricted ecology. **Brazilian Journal of Genetics**, v. 7, n. 1, p. 21–39, 1984.

HURST, G. D. D.; JIGGINS, F. M. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. **Proceedings of the Royal Society B: Biological Sciences**, v. 272, n. 1572, p. 1525–1534, 7 ago. 2005.

KELEMEN, L.; MORITZ, C. Comparative Phylogeography of a Sibling Pair of Rainforest Drosophila Species (drosophila Serrata and D. Birchii). **Evolution**, v. 53, n. 4, p. 1306–1311, 1999.

KOHLER, R. E. Drosophila: A life in the laboratory. **Journal of the History of Biology**, v. 26, n. 2, p. 281–310, 1 jun. 1993.

LAWVER, L. A.; GAHAGAN, L. M.; COFFIN, M. F. The Development of Paleoseaways Around Antarctica. In: **The Antarctic Paleoenvironment: A Perspective on Global Change: Part One**. [s.l.] American Geophysical Union (AGU), 1992. p. 7–30.

MANFRIN, M. H.; SENE, F. M. Cactophilic Drosophila in South America: A Model for Evolutionary Studies. **Genetica**, v. 126, n. 1, p. 57–75, 1 jan. 2006.

MARKOW, T. A.; O'GRADY, P. M. Phylogenetic relationships of Drosophilidae. In: **Drosophila: A Guide to Species Identification and Use**. Cambridge, MA: Academic Press, 2006. p. 3–64.

MENDEL, G. Versuche uber pflanzen-hybriden. **Verhandlungen des naturforschenden Vereins in Brünn**, v. 4, p. 3–47, 1866.

MIX, A. C.; BARD, E.; SCHNEIDER, R. Environmental processes of the ice age: land, oceans, glaciers (EPILOG). **Quaternary Science Reviews**, v. 20, n. 4, p. 627–657, 1 fev. 2001.

MONGE-NÁJERA, J. ECOLOGICAL BIOGEOGRAPHY: A REVIEW WITH EMPHASIS ON CONSERVATION AND THE NEUTRAL MODEL BIOGEOGRAFIA ECOLOGICA: REVISION CON ENFASIS EN CONSERVACION Y EL "MODELO NEUTRAL". **Gayana (Concepción)**, v. 72, n. 1, p. 102–112, jun. 2008.

MORAES, E. M. et al. Phylogeography of the cactophilic species Drosophila gouveai: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil. **Journal of Biogeography**, v. 36, n. 11, p. 2136–2147, 2009.

MORGAN, T. H. Sex Limited Inheritance in Drosophila. **Science**, v. 32, n. 812, p. 120–122, 22 jul. 1910.

MORRONE, J. J.; CRISCI, J. V. HISTORICAL BIOGEOGRAPHY: Introduction to Methods. **Annual Review of Ecology and Systematics**, v. 26, n. 1, p. 373–401, 1 nov. 1995.

O'DEA, A. et al. Formation of the Isthmus of Panama. **Science Advances**, v. 2, n. 8, p. e1600883, 1 ago. 2016.

PEARCE, C. M. et al. Parquin and carboxyparquin, toxic kaurene glycosides from the shrub Cestrum parqui. **Journal of the Chemical Society, Perkin Transactions 1**, n. 5, p. 593–600, 1 jan. 1992.

ROBE, L. J. et al. Molecular phylogeny of the subgenus Drosophila (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: a nuclear versus mitochondrial gene approach. **Molecular Phylogenetics and Evolution**, v. 36, n. 3, p. 623–640, set. 2005.

ROBE, L. J. et al. The Drosophila flavopilosa species group (Diptera, Drosophilidae): an array of exciting questions. **Fly**, v. 7, n. 2, p. 59–69, 21 abr. 2013.

ROBE, L. J.; LORETO, E. L. S.; VALENTE, V. L. S. Radiation of the „Drosophila" subgenus (Drosophilidae, Diptera) in the Neotropics. **Journal of Zoological Systematics and Evolutionary Research**, v. 48, n. 4, p. 310–321, 1 nov. 2010.

RULL, V. Speciation timing and neotropical biodiversity: the Tertiary-Quaternary debate in the light of molecular phylogenetic evidence. **Molecular Ecology**, v. 17, n. 11, p. 2722–2729, jun. 2008.

RULL, V. Neotropical biodiversity: timing and potential drivers. Trends in Ecology and Evolution, v. 26, n. 10, p. 508-513, out. 2011.

SANMARTÍN, I. Historical Biogeography: Evolution in Time and Space. **Evolution: Education and Outreach**, v. 5, n. 4, p. 555–568, 1 dez. 2012.

SANTOS, R. DE C. O. DOS; VILELA, C. R. Breeding sites of Neotropical Drosophilidae (Diptera): IV. living and fallen flowers of Sessea brasiliensis and Cestrum spp. (Solanaceae). **Revista Brasileira de Entomologia**, v. 49, n. 4, p. 544–551, dez. 2005.

SCHMITZ, H. J.; VALENTE, V. L. DA S. The flower flies and the unknown diversity of Drosophilidae (Diptera): a biodiversity inventory in the Brazilian fauna. **Papéis Avulsos de Zoologia**, v. 59, 2019.

SEPEL, L. M. N. et al. Seasonal fluctuations of D. cestri and D. incompta, two species of the flavopilosa group. **Drosophila Information Service**, v. 83, p. 122–126, 2000.

SILLERO, N. et al. Niche evolution and thermal adaptation in the temperate species Drosophila americana. **Journal of Evolutionary Biology**, v. 27, n. 8, p. 1549–1561, ago. 2014.

SINGH, B. P. Evidence of growth fault and forebulge in the Late Paleocene (~57.9–54.7 Ma), western Himalayan foreland basin, India. **Earth and Planetary Science Letters**, v. 216, n. 4, p. 717–724, 10 dez. 2003.

SOARES, E. L. C.; VIGNOLI-SILVA, M.; MENTZ, L. A. O gênero Cestrum L. (Solanaceae) no Rio Grande do Sul, Brasil. **Pesquisas, Botânica**, v. 58, p. 263–282, 2007.

STECK, A.; HUNZIKER, J. The Tertiary structural and thermal evolution of the Central Alps—compressional and extensional structures in an orogenic belt. **Tectonophysics**, Late Orogenic Extension. v. 238, n. 1, p. 229–254, 15 nov. 1994.

SUGUIO, K. Recent progress in Quaternary geology of Brazil. **Episodes**, v. 22, n. 3, p. 217–218, 1999.

SUNNUCKS, NULL. Efficient genetic markers for population biology. **Trends in Ecology & Evolution**, v. 15, n. 5, p. 199–203, maio 2000.

TEMPLETON, A. R. **Genética de Populações e Teoria Microevolutiva**. Ribeirão Preto, SP: Sociedade Brasileira de Genética, 2011.

THROCKMORTON, L. H. The Phylogeny, Ecology, and Geography of Drosophila. In: KING, R. C. (Ed.). . **Handbook of Genetics**. New York, NY: Plenum Publishing, 1975. v. 3p. 421–469.

TOSI, D. et al. On a new cave-dwelling species of bat-guano-breeding drosophila closely related to D. repleta Wollaston (Diptera, Drosophilidae). **Brazilian Journal of Genetics**, v. 13, n. 1, p. 19–31, 1990.

TURCHETTO-ZOLET, A. C. et al. Phylogeographical patterns shed light on evolutionary process in South America. **Molecular Ecology**, v. 22, n. 5, p. 1193–1213, mar. 2013.

VITT, L. J.; CALDWELL, J. P. Biogeography and Phylogeography. In: **Herpetology: An Introductory Biology of Amphibians and Reptiles**. 4th. ed. London, UK: Academic Press, 2013. p. 381–406.

WALKER, J. D. et al. The Geological Society of America Geologic Time Scale v. 5.0. **Geological Society of America**, 2018.

WHEELER, M. R.; TAKADA, H.; BRNCIC, D. The flavopilosa species group of Drosophila. In: **Studies in Genetics II**. [s.l.] University of Texas Publication, 1962. p. 395–413.

YASSIN, A. Phylogenetic classification of the Drosophilidae Rondani (Diptera): the role of morphology in the postgenomic era. **Systematic Entomology**, v. 38, n. 2, p. 349–364, 2013.

CAPÍTULO II – RELAÇÕES FILOGENÉTICAS NO GRUPO *FLAVOPILOSA* DE *DROSOPHILA*

Manuscrito preparado de acordo com as regras do periódico *Papéis Avulsos de Zoologia*†

Taxonomic subdivision and divergence times of Brazilian species of the *flavopilosa* group as assessed from a phylogenetic multi-locus perspective

Henrique da Rocha Moreira[1], Tuane Letícia Carvalho[2], Élgion Lúcio da Silva Loreto[2,3], Lizandra Jaqueline Robe[2] and Maríndia Deprá[1,3*]

[1]Programa de Pós-Graduação em Biologia Animal, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil;
[2]Programa de Pós-Graduação em Biodiversidade Animal, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brazil;
[3]Programa de Pós-Graduação em Genética e Biologia Molecular, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil;

*Corresponding author: Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul
Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil
E-mail: marindiadepra@gmail.com

† Normas e regras de submissão disponíveis em http://www.revistas.usp.br/paz/about/submissions

ABSTRACT

Among other anthophilous species groups of *Drosophila*, the *flavopilosa* group stands out not only because of its restricted ecology to *Cestrum* flowers, but also due to the cryptic morphology of its 17 species. Previous studies confirmed the monophyly of the group and the success of DNA barcoding on identifying a subset of its species, but several species remain yet to be evaluated in regard to these properties. Moreover, the taxonomy of the group remains widely incomplete, and only four of the six Brazilian species are currently assigned to subgroups. Here, we aim to test the phylogenetic subdivision and divergence times among five species of the *flavopilosa* group, helping to infer the phylogenetic position of *Drosophila mariaehelenae*, update the taxonomy of the complex and reevaluate the status of DNA barcoding as a valid tool for specimen identification in the lineage. Male individuals of *D. mariaehelenae* hatched from flowers of *Cestrum* sampled in Campo Grande (MS, Brazil) had their DNA individually extracted, as well as additional specimens of *D. flavopilosa* from Pelotas (RS, Brazil). The mitochondrial Cytochrome Oxidase subunit I (*COI*) gene and the nuclear alpha-methyldopa (*Amd*) and dopa decarboxylase (*Ddc*) genes were amplified for those specimens, and available sequences for these genes were downloaded from GenBank for other species of the group. Genetic intra- and interspecific distances were calculated for the *COI* gene and phylogenies were reconstructed for each gene individually and simultaneously by Bayesian Inference. A chronophylogeny was also estimated to access divergence times. Overall, the *flavopilosa* group was not recovered as a monophyletic unit only for *COI* phylogeny, which may reflect saturation or long branch attraction. Even so, the close relationships of *D. flavopilosa* with *D. cestri* and *D. cordeiroi* was strongly supported by all nuclear markers, suggesting the last two species should be formally allocated in the *flavopilosa* subgroup. Furthermore, *D. mariaehelenae* positioned as sister to *D. incompta*, supporting its presence in the *nesiota* subgroup. The divergence time estimation suggests the diversification of the *flavopilosa* group took place before the closure of the Central American Seaway, about 10 Mya.

Keywords: DNA barcode; ecological specialization; molecular phylogeny

INTRODUCTION

Drosophilidae comprises species with the most diverse ecological habits. Despite of being world widely known as the "fruit flies" due to the model organism *Drosophila melanogaster* Meigen, 1830, this taxon includes anthophilous (Pipkin et al., 1966), cactophilous (Manfrin and Sene, 2006) and mycophagous species (Courtney et al., 1990). Among anthophilous species, the *Drosophila flavopilosa* Frey, 1919 group stands out for its ecology, which is restricted to flowers of *Cestrum* L. (Solanaceae) (Brncic, 1966). This group is placed within the *virilis-repleta* radiation (Robe et al., 2010b), currently *Siphlodora* subgenus (Yassin, 2013) and encompasses morphologically cryptic species. The recognition of its species is only possible through the examination of internal reproductive traits, mainly the aedeagus morphology (Wheeler et al., 1962). Genetic approaches – such as DNA barcoding – have given satisfactory results on its species identification (Robe et al., 2013), and seem to be an alternative for the identification of individuals, especially females. Nevertheless, tests provided so far only included a subset of the species of the group, and for one of the species tests, only one specimen was included (Robe et al., 2013).

According to Bächli (2020), the 17 species of the *flavopilosa* group are further divided into the *nesiota* (*D. gentica* Wheeler & Takada, 1962; *D. incompta* Wheeler & Takada, 1962; *D. mariaehelenae* Vilela, 1984; and *D. nesiota* Wheeler & Takada, 1962) and *flavopilosa* (*D. acroria* Wheeler & Takada, 1962; *D. crossoptera* Wheeler & Takada, 1962; *D. flavopilosa* Frey, 1919; *D. hollisae* Vilela & Pereira, 1992; and *D. lauta* Wheeler & Takada, 1962) subgroups. The other eight species – *D. cestri* Brncic, 1978; *D. cordeiroi* Brncic, 1978; *D. korefae* Vela & Rafael, 2004; *D. melina* Wheeler, 1962; *D. pseudokorefae* Guillín & Rafael, 2018; *D. sisa* Vela & Rafael, 2005; *D. suni* Vela & Rafael, 2005; *D. taxohuaycu* Vela & Rafael, 2005 – are not yet taxonomically positioned. Among these, *D. cestri, D. cordeiroi, D. flavopilosa, D. holisae, D. incompta*, and *D. mariaehelenae* occur in Brazilian territories (Gottschalk et al., 2008), especially in the south and southeastern regions of the country (Robe et al., 2013). Although Robe et al. (2013) suggested the inclusion of *D. cestri* and *D. cordeiroi* into the *flavopilosa* subgroup, given the close relationships of both with *D. flavopilosa* (Robe et al., 2010b), this modification was not yet formally recognized. Moreover, the phylogenetic position of *D. hollisae* and *D. mariaehelenae* were not tested with molecular approaches.

Thus, the aim of this work is to provide an update on the phylogenetic subdivision and taxonomy of the *flavopilosa* group of *Drosophila*, (1) re-evaluating the status of DNA barcoding as a valid tool for specimen identification in the group; (2) describing the

phylogenetic position of *D. mariaehelenae*; (3) testing the monophyly of the group and its subgroups; and (4) estimating the divergence times within this species complex.

MATERIAL AND METHODS

Area of study and specimens sampling

Flowers of *Cestrum* were collected in Campo Grande (State of Mato Grosso do Sul, Brazil) (20°30'11.6"S, 54°36'54.1"W) and Pelotas (State of Rio Grande do Sul, Brazil) (31°45'55.44"S, 52°20'15.32"W). The flowers were taken to the laboratory, placed in 500mL vials containing vermiculite and kept in a chamber with constant temperature (21±1°C). After hatching, adult flies were captured with an entomological aspirator (Machado et al., 2014). Males were identified to species level based on genitalia patterns, according to pictures provided by Wheeler et al. (1962).

DNA isolation and sequencing

Total genomic DNA was extracted individually for each male specimen following a phenol-chloroform protocol (Jowett, 1986). Fragments of the mitochondrial gene cytochrome c oxidase subunit I (*COI*) and the nuclear genes alpha methyldopa (*Amd*) and dopa decarboxylase (*Ddc*) were amplified using the primer pairs LCO1490 and HCO2198 (Folmer et al., 1994), AmdEx4F and Amd-bw (Robe et al., 2010b; Tatarenkov et al., 2001) and BPF and BPR (Tatarenkov et al., 1999), respectively. PCR reactions were carried out using GoTaq® Hot Start Green Master Mix (Promega, Madison WI, USA), according to the manufacturer protocol, with 0.2 μM of each primer and 100-200 ng of DNA.

Amplification of the *COI* gene was performed by means of an initial step of denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 45 s, annealing at 57°C for 45 s and extension at 72°C for 1 min, and by a final extension cycle at 72°C for 5 min. Cycling conditions for *Amd* were based on two steps: 5 cycles of denaturation at 95°C for 1 min, annealing at 58°C (-1°C per cycle) for 45 s and extension at 72°C for 1 min, followed by 30 cycles of denaturation at 95°C for 1 min, annealing at 53°C for 45 s and extension at 72°C for 1 min 15 s. Amplification of the *Ddc* gene was performed with the same PCR cycles used for *Amd* gene, adjusting the first and second annealing temperatures to 60°C and 55°C, respectively.

To check whether the amplification was successful, 5μl of the 20μl PCR product was applied to 0.8% agarose gel, stained with GelRed® (Biotium). The obtained amplicons were

purified with Exonuclease I (10U/µl) and FastAP Thermosensitive Alkaline Phosphatase (1U/µl) (Thermo Scientific Inc) and directly sequenced. Sequencing was performed by Macrogen Inc. (Seoul, South Korea) using BigDye technology and the same PCR primers.

Sequences curation

The electropherograms were assembled and inspected in the Gap4 Software of the Staden Package (Staden, 1996). The consensus sequences obtained were then checked for identity in the BLASTn tool (NCBI website) and aligned using the MEGA 7.0 software (Kumar et al., 2016). Each polymorphic site detected in the alignments was checked and corrected, if necessary. Additionally, all sequences deposited on GenBank of *COI*, *Amd* and *Ddc* genes belonging to species of the *flavopilosa* group were downloaded, along with sequences of *Drosophila gymnobasis* Hardy & Kaneshiro, 1971, *Drosophila hydei* Sturtevant, 1921, *Drosophila melanogaster* Meigen, 1830, *Drosophila virilis* Sturtevant, 1916 and *Scaptomyza adusta* (Loew, 1862), which were used as outgroup. Sequences generated in this work will be further submitted to the GenBank.

DNA barcoding and phylogenetic analyses

The intra- and interspecific genetic distances were calculated for the *COI* gene in the PAUP* 4 software (Cummings, 2014), under the Kimura 2-parameter (K2P) nucleotide substitution model. The phylogenetic relationships were accessed through the reconstruction of Bayesian Inference (BI) phylogenetic trees in MrBayes 3.2 (Ronquist and Huelsenbeck, 2003). These analyses were performed individually for each marker, and simultaneously through a multi-locus approach. For the multi-locus dataset, species with more than one sequence of each gene had their respective consensus sequence calculated with Unipro UGENE software (Okonechnikov et al., 2012). The best nucleotide substitution model was chosen separately for each gene based on the Akaike information criterion (AIC), estimated with MrModeltest2 (Nylander, 2004) implemented on PAUP* 4 (Cummings, 2014). The MCMC chains were run for a minimum of 20,000,000, 55,000,000 and 100,000,000 generations for the *COI* gene, the nuclear genes and the multilocus dataset, respectively, with sampling every 10,000 generations and a 25% burn-in. Besides, for COI, a median-joining haplotype network was reconstructed with the Network 5 software (Bandelt et al., 1999) to better assess the relationships among the haplotypes of *D. incompta* and *D. mariaehelenae*.

Divergence times were estimated with BEAST 1.10 (Suchard et al., 2018). The MCMC chains were run for 600,000,000 generations, sampling every 10,000 generations. The analysis was performed using the multi-locus dataset under a lognormal relaxed-clock model, using a random start tree that was refined under a Yule speciation model. The convergence was analyzed in Tracer (Rambaut et al., 2018) and the maximum credibility tree was obtained in TreeAnotator (Suchard et al., 2018) with 25% burn-in. The calibration points were (1) the split between the *Scaptomyza* and the Hawaiian *Drosophila* lineages (17.5 Ma) (Grimaldi, 1987; Iturralde-Vinent and MacPhee, 1996); and (2) the split between the *Drosophila* and *Sophophora* subgenera (40 Ma) (Russo et al., 1995).

RESULTS

A total of four male individuals of *Drosophila mariaehelenae* emerged from *Cestrum* flowers collected in Campo Grande (MS). In addition, a total of 25, eight and four sequences of *COI*, *Amd* and *Ddc*, respectively, were downloaded from Genbank (see accession numbers on Table 1). The final datasets of *COI*, *Amd* and *Ddc*, and the multi-locus dataset, had a total of 673, 839, 1,105 and 2,310-bp length after curation, respectively. These matrices encompass sequences from 29, 10, 5 and 5 individuals of the ingroup, respectively. The best nucleotide substitution models were GTR+I for *COI* and GTR+G for both *Amd* and *Ddc* (see the complete results on Tables S1, S2 and S3).

Genetic distances for *COI* resulted in lower intra- than interspecific values for all species, and maximum intraspecific values never exceeded the minimum interspecific ones (Table 2). Regarding the phylogenetic reconstructions, although *COI* supported (PP = 1.00) the reciprocal monophyly of all the species tested, it did not recover) neither the *flavopilosa* group nor the *nesiota* subgroup as monophyletic lineages (Fig. 1A). The median-joining haplotype network reconstructed with *COI* sequences of *D. incompta* and *D. mariaehelenae* showed highly divergent haplotypes between the two species (Fig. 2).

Conversely, *Amd* (Fig. 1B) and *Ddc* (Fig. 1C) not only recovered the monophyly of the *flavopilosa* group (PP = 1.00, in both cases), but also supported its subdivision in two strongly supported sister clades: a clade containing *D. cestri*, *D. cordeiroi* and *D. flavopilosa* (PP = 0.96 and 0.94, respectively); and a clade with *D. incompta* and *D. mariaehelenae* (PP = 1.00, in both cases). For *Amd,* nevertheless, *D. incompta* appeared as a paraphyletic in regard to *D. mariaehelenae* (PP = 0.64) and placed *D. hydei* (*repleta* group) as outgroup (PP = 1.00) for the

clade containing *Scaptomyza*, Hawaiian *Drosophila*, *D. virilis* and the *flavopilosa* group, though poorly supported (PP = 0.65).

Although only a single individual of each species of the *flavopilosa* group was sequenced for *Ddc* (Fig. 1C), hindering inspection of reciprocal monophyly of the species, the topology recovered by this marker differed from that supported by *Amd* in presenting *D. cordeiroi* as closer to *D. flavopilosa* instead of *D. cestri* (PP = 0.54). The *Ddc* dataset also recovered the *Siphlodora* subgenus (*virilis-repleta* radiation) (PP = 0.97) and placed *D. hydei* as sister species of *D. virilis* and the *flavopilosa* group.

The multilocus (Fig. 1D) dataset basically repeated the general topology presented by *Ddc*, and added support for most of its relationships. The chronophylogeny showed that the diversification of the *flavopilosa* began around 10 Mya, whereas each of the two lineages diversified at means between 6 and 2 Mya (Fig. 3).

DISCUSSION

Until now, molecular studies within the *flavopilosa* group of *Drosophila* only included four of the 17 species of the group (*D. cestri*, *D. cordeiroi*, *D. flavopilosa* and *D. incompta*), and only two of these are formally included in each of the two subgroups usually attributed to the group (Bächli, 2020). This gap is probably an outcome of collection bias, once samplings of the *flavopilosa* group are widely concentrated in the South and Southeast regions of Brazil (Bächli, 2020). In addition, as each member of the *flavopilosa* group uses a different array of *Cestrum* species as host (Santos and Vilela, 2005), the collection efforts may be (not intentionally) biased on *Cestrum* species used only by that array of *Drosophila* species. Thus, this study is the first to analyze sequences obtained from specimens of *D. mariaehelenae* in a phylogenetic context. These specimens were collected at the state of Mato Grosso do Sul, which still had no registers of species from the *flavopilosa* group (Bächli, 2020).

As previously reported by Robe et al. (2013), the *COI* barcoding provides satisfactory results for specimen identification in the *flavopilosa* group, even though their analysis counted with a single sequence of *D. flavopilosa* for the estimations of genetic distances. Our study, besides the new sequences of *D. mariaehelenae*, included two additional sequences of *D. flavopilosa* into the dataset to assess whether this conclusion is robust to additional intra and interspecific sampling. Our results corroborate the findings of Robe et al. (2013) and reinforce *COI* as a valid marker for specimen identification within this cryptic species complex, especially for female individuals – which lack specific identification keys. In fact, all the tested

species satisfied the DNA barcoding premises of reciprocal monophyly (Hickerson et al., 2006) and positive barcoding gap values (Meyer and Paulay, 2005; Stoeckle, 2003).

Besides harboring the traditional marker for animal DNA barcoding approaches (Hebert et al., 2003), the mitochondrial DNA (mtDNA) has been widely used for inferring phylogenetic relationships between animal species with recent diversification, especially due to the faster evolutionary rates compared with the nuclear DNA (nDNA) (Avise et al., 1986). Nonetheless, this marker did not prove here to be a good choice for assessing the phylogenetic relationships of the *flavopilosa* group, which seem to have diversified about 10 Mya. In fact, this marker was not able to recover the monophyly of the target group, which was strongly supported by the other two nuclear markers. Beyond discussions about incongruences between mtDNA and nDNA evolutionary histories (Shaw, 2002), which seem to be quite common in *Drosophila* (De Ré et al., 2014, 2010; Robe et al., 2010; Zanini et al., 2018), this outcome is probably an artifact related to saturation and long branch attraction, which are often attributed to mitochondrial markers (Reyes et al., 1999; Robe et al., 2005; Rubinoff and Holland, 2005; Wilcox et al., 2004).

Although *Amd* and *Ddc* showed minor differences in the phylogenetic relationships within the *flavopilosa* group, nDNA seems to be essential to clarify the phylogeny of this group. In fact, these two markers taken individually or in combination were able to support not only the monophyly of the *flavopilosa* group, but also its subdivision in two lineages: one presenting *D. incompta* and *D. mariaehelenae* as sister species; the other clustering *D. flavopilosa* with *D. cestri* and *D. cordeiroi*. The first grouping supports the monophyly of the *nesiota* subgroup, as initially proposed by Vilela (1983) through the use of morphological data. Conversely, the later lineage confirms the positioning of *D. cestri* and *D. cordeiroi* within the *flavopilosa* subgroup, as initially proposed by Robe et al. (2010b) with the use of molecular data. Thus, the phylogeny recovered here supports the division of the group into the *flavopilosa* and *nesiota* subgroups, as proposed by Wheeler et al. (1962) based on morphological characters and assigns two unplaced species to their corresponding subgroup. More controversial points of the phylogeny, which should be further assessed through the inclusion of additional specimens and markers are the paraphyly of *D. incompta* in regard to *D. mariaehelenae*, as suggested by *Amd*, and the floating relationships between *D. cestri*, *D. cordeiroi* and *D. flavopilosa*, which are incongruent between both nuclear markers.

The estimation of divergence times indicates that the diversification of the *flavopilosa* group began around 10 Mya, within the Neogene (Tertiary) – about 15 my after the origin of its host plant, the *Cestrum* genus (Clarkson et al., 2017). All the diversifications within the

group happened before the closure of the Central American Seaway (the formation of the Isthmus of Panama, around 3 Mya) (O'Dea et al., 2016), except the split between *D. incompta* and *D. mariaehelenae* (around 2.5 Mya). This palaeogeological event thus may not be related with the speciation processes within the *flavopilosa* subgroup, even though further paleoecological investigations are still needed to better support this scenario.

CONCLUSION

Overall, our results agree with previous studies and confirm DNA barcoding as a valid technique to recognize species within the *flavopilosa* group of *Drosophila* (Robe et al., 2013) – especially for females, which lack identification keys. The subdivision of the group into the *flavopilosa* and *nesiota* subgroups (Wheeler et al., 1962), as well as the positioning of *D. cestri* and *D. cordeiroi* within the *flavopilosa* subgroup (Robe et al., 2013) are highly supported. In addition, the molecular data of *D. mariaehelenae* corroborates its taxonomic position within the *nesiota* subgroup (based on morphological characters) as sister species of *D. incompta* (Vilela, 1983). Although the diversification of the group seems to have occurred around 10 Mya, most splits occurred in the last 5 Mya, but still before the connection of North and South America through the Isthmus of Panama.

REFERENCES

Avise, J.C., Ellis, D., Clarke, B.C., Robertson, A., Jeffreys, A.J., 1986. Mitochondrial DNA and the evolutionary genetics of higher animals. Philosophical Transactions of the Royal Society of London. B, Biological Sciences 312, 325–342. https://doi.org/10.1098/rstb.1986.0011

Bächli, G., 2020. TaxoDros: The Database on Taxonomy of Drosophilidae [WWW Document]. URL https://www.taxodros.uzh.ch/ (accessed 1.7.20).

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Brncic, D., 1966. Ecological and Cytogenetic Studies of Drosophila Flavopilosa, a Neotropical Species Living in Cestrum Flowers. Evolution 20, 16–29. https://doi.org/10.1111/j.1558-5646.1966.tb03340.x

Clarkson, J.J., Dodsworth, S., Chase, M.W., 2017. Time-calibrated phylogenetic trees establish a lag between polyploidisation and diversification in Nicotiana (Solanaceae). Plant Syst Evol 303, 1001–1012. https://doi.org/10.1007/s00606-017-1416-9

Courtney, S.P., Kibota, T.T., Singleton, T.A., 1990. Ecology of Mushroom-feeding Drosophilidae, in: Begon, M., Fitter, A.H., Macfadyen, A. (Eds.), Advances in Ecological Research. Academic Press, pp. 225–274. https://doi.org/10.1016/S0065-2504(08)60056-2

Cummings, M.P., 2014. PAUP* (Phylogenetic Analysis Using Parsimony (and Other Methods)), in: Dictionary of Bioinformatics and Computational Biology. American Cancer Society. https://doi.org/10.1002/9780471650126.dob0522.pub2

De Ré, F.C., Loreto, E.L.S., Robe, L.J., 2010. Gene and species trees reveal mitochondrial and nuclear discordance in the Drosophila cardini group (Diptera: Drosophilidae). Invertebrate Biology 129, 353–367. https://doi.org/10.1111/j.1744-7410.2010.00207.x

De Ré, F.C., Wallau, G.L., Robe, L.J., Loreto, E.L.S., 2014. Characterization of the complete mitochondrial genome of flower-breeding Drosophila incompta (Diptera, Drosophilidae). Genetica 142, 525–535. https://doi.org/10.1007/s10709-014-9799-9

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Marine Biol. Biotechnol. 3, 294–299.

Gottschalk, M.S., Hofmann, P.R.P., Valente, V.L.S., 2008. Diptera, Drosophilidae: historical occurrence in Brazil. Check List 4, 485–518. https://doi.org/10.15560/4.4.485

Grimaldi, D.A., 1987. Phylogenetics and taxonomy of Zygothrica (Diptera: Drosophilidae). Bull Am Mus Nat Hist 186, 103–268.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. Biol. Sci. 270, 313–321. https://doi.org/10.1098/rspb.2002.2218

Hickerson, M.J., Meyer, C.P., Moritz, C., 2006. DNA Barcoding Will Often Fail to Discover New Animal Species over Broad Parameter Space. Syst Biol 55, 729–739. https://doi.org/10.1080/10635150600969898

Iturralde-Vinent, M.A., MacPhee, R.D.E., 1996. Age and Paleogeographical Origin of Dominican Amber. Science 273, 1850–1852. https://doi.org/10.1126/science.273.5283.1850

Jowett, T., 1986. Preparation of nucleic acids, in: Drosophila: A Practical Approach. IRL Press, pp. 275–286.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054

Machado, S., Santos, J.P.J., Robe, L.J., Loreto, E.L.S., 2014. An efficient and cheap entomological aspirator to collect mycophylic and anthophilic adult Drosophilidae flies. Dros Inf Serv 97, 169–171.

Manfrin, M.H., Sene, F.M., 2006. Cactophilic Drosophila in South America: A Model for Evolutionary Studies. Genetica 126, 57–75. https://doi.org/10.1007/s10709-005-1432-5

Meyer, C.P., Paulay, G., 2005. DNA Barcoding: Error Rates Based on Comprehensive Sampling. PLOS Biology 3, e422. https://doi.org/10.1371/journal.pbio.0030422

Nylander, J.A.A., 2004. MrModeltest v2. Uppsala University.

O'Dea, A., Lessios, H.A., Coates, A.G., Eytan, R.I., Restrepo-Moreno, S.A., Cione, A.L., Collins, L.S., Queiroz, A. de, Farris, D.W., Norris, R.D., Stallard, R.F., Woodburne, M.O., Aguilera, O., Aubry, M.-P., Berggren, W.A., Budd, A.F., Cozzuol, M.A., Coppard, S.E., Duque-Caro, H., Finnegan, S., Gasparini, G.M., Grossman, E.L., Johnson, K.G., Keigwin, L.D., Knowlton, N., Leigh, E.G., Leonard-Pingel, J.S., Marko, P.B., Pyenson, N.D., Rachello-Dolmen, P.G., Soibelzon, E., Soibelzon, L., Todd, J.A., Vermeij, G.J., Jackson, J.B.C., 2016. Formation of the Isthmus of Panama. Science Advances 2, e1600883. https://doi.org/10.1126/sciadv.1600883

Okonechnikov, K., Golosova, O., Fursov, M., 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28, 1166–1167. https://doi.org/10.1093/bioinformatics/bts091

Pipkin, S.B., Rodriguez, R.L., Leon, J., 1966. Plant Host Specificity Among Flower-Feeding Neotropical Drosophila (Diptera: Drosophilidae). The American Naturalist 100, 135–156. https://doi.org/10.1086/282407

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Syst Biol 67, 901–904. https://doi.org/10.1093/sysbio/syy032

Reyes, S.G., Cooper, S.J.B., Schwarz, M.P., 1999. Species Phylogeny of the Bee Genus Exoneurella Michener (Hymenoptera: Apidae: Allodapini): Evidence from Molecular and Morphological Data Sets. Ann Entomol Soc Am 92, 20–29. https://doi.org/10.1093/aesa/92.1.20

Robe, L. J., Cordeiro, J., Loreto, E.L.S., Valente, V.L.S., 2010a. Taxonomic boundaries, phylogenetic relationships and biogeography of the Drosophila willistoni subgroup (Diptera: Drosophilidae). Genetica 138, 601–617. https://doi.org/10.1007/s10709-009-9432-5

Robe, L.J., De Ré, F.C., Ludwig, A., Loreto, E.L.S., 2013. The Drosophila flavopilosa species group (Diptera, Drosophilidae): an array of exciting questions. Fly 7, 59–69. https://doi.org/10.4161/fly.23923

Robe, Lizandra J., Loreto, E.L.S., Valente, V.L.S., 2010b. Radiation of the „Drosophila" subgenus (Drosophilidae, Diptera) in the Neotropics. Journal of Zoological Systematics and Evolutionary Research 48, 310–321. https://doi.org/10.1111/j.1439-0469.2009.00563.x

Robe, L.J., Valente, V.L.S., Budnik, M., Loreto, E.L.S., 2005. Molecular phylogeny of the subgenus Drosophila (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: a nuclear versus mitochondrial gene approach. Mol. Phylogenet. Evol. 36, 623–640. https://doi.org/10.1016/j.ympev.2005.05.005

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Rubinoff, D., Holland, B.S., 2005. Between Two Extremes: Mitochondrial DNA is neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference. Syst Biol 54, 952–961. https://doi.org/10.1080/10635150500234674

Russo, C.A., Takezaki, N., Nei, M., 1995. Molecular phylogeny and divergence times of drosophilid species. Mol Biol Evol 12, 391–404. https://doi.org/10.1093/oxfordjournals.molbev.a040214

Santos, R. de C.O. dos, Vilela, C.R., 2005. Breeding sites of Neotropical Drosophilidae (Diptera): IV. living and fallen flowers of Sessea brasiliensis and Cestrum spp. (Solanaceae). Revista Brasileira de Entomologia 49, 544–551. https://doi.org/10.1590/S0085-56262005000400015

Shaw, K.L., 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. PNAS 99, 16122–16127. https://doi.org/10.1073/pnas.242585899

Staden, R., 1996. The staden sequence analysis package. Mol Biotechnol 5, 233. https://doi.org/10.1007/BF02900361

Stoeckle, M., 2003. Taxonomy, DNA, and the Bar Code of Life. BioScience 53, 796–797. https://doi.org/10.1641/0006-3568(2003)053[0796:TDATBC]2.0.CO;2

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol 4. https://doi.org/10.1093/ve/vey016

Tatarenkov, A., Kwiatowski, J., Skarecky, D., Barrio, E., Ayala, F.J., 1999. On the Evolution of Dopa decarboxylase (Ddc) and Drosophila Systematics. J Mol Evol 48, 445–462. https://doi.org/10.1007/PL00006489

Tatarenkov, A., Žurovcová, M., Ayala, F.J., 2001. Ddc and amd Sequences Resolve Phylogenetic Relationships of Drosophila. Molecular Phylogenetics and Evolution 20, 321–325. https://doi.org/10.1006/mpev.2001.0967

Vilela, C.R., 1983. Occurrence of the Drosophila flavopilosa species group (Diptera, Drosophilidae) in the state of São Paulo (Brazil) with description of one new species. Revista Brasileira de Zoologia 2, 63–69. https://doi.org/10.1590/S0101-81751983000200004

Wheeler, M.R., Takada, H., Brncic, D., 1962. The flavopilosa species group of Drosophila, in: Studies in Genetics II. University of Texas Publication, pp. 395–413.

Wilcox, T.P., García de León, F.J., Hendrickson, D.A., Hillis, D.M., 2004. Convergence among cave catfishes: long-branch attraction and a Bayesian relative rates test. Molecular Phylogenetics and Evolution 31, 1101–1113. https://doi.org/10.1016/j.ympev.2003.11.006

Yassin, A., 2013. Phylogenetic classification of the Drosophilidae Rondani (Diptera): the role of morphology in the postgenomic era. Systematic Entomology 38, 349–364. https://doi.org/10.1111/j.1365-3113.2012.00665.x

Zanini, R., Müller, M.J., Vieira, G.C., Valiati, V.H., Deprá, M., Valente, V.L. da S., 2018. Combining morphology and molecular data to improve Drosophila paulistorum (Diptera, Drosophilidae) taxonomic status. Fly 12, 81–94. https://doi.org/10.1080/19336934.2018.1429859

FIGURES

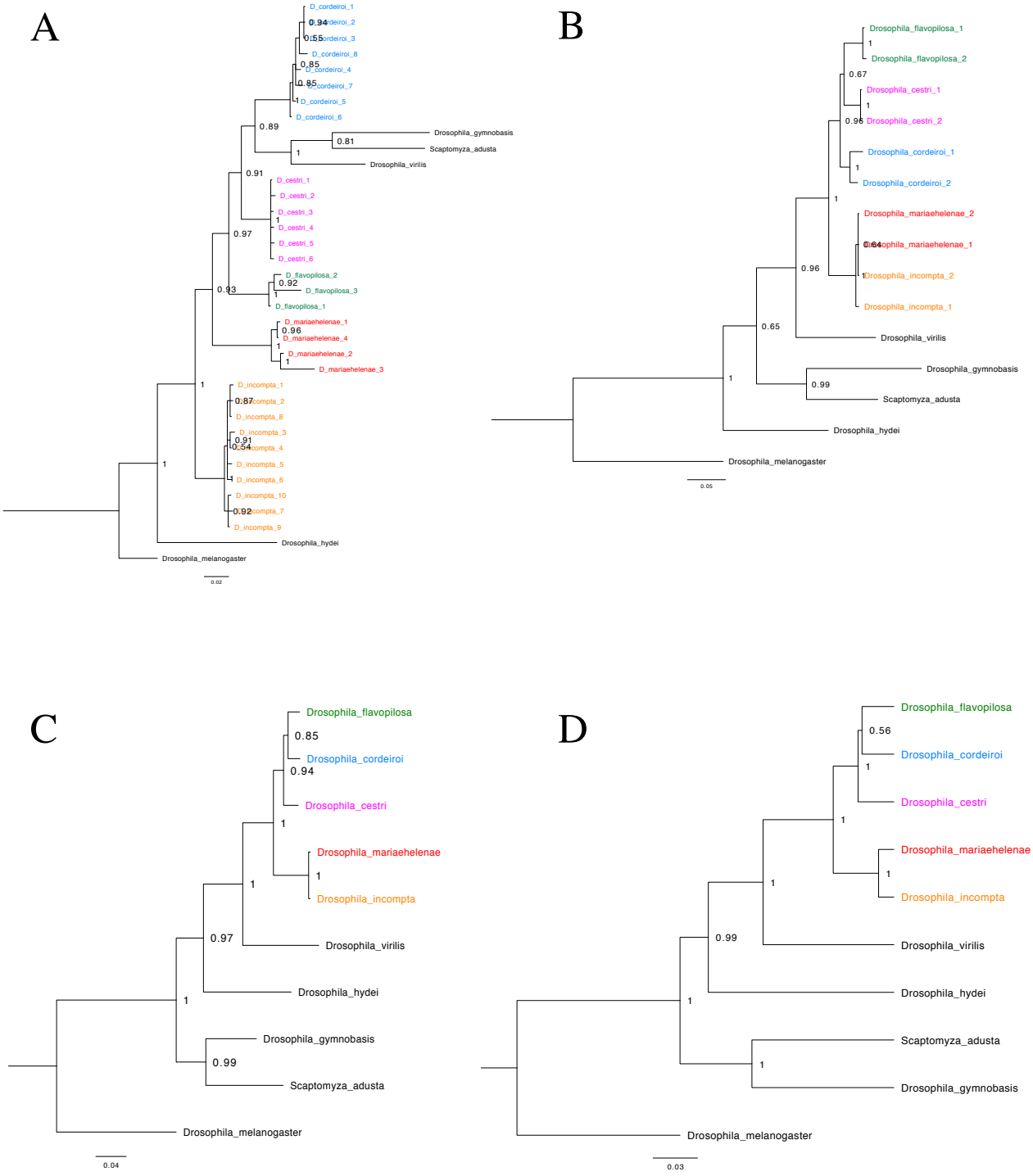


Figure 1. Majority-rule consensus trees recovered for the *flavopilosa* group of *Drosophila* through BI performed in MrBayes based on (A) the mitochondrial gene cytochrome c oxidase subunit I (*COI*); the nuclear genes (B) alpha-metyldopa (*Amd*) and (C) dopa decarboxylase (*Ddc*); and (D) a multi-locus (*COI* + *Amd* + *Ddc*) dataset. Numbers above internal nodes represent the posterior probability of each clade. Branch lengths are proportional to the scale, given in substitutions per site.

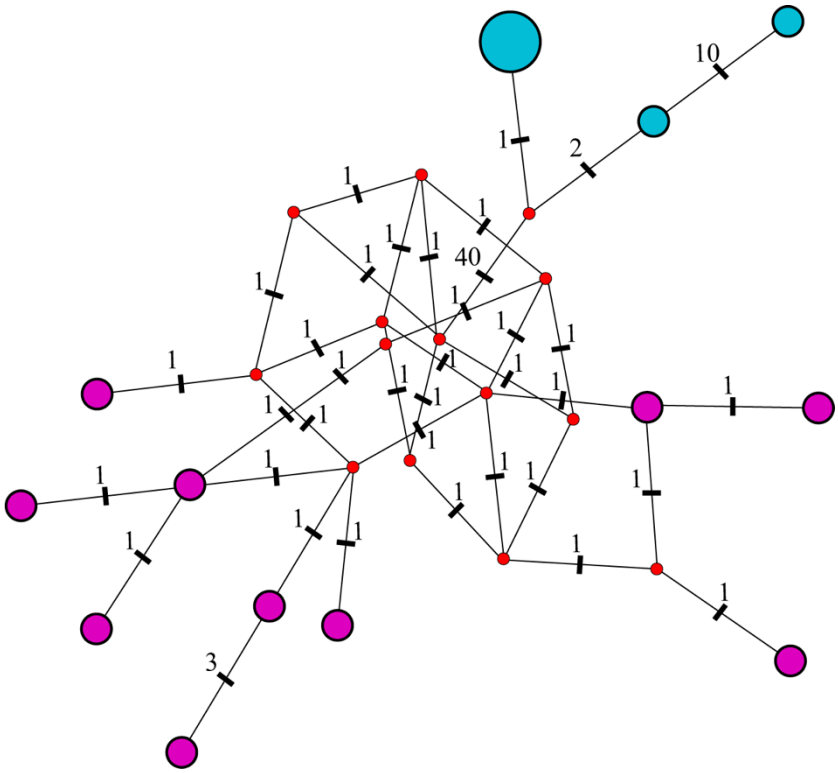


Figure 2. Median-joining network for the *COI* haplotypes of *D. incompta* (pink circles) and *D. mariaehelenae* (blue circles) performed in the software Network 5. The number of mutational steps between two haplotypes are represented above each perpendicular lines. Red circles correspond to median vectors. The haplotype size corresponds to its frequency.

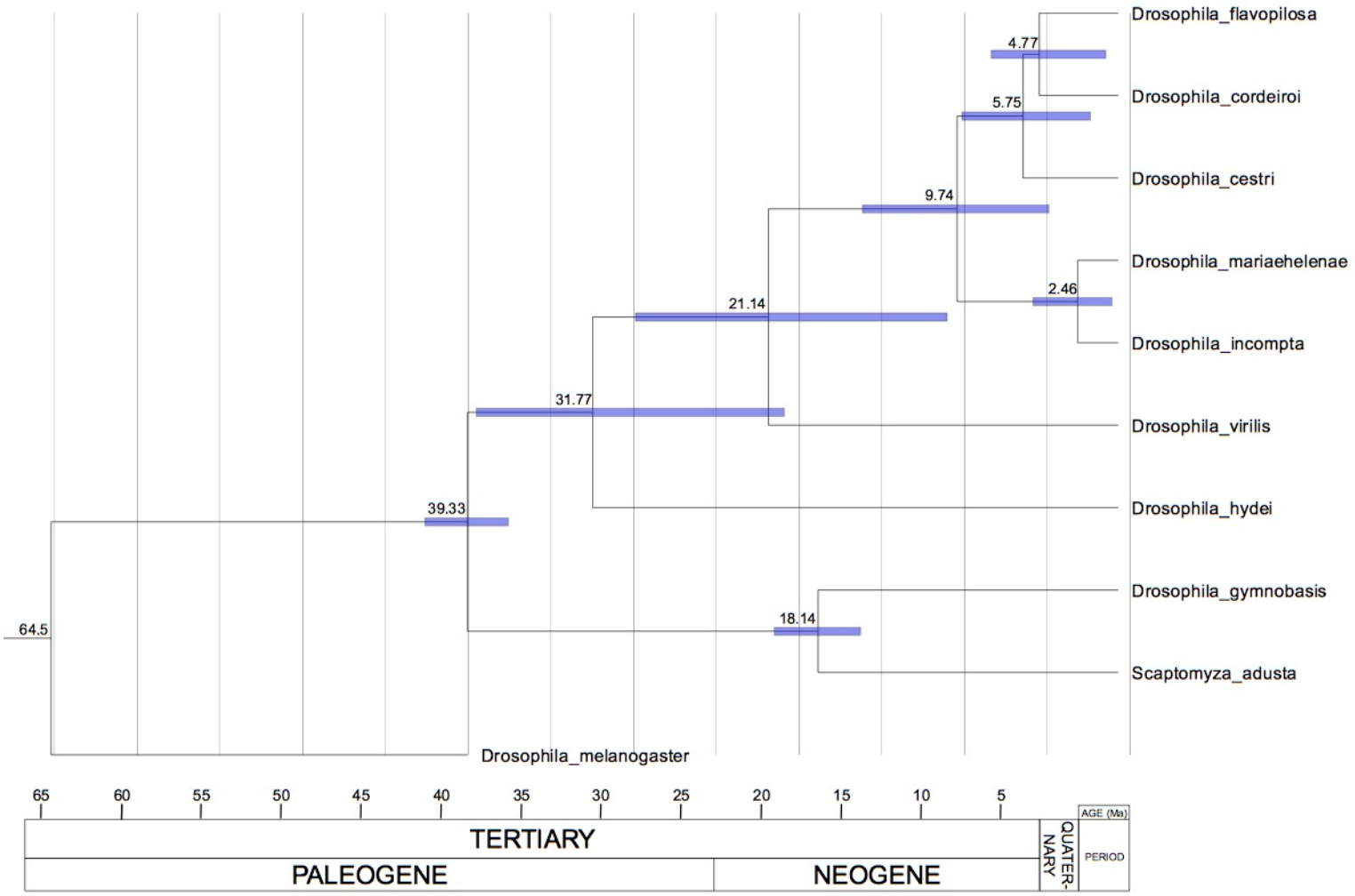


Figure 3. Chronophylogeny recovered for the *flavopilosa* group of *Drosophila* through BI performed in Beast based on a multi-locus (*COI* + *Amd* + *Ddc*) dataset. Numbers above internal nodes represent the mean age of each split, whose 95% confidence limits are reflected by the respective node bars.

TABLES

Table 1. Sequences downloaded from GenBank and their respective sources, geographical coordinates and accession numbers.

Species	Source	Latitude	Longitude	*COI*	*Amd*	*Ddc*
D. cestri	Canguçu (RS)	31°23'47.32"S	52°40'43.63"W	JX993112.1	--	--
	Pelotas (RS)	31°45'55.44"S	52°20'15.32"W	JX993111.1	--	--
	Saldanha Marinho (RS)	28°23'40.45"S	53°5'51.09"W	JX993108.1	--	--
				JX993109.1	--	--
	Santa Maria (RS)	29°41'14.3"S	53°48'55.72"W	JX993088.1	EU444560.1	EU446065.1
	São Sepé (RS)	30°9'50.36"S	53°34'18.56"W	JX993110.1	--	--
	--	--	--	--	AY699246.1	--
D. cordeiroi	Canguçu (RS)	31°23'47.32"S	52°40'43.63"W	JX993094.1	--	--
				JX993095.1	--	--
	Curitiba (PR)	25°25'44.23"S	49°16'1.69"W	JX993092.1	--	--
	Pelotas (RS)	31°45'55.44"S	52°20'15.32"W	JX993093.1	--	--
	Santa Maria (RS)	29°41'14.3"S	53°48'55.72"W	JX993089.1	EU444561.1	EU446066.1
				JX993090.1	--	--
				JX993091.1	--	--
	Viamão (RS)	30°4'19.98"S	51°5'49.05"W	JX993096.1	--	--
	--	--	--	--	EU068743.1	--
D. flavopilosa	Santa Maria (RS)	29°41'14.3"S	53°48'55.72"W	--	EU444564.1	EU446069.1
	Pelotas (RS)	31°45'55.44"S	52°20'15.32"W	JX993097.1	--	--
				*	--	--
				*	--	--
	--	--	--	--	EU068742.1	--
D. incompta	Canguçu (RS)	31°23'47.32"S	52°40'43.63"W	JX993107.1	--	--
	Cruz Alta (RS)	28°38'43.76"S	53°36'19.28"W	JX993103.1	--	--
	Curitiba (PR)	25°25'44.23"S	49°16'1.69"W	JX993104.1	--	--
	Pelotas (RS)	31°45'55.44"S	52°20'15.32"W	JX993106.1	--	--
	Saldanha Marinho (RS)	28°23'40.45"S	53°5'51.09"W	JX993101.1	--	--
				JX993102.1	--	--
	Santa Maria (RS)	29°41'14.3"S	53°48'55.72"W	JX993099.1	EU444569.1	EU446075.1
				JX993100.1	--	--
	São João do Polesine (RS)	29°37'17.02"S	53°26'57.82"W	JX993098.1	--	--
	São Sepé (RS)	30°9'50.36"S	53°34'18.56"W	JX993105.1	--	--
	--	--	--	--	AY699247.1	--
D. mariaehelenae	Campo Grande (MS)	20°28'10.96"S	54°37'12.44"W	*	*	*
				*	*	--
				*	--	--
				*	--	--

Amd, alpha metyldopa; *COI*, cytochrome oxidase c subunit I; *Ddc*, DOPA decarboxylase. Asterisks indicate new sequences added by the present study.

Table 2. Genetic K2P distances presented by the species belonging to the *D. flavopilosa* group for *COI* sequences.

	Genetic distances range (min - max)		
Species	**Intraspecific**	**Interspecific**	**Barcoding gap**
D. cestri	0.002 – 0.005	0.05 – 0.11	4.5%
D. cordeiroi	0.00 – 0.011	0.05 – 0.119	3.9%
D. flavopilosa	0.00 – 0.011	0.053 – 0.118	4.2%
D. incompta	0.002 – 0.011	0.074 – 0.11	6.3%
D. mariaehelenae	0.00 – 0.03	0.076 – 0.119	4.6%

SUPPLEMENTARY MATERIAL

Table S 1. Results of the best nucleotide substitution model based on AIC for the *COI* dataset, estimated with MrModeltest2 and PAUP* 4.

Model	-lnL	K	AIC	delta	Weight	CumWeight
GTR+I	23.438.904	9	47.057.808	0.0000	0.3785	0.3785
GTR+G	23.440.708	9	47.061.416	0.3608	0.3160	0.6945
GTR+I+G	23.431.045	10	47.062.090	0.4282	0.3055	10.000
HKY+G	23.748.574	5	47.597.148	539.341	7.35e-13	10.000
HKY+I+G	23.740.571	6	47.601.143	543.335	6.02e-13	10.000
HKY+I	23.757.759	5	47.615.518	557.710	2.93e-13	10.000
SYM+G	24.000.952	6	48.121.904	1.064.097	2.96e-24	10.000
SYM+I+G	23.996.145	7	48.132.290	1.074.482	1.76e-24	10.000
SYM+I	24.014.045	6	48.148.091	1.090.283	8.00e-25	10.000
F81+G	24.064.050	4	48.208.101	1.150.293	3.98e-26	10.000
GTR	24.026.846	8	48.213.691	1.155.884	3.01e-26	10.000
F81+I+G	24.057.734	5	48.215.469	1.157.661	2.75e-26	10.000
F81+I	24.072.947	4	48.225.894	1.168.086	1.63e-26	10.000
SYM	24.496.782	5	49.093.564	2.035.757	2.80e-45	10.000
HKY	24.510.518	4	49.101.035	2.043.228	1.40e-45	10.000
K80+G	24.695.710	2	49.431.421	2.373.613	0.00e+00	10.000
K80+I+G	24.689.512	3	49.439.023	2.381.216	0.00e+00	10.000
K80+I	24.706.135	2	49.452.271	2.394.463	0.00e+00	10.000
F81	24.738.850	3	49.537.700	2.479.893	0.00e+00	10.000
JC+G	24.862.356	1	49.744.712	2.686.904	0.00e+00	10.000
JC+I+G	24.856.360	2	49.752.720	2.694.912	0.00e+00	10.000
JC+I	24.871.760	1	49.763.521	2.705.713	0.00e+00	10.000
K80	25.287.415	1	50.594.829	3.537.021	0.00e+00	10.000
JC	25.446.023	0	50.892.046	3.834.238	0.00e+00	10.000

-lnL: negative log likelihood

K: number of estimated (free) parameters

AIC: Akaike Information Criterion

delta: Akaike difference

weight: Akaike weight

cumWeight: cumulative Akaike weight

Table S 2. Results of the best nucleotide substitution model based on AIC for the *Amd* dataset, estimated with MrModeltest2 and PAUP* 4.

Model	-lnL	K	AIC	delta	Weight	CumWeight
GTR+G	21.539.868	9	43.259.736	0.0000	0.4239	0.4239
GTR+I	21.544.712	9	43.269.424	0.9688	0.2611	0.6850
GTR+I+G	21.538.064	10	43.276.128	16.392	0.1868	0.8717
HKY+G	21.598.640	5	43.297.280	37.544	0.0649	0.9366
HKY+I	21.605.691	5	43.311.382	51.646	0.0320	0.9687
HKY+I+G	21.595.940	6	43.311.880	52.144	0.0313	0.9999
K80+G	21.703.806	2	43.447.612	187.876	3.53e-05	0.9999
K80+I+G	21.700.603	3	43.461.206	201.470	1.79e-05	10.000
K80+I	21.712.234	2	43.464.468	204.731	1.52e-05	10.000
SYM+G	21.674.331	6	43.468.662	208.926	1.23e-05	10.000
SYM+I	21.681.301	6	43.482.603	222.866	6.13e-06	10.000
SYM+I+G	21.671.726	7	43.483.452	223.716	5.88e-06	10.000
GTR	21.757.996	8	43.675.991	416.255	3.88e-10	10.000
HKY	21.806.580	4	43.693.159	433.423	1.64e-10	10.000
K80	21.909.812	1	43.839.624	579.888	1.08e-13	10.000
SYM	21.881.638	5	43.863.276	603.540	3.32e-14	10.000
F81+G	21.897.441	4	43.874.883	615.146	1.86e-14	10.000
F81+I	21.902.476	4	43.884.951	625.215	1.12e-14	10.000
F81+I+G	21.896.011	5	43.892.021	632.285	7.89e-15	10.000
JC+G	21.988.879	1	43.997.759	738.022	3.99e-17	10.000
JC+I	21.995.017	1	44.010.034	750.298	2.16e-17	10.000
JC+I+G	21.987.039	2	44.014.077	754.341	1.77e-17	10.000
F81	22.081.665	3	44.223.330	963.594	5.05e-22	10.000
JC	22.174.385	0	44.348.770	1.089.033	9.53e-25	10.000

-lnL: negative log likelihood

K: number of estimated (free) parameters

AIC: Akaike Information Criterion

delta: Akaike difference

weight: Akaike weight

cumWeight: cumulative Akaike weight

Table S 3. Results of the best nucleotide substitution model based on AIC for the *Ddc* dataset, estimated with MrModeltest2 and PAUP* 4.

Model	-lnL	K	AIC	delta	Weight	CumWeight
GTR+G	25.742.202	9	51.664.404	0.0000	0.2316	0.2316
HKY+G	25.782.749	5	51.665.498	0.1094	0.2193	0.4508
HKY+I	25.784.956	5	51.669.912	0.5508	0.1758	0.6267
GTR+I	25.747.405	9	51.674.810	10.405	0.1376	0.7643
GTR+I+G	25.737.815	10	51.675.630	11.226	0.1321	0.8964
HKY+I+G	25.780.293	6	51.680.586	16.182	0.1031	0.9996
SYM+G	25.845.806	6	51.811.611	147.207	0.0001	0.9997
SYM+I+G	25.841.116	7	51.822.231	157.827	8.66e-05	0.9998
SYM+I	25.851.152	6	51.822.305	157.900	8.63e-05	0.9999
K80+G	25.896.243	2	51.832.485	168.081	5.19e-05	0.9999
K80+I	25.898.921	2	51.837.842	173.438	3.97e-05	10.000
K80+I+G	25.893.193	3	51.846.387	181.982	2.59e-05	10.000
GTR	25.961.331	8	52.082.661	418.257	1.92e-10	10.000
HKY	26.012.627	4	52.105.254	440.850	6.19e-11	10.000
SYM	26.065.574	5	52.231.147	566.743	1.14e-13	10.000
K80	26.125.786	1	52.271.572	607.168	1.51e-14	10.000
F81+G	26.310.825	4	52.701.650	1.037.246	6.94e-24	10.000
F81+I	26.311.729	4	52.703.457	1.039.053	6.34e-24	10.000
F81+I+G	26.310.479	5	52.720.957	1.056.553	2.64e-24	10.000
JC+G	26.413.242	1	52.846.484	1.182.080	4.97e-27	10.000
JC+I	26.414.617	1	52.849.233	1.184.829	4.33e-27	10.000
JC+I+G	26.412.507	2	52.865.015	1.200.610	1.97e-27	10.000
F81	26.509.148	3	53.078.296	1.413.892	4.60e-32	10.000
JC	26.610.027	0	53.220.054	1.555.649	3.84e-35	10.000

-lnL: negative log likelihood

K: number of estimated (free) parameters

AIC: Akaike Information Criterion

delta: Akaike difference

weight: Akaike weight

cumWeight: cumulative Akaike weight

CAPÍTULO III – RELAÇÕES FILOGEOGRÁFICAS NO GRUPO *FLAVOPILOSA* DE *DROSOPHILA*

Manuscrito preparado de acordo com as regras do periódico *Zoological Journal of the Linnean Society*‡

The role of ecological specialization on population structure over time, as inferred from phylogeographical patterns evaluated for two species of the *Drosophila flavopilosa* group (Diptera, Drosophilidae)

Henrique da Rocha Moreira[1], Francine Cenzi De Ré[2], Élgion Lúcio da Silva Loreto[2,3], Lizandra Jaqueline Robe[2] and Maríndia Deprá[1,3*]

[1]Programa de Pós-Graduação em Biologia Animal, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil;
[2]Programa de Pós-Graduação em Biodiversidade Animal, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brazil;
[3]Programa de Pós-Graduação em Genética e Biologia Molecular, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil;

*Corresponding author: Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul
Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil
E-mail: marindiadepra@gmail.com

‡ Normas e regras de submissão disponíveis em https://academic.oup.com/zoolinnean/pages/General_Instructions

Abstract

The *Drosophila flavopilosa* group encompasses a cryptic complex of 17 Neotropical anthophilous species ecologically restricted to flowers of *Cestrum* L. (Solanaceae) as feeding and reproduction sites. Here, we aim to access the patterns of distribution of genetic diversity encountered in populations of *D. cestri* and *D. incompta*, in order to understand the influence of ecological specialization in the population dynamics of these species, while assessing the putative strategies that enable them to survive during periods of resource scarcity. The genes Cytochrome Oxidase subunit I (*COI*) and Ebony (*E*) were characterized for individuals of both species collected in the area where their distribution overlap, that is, the Southern and South-Eastern region of Brazil and analyzed through a set of phylogeographic approaches. The results for both species and genes are congruent and revealed high haplotype or allele diversity, with low nucleotide diversity and negative results for the neutrality tests, suggesting recent population expansions. The haplotype networks and the phylogenetic trees for the concatenated genes, together with the F_{ST}, Mantel and BAPS tests, further suggest a shallow population structure. Nevertheless, although the demographic history inferred by skyline plots confirmed that both species experienced recent population expansions, the datings of these events were quite different: whereas for *D. incompta* expansion probably started about 1.5 Mya, for *D. cestri* this event started about 300 kya. Such differences suggest that the similarities encountered for both species in the patterns of distribution of genetic diversity could be consequences of their shared ecological habits, being putatively related to their survival strategies in face of the ephemerity of the resources. We further suggest that these species are probably organized into metapopulations, which are constituted by subpopulations with high levels of migration.

Keywords: *Drosophila cestri – Drosophila incompta* – phylogeography – population structure – specialist species

Introduction

In ecology, species that exploit a restricted range or few resources are commonly referred as specialists, in contrast to generalist species, which exploit a broad range of resources (Mills *et al*., 2019). It is commonly assumed that these species have different population dynamics (Kolasa & Li, 2003), as detected from differences in their patterns of genetic diversity and population structure (Carvalho, 2019). In this sense, specialist species are estimated to be more affected by past and current climate changes (Colles *et al*., 2009), which is also related to their higher extinction rates (Clavel *et al*., 2010). This poses serious flaws to the survival of specialist species in larger time ranges, and reveals the importance of assessing their survival strategies over evolutionary time.

The high ecological diversity found within Drosophilidae is not only linked with the array of feeding or reproduction substrates, but also with the different levels of specialization presented by its species. In fact, anthophilous species of *Drosophila* are a good example of how an ecological trait can differ even among closely related species (Schmitz & Valente, 2019). Even so, there are cases where an entire group of species share a restricted set of resources. This is the case for the *Drosophila flavopilosa* Frey, 1919 group, which encompasses a cryptic complex of 17 anthophilous species ecologically restricted to flowers of *Cestrum* L. (Solanaceae) as feeding and reproduction sites (Wheeler *et al*., 1962; Bächli, 2020). In this sense, the geographical distribution of the group is totally dependent on the distribution of *Cestrum* species, which are highly abundant in the Neotropics (Robe *et al*., 2013).

Even so, the geographic distribution and the array of host species usually differ between species of the *D. flavopilosa* group. In fact, it seems that each species is specialized to a unique but somewhat overlapped array of *Cestrum* species (Santos & Vilela, 2005). Moreover, whereas some species of the group are widely endemic, others are widespread along the Neotropical region (Robe *et al*., 2013). The geographical and host range overlaps allow different species to be encountered in sympatry or even syntopy (Sepel *et al*., 2000). This is the case, for example, of *Drosophila cestri* Brncic, 1978 and *Drosophila incompta* Wheeler & Takada, 1962, which co-occur in a small area of the Southern Neotropics (Robe *et al*., 2013). Nevertheless, although the former species is possibly restricted to the Southeast region of Brazil and Uruguay, the latter has a disjunct geographical distribution expanding from the South of Mexico to the North of Argentina (Bächli, 2020).

The close ecological relationship with *Cestrum* flowers entailed several morphological and physiological adaptations to species of the *D. flavopilosa* group. Among these, the ability

to explore toxic compounds of the host species (Pearce *et al*., 1992) and the short and quick embryological development due to the ephemerity of the host flowers (Brncic, 1983; Aguilar & Galetto, 2004) can be highlighted. In addition, the hosts do not seem to have a specific pattern of phenology (Santos & Vilela, 2005), which imposes the need of demographic or physiologic adaptations that are not understood. Here, we aim to access the patterns of distribution of genetic diversity encountered in Brazilian populations of *D. cestri* and *D. incompta* in order to clarify the historical and demographic processes that could explain them. While addressing the similarities and incongruences between species, we aim to understand how the *flavopilosa* group survive during the periods of resource scarcity and identify the phylogeographical patterns and processes related with ecologically specialized species in the Neotropics.

Material and methods

Area of study and specimen sampling

Flowers of *Cestrum* were collected in the Brazilian states of Minas Gerais (MG), Paraná (PR), Rio Grande do Sul (RS) and Santa Catarina (SC) (Fig. 1). The flowers were taken to the laboratory, placed in 500 mL vials containing vermiculite and kept in a chamber with constant temperature (21±1°C). After hatching, the adult flies were captured with an entomological aspirator (Machado *et al*., 2014) and immediately fixed in absolute ethanol. Males of *D. cestri* and *D. incompta* were identified to species level based on internal morphology and genitalia patterns, according to pictures provided by Wheeler *et al*. (1962). Females were not included in our analyses due to the lack of specific morphological identification keys.

DNA isolation and sequencing

Total genomic DNA was extracted individually for each specimen following a phenol-chloroform protocol (Jowett, 1986). Fragments of the mitochondrial gene Cytochrome c oxidase subunit I (*COI*) and the nuclear gene Ebony (*E*) were amplified using the primer pairs LCO1490 and HCO2198 (Folmer *et al*., 1994) and 525F (5'-CCCATSACCTCKGTGGAGCCGTA-3') and 526R 5'-CTGCATCGCATCTTYGAGGAGCA-3'), respectively. PCR reactions were carried out using GoTaq® Hot Start Green Master Mix (Promega, Madison WI, USA), according to the manufacturer protocol, with 0.2 μM of each primer, and 100-200 ng of DNA.

Amplification for the *COI* gene was performed by means of an initial step of denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 45 s, annealing at 57°C for 45 s and extension at 72°C for 1 min, and by a final extension step at 72°C for 5 min. Amplification for the *E* gene was performed by means of an initial step of denaturation at 94°C for 5 min, followed by 40 cycles of denaturation at 94°C for 45 s, annealing at 58°C for 45 s and extension at 72°C for 1 min, and by a final extension step at 72°C for 7 min.

To check whether the amplification was successful, 5 μl of the 20 μl PCR product were applied to 0.8% agarose gel, stained with GelRed® (Biotium). The obtained amplicons were purified with Exonuclease I (10U/μl) and FastAP Thermosensitive Alkaline Phosphatase (1U/μl) (Thermo Scientific Inc) and directly sequenced. Sequencing was performed by Macrogen Inc. (Seoul, South Korea) using BigDye technology and the same PCR primers.

Data analysis

The electropherograms were assembled and inspected in the Gap4 Software of the Staden Package (Staden, 1996). The consensus sequences obtained were then checked for identity in the BLASTn tool (NCBI website) and aligned with the ClustalW algorithm on MEGA 7 (Kumar *et al*., 2016). Heterozygous sites found in *E* sequences and polymorphic sites detected in the alignments were individually checked and corrected, if necessary. The alleles for the *E* gene were reconstructed with PHASE, implemented in DnaSP 6 (Rozas *et al*., 2017).

Phylogeographical analysis

General levels of genetic diversity, including the number of haplotypes (h) [or alleles (a), for *E*], haplotype diversity (Hd) [or expected heterozygosity (He), for *E*], nucleotide diversity (π) and number of polymorphic sites (s), were obtained in DnaSP 6 (Rozas *et al*., 2017). This software was also used to perform the Tajima's *D* (Tajima, 1989) and Fu's and Li's *D* and *F* (Fu, 1997) neutrality tests. The differentiation levels presented by different sampling locations were estimated by pairwise F_{ST} in Arlequin 3.5 (Excoffier & Lischer, 2010). A Mantel Test was performed for each species in the software Alleles In Space (Miller, 2005), with significance measured though 1,000 permutations. Additionally, the general structure and the most likely number of clusters was evaluated through a Bayesian analysis performed in BAPS 6.0 (Corander *et al*., 2003). Finally, evolutionary relationships between the resulting

COI haplotypes were evaluated by the reconstruction of median-joining haplotype networks on Network 5 (Bandelt *et al*., 1999).

Phylogenetic relationships were further accessed through Bayesian Inferences (BI) for the multi-locus datasets (concatenated *COI* and *E* sequences), employing sequences of *D. virilis* and *D. melanogaster* as outgroups. In each case, the best DNA substitution model was chosen for each gene individually based on the Akaike information criterion (AIC) estimated by MrModeltest2 (Nylander, 2004), implemented with the software PAUP* 4 (Cummings, 2014). The first BI was performed in MrBayes 3.2 (Ronquist & Huelsenbeck, 2003), using a partitioned scheme in which each gene has its own substitution model and parameters. This analysis was performed with two MCMC runs, each of which consisted of 200,000,000 generations, saving every 10,000 and burning the initial 25%. An additional BI was then performed in BEAST 1.8.4 (Drummond *et al*., 2012) (employing sequences of *D. albomicans*, *D. americana*, *D. grimshawi*, *D. hydei*, *D. virilis* and *Scaptomyza flava* as outgroups), in order to measure divergence and diversification times, and obtain valid estimates of evolutionary rates. This analysis was performed using unlinked substitution and molecular clock models, with a normal prior distributions of 17.5 Mya (±1.25) to the divergence between *Scaptomyza* and *D. grimshawi* (Grimaldi, 1987; Iturralde-Vinent & MacPhee, 1996). The MCMC run consisted of 100,000,000 generations, saving every 10,000. Convergence of the run and effective sample size were evaluated in Tracer 1.7 (Rambaut *et al*., 2018), and the maximum clade credibility tree was reconstructed in TreeAnnotator (Drummond *et al*., 2012), using the first 10% samplings as burn-in.

The inference of past population dynamics of each species was assessed by the reconstruction of Extended Bayesian Skyline Plots (EBSP) implemented on BEAST 2.5 (Bouckaert *et al*., 2019). These analyses were performed using unlinked substitution and molecular clock models, employing the evolutionary rates estimated through the chronophylogenetic analysis described above as normal priors. The analysis was run for 500,000,000 generations, with samplings every 10,000.

Results

Sampling and sequence analysis

A total of 119 and 183 individuals of *D. cestri* and *D. incompta* were collected, respectively, along 24 sample points across the Brazilian states of Minas Gerais (n = 1), Paraná (n = 2), Rio Grande do Sul (n = 19) and Santa Catarina (n = 2) (Fig. 1, Table 1). Regarding the

molecular markers, 626 and 623 bp of the *COI* gene were sequenced for 94 and 183 individuals, while 881 and 700 bp of the *E* gene were sequenced for 88 and 128 individuals of *D. cestri* and *D. incompta*, respectively (Table 1).

Genetic diversity and Neutrality tests

The general diversity indexes presented by each species for both genes and the multilocus datasets are presented on Table 2. Overall, both *COI* and *E* genes showed similar and high values for the haplotype diversity and expected heterozygosity, respectively, but low nucleotide diversity across both *D. cestri* and *D. incompta*. Both genes also presented negative and statistically significant ($P < 0.05$) results for the Tajima's D and Fu & Li's D and F neutrality tests.

Population structure

The F_{ST} values estimated for *D. cestri* indicated medium to high genetic differentiation (Hartl & Clark, 1997), even though the *COI* sequences comparisons (Table 3, Fig. 2A) scored higher values than the *E* gene (Table 3, Fig. 2B), especially between the populations of Santa Rosa, RS and Bagé, RS (0.7); Cachoeira do Sul, RS (0.44); Curitiba, PR (0.66); Pelotas, RS (0.6); Santa Maria, RS (0.58). Otherwise, the values for *D. incompta* showed low to medium genetic differentiation between the sampled populations, with more congruent F_{ST} scores between *COI* (Table 4, Fig. 2C) and *E* (Table 4, Fig. 2D). Higher values for the *COI* gene were scored between the populations of Florianópolis, SC and Santiago (0.5); Bagé, RS (0.34); Cruz Alta, RS (0.25); Belo Horizonte, MG (0.31).

The shallow population structure was reinforced by the *COI* haplotype networks (Fig. 3). Both *D. cestri* and *D. incompta* presented a star-like network pattern, in which a few central and highly frequent haplotypes are connected by few mutational steps to peripheral and exclusive or little shared haplotypes. Additionally, the majority rule consensus trees reconstructed through BI for the multilocus dataset of both species (Fig. 4) had essentially polytomic relationships between the sampled specimens and, therefore, did not present any clade that would correspond to specific regions or sampling points. In addition, the Mantel Test performed for both species and genes (*D. cestri*: r = -0,1703 and r = 0,0396; *D. incompta*: r = -0,0482 and r = 0,0747, for *COI* and *E* genes, respectively) scored no significant correlation between genetic and geographical distances (Fig. S1) and the BAPS placed all populations within one group (Fig. S2), thus supporting the previous results.

Demographic history

The results of the EBSP (Fig. 5) indicate population expansion for both species and corroborate the previous results of the networks and neutrality tests. Despite the general similarity in the patterns obtained, *D. cestri* presents signals of a small population expansion that occurred in the last 300 kya, while *D. incompta* displays a more substantial population expansion that began around 1.5 Mya.

Discussion

The sequences here characterized from both genomes recovered very similar patterns within and between *D. cestri* and *D. incompta*, even though mtDNA and nDNA markers usually describe different aspects on the evolutionary history of a species due to the distinct inheritance patterns and effective population sizes (Sunnucks, 2000). Despite the high values of haplotype diversity or expected heterozigosity, the low nucleotide diversity and the neutrality tests for both species indicate only small differences between haplotypes or alleles and suggest a recent population expansion. The median-joining networks also display and emphasize those results, showing a star-like-pattern that is characteristic of recent population expansions. Moreover, for both species, the Mantel tests discarded the isolation-by-distance hypothesis and the BAPS analysis placed all populations in one group. Such patterns of weak population structure seem to be common in somehow ecologically restricted insects, as seen in *Mycetophylax simplex* (Cardoso *et al*., 2015), an ant endemic to the coast of Brazil and highly specialized to the sand dunes habitat; and for Euglossini bees, a tribe that exclusively pollinates orchid species and that reach up to 2.500 km of gene flow (Dick *et al*., 2004).

Nevertheless, although both *D. cestri* and *D. incompta* present signals of population expansion, with absence of a deeper population structure, the timing of the expansion events are quite different. In fact, although both species expanded in face of the Quaternary climate changes (Rull, 2008), *D. cestri* showed only a slight population expansion, that occurred in the last 300 kya, whereas *D. incompta* has been experiencing a deep population expansion in the last 1.5 Mya. As the beginning of both expansions can be located in interglacial periods (Gibbard *et al*., 2010; Dahl-Jensen *et al*., 2013), the two species of the *D. flavopilosa* group can be added to the list of *Drosophila* species presenting evidences of demographic effects to the climatic oscillations of the Quaternary (Brehm *et al*., 2004; Franco & Manfrin, 2013; De Ré *et al*., 2014a; Gustani *et al*., 2015; Barrios-Leal *et al*., 2018). Interestingly, although many

Cestrum plants are usually encountered in transitional areas (de Rojas & D'Arcy, 1998), the patterns found for the *D. flavopilosa* group resembles those encountered for species inhabiting more humid habitats [as *D. maculifrons* (De Ré *et al*., 2014a) and *D. ornatifrons* (Gustani *et al*., 2015)] than for species inhabiting xeromorphic environments [as *D. seriema* (Franco & Manfrin, 2013) and *D. buzzatii* (Brito *et al*., 2002)].

Beyond this scenario, despite the differences in the expansion dating, the fact that both species presented similar patterns of population structure allows us some further speculations. In fact, although it could be argued that the absence of population structure could be a consequence of the recent population expansions, with insufficient time to attain significant population structure (Ceballos *et al*., 2012), the fact that the species presenting older expansion (*D. incompta*) presents lower mean F_{ST} values (0.087 versus 0.112, as encountered for *COI*; 0.031 versus 0.088, as encountered for *E*; all values with $P < 0.05$) encourages us to present similarities as an outcome of the shared ecological patterns. In this case, similar patterns of distribution of genetic diversity could arrive as a consequence of specialization or even as survival strategies of the species in face of the ephemerality of the resources.

Two hypotheses have been raised so far to explain the survival of specimens belonging to the *flavopilosa* group of *Drosophila* during the periods of scarce or none flourishing of *Cestrum* species. The first, stated by Hofmann (1985), considers the presence of apparently senile adults of *D. incompta* as a signal that a few individuals survive those critical periods and found new populations as the resource becomes available again; in this scenario, we would expect to detect signals of one or more population bottlenecks. The second hypothesis, more recently stated by De Ré *et al*. (2014b), view the high levels of polymorphism found in the mitochondrial DNA of *D. incompta* as a consequence from cyclic migration and population expansions/retractions during the search for available resources; in this scenario, we would expect a shallow population structure. Our results, therefore, are congruent with the latter hypothesis, although do not completely discard the former.

In this case, for both *D. cestri* and *D. incompta*, F_{ST} values indicated punctual cases of high genetic differentiation, especially concerning the *COI* gene. In both cases, higher values of F_{ST} were presented by peripheral populations, not necessarily the more distant ones. Such a pattern may reflect the fact that we are dealing with a major metapopulation, with outliers putatively reflecting adjacent metapopulations or contact zones. This model was firstly described by Levins, (1969) and characterizes the process of populations displacement, extinction and establishment of new populations across time (Hanski & Gilpin, 1991); or, in

Levins' own words, "*a population of populations*". The idea of founding new populations by a few individuals after a former extinction is compatible with all the results provided here, and may be further supported by data showing different phenological seasons for different *Cestrum* species (Santos & Vilela, 2005).

Acknowledgements

The authors acknowledge Dr Adriana Ludwig, MSc Ana Paula T. Costa, Dr Juliana Cordeiro and MSc Tuane Letícia Carvalho for helping with *Cestrum* and *Drosophila* samples ant to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Finance Code 001 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for providing scholarship and financial supports.

References

Aguilar R, Galetto L. 2004. Effects of forest fragmentation on male and female reproductive success in Cestrum parqui (Solanaceae). *Oecologia* 138: 513–520.

Bächli G. 2020. TaxoDros: The Database on Taxonomy of Drosophilidae.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37–48.

Barrios-Leal DY, Franco FF, Silva ECC, Santos CKB, Sene FM, Manfrin MH. 2018. Deep intraspecific divergence in Drosophila meridionalis, a cactophilic member of the New World Drosophila repleta group. *Biological Journal of the Linnean Society* 123: 163–178.

Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, Maio ND, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, Plessis L du, Popinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond AJ. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology* 15: e1006650.

Brehm A, Harris DJ, Hernández M, Perez JA, Larruga JM, Pinto FM, González AM. 2004. Phylogeography of Drosophila subobscura from North Atlantic islands inferred from mtDNA A+T rich region sequences. *Molecular Phylogenetics and Evolution* 30: 829–834.

Brito RA de, Manfrin MH, Sene FM. 2002. Mitochondrial DNA phylogeography of Brazilian populations of Drosophila buzzatii. *Genetics and Molecular Biology* 25: 161–171.

Brncic D. 1983. The flavopilosa group of species as an example of flower-breeding species. *The genetics and biology of Drosophila*. New York, NY: Academic Press, 360–377.

Cardoso DC, Cristiano MP, Tavares MG, Schubart CD, Heinze J. 2015. Phylogeography of the sand dune ant Mycetophylax simplex along the Brazilian Atlantic Forest coast: remarkably low mtDNA diversity and shallow population structure. *BMC Evolutionary Biology* 15: 1–13.

Carvalho TL. 2019. Estruturação populacional de espécies de drosofilídeos antófilas com diferentes graus de especialização. Thesis, Universidade Federal de Santa Maria.

Ceballos SG, Lessa EP, Victorio MF, Fernández DA. 2012. Phylogeography of the sub-Antarctic notothenioid fish *Eleginops maclovinus*: evidence of population expansion. *Marine Biology* 159: 499-505.

Clavel J, Julliard R, Devictor V. 2010. Worldwide decline of specialist species: toward a global functional homogenization?. *Frontiers in Ecology and the Environment* 9(4): 222-228

Colles A, Liow LH, Prinzing A. 2009. Are specialists at risk under environmental change? Neological, paleoecological and phylogenetic approaches. *Ecology Letters* 12(8): 849-863.

Corander J, Waldmann P, Sillanpää MJ. 2003. Bayesian Analysis of Genetic Differentiation Between Populations. *Genetics* 163: 367–374.

Cummings MP. 2014. PAUP* (Phylogenetic Analysis Using Parsimony (and Other Methods)). *Dictionary of Bioinformatics and Computational Biology*. American Cancer Society, .

Dahl-Jensen D, Albert MR, Aldahan A, Azuma N, Balslev-Clausen D, Baumgartner M, Berggren AM, Bigler M, Binder T, Blunier T, Bourgeois JC, Brook EJ, Buchardt SL, Buizert C, Capron E, Chappellaz J, Chung J, Clausen HB, Cvijanovic I, Davies SM, Ditlevsen P, Eicher O, Fischer H, Fisher DA, Fleet LG, Gfeller G, Gkinis V, Gogineni S, Goto-Azuma K, Grinsted A, Gudlaugsdottir H, Guillevic M, Hansen SB, Hansson M, Hirabayashi M, Hong S, Hur SD, Huybrechts P, Hvidberg CS, Iizuka Y, Jenk T, Johnsen SJ, Jones TR, Jouzel J, Karlsson NB, Kawamura K, Keegan K, Kettner E, Kipfstuhl S, Kjær HA, Koutnik M, Kuramoto T, Köhler P, Laepple T, Landais A, Langen PL, Larsen LB, Leuenberger D, Leuenberger M, Leuschen C, Li J, Lipenkov V, Martinerie P, Maselli OJ, Masson-Delmotte V, McConnell JR, Miller H, Mini O, Miyamoto A, Montagnat-Rentier M, Mulvaney R, Muscheler R, Orsi AJ, Paden J, Panton C, Pattyn F, Petit JR, Pol K, Popp T, Possnert G, Prié F, Prokopiou M, Quiquet A, Rasmussen SO, Raynaud D, Ren J, Reutenauer C, Ritz C, Röckmann T, Rosen JL, Rubino M, Rybak O, Samyn D, Sapart CJ, Schilt A, Schmidt AMZ, Schwander J, Schüpbach S, Seierstad I, Severinghaus JP, Sheldon S, Simonsen SB, Sjolte J, Solgaard AM, Sowers T, Sperlich P, Steen-Larsen HC, Steffen K, Steffensen JP, Steinhage D, Stocker TF, Stowasser C, Sturevik AS, Sturges WT, Sveinbjörnsdottir A, Svensson A, Tison JL, Uetake J, Vallelonga P, van de Wal RSW, van der Wel G, Vaughn BH, Vinther B, Waddington E, Wegner A, Weikusat I, White JWC, Wilhelms F, Winstrup M, Witrant E, Wolff EW, Xiao C, Zheng J, NEEM community members. 2013. Eemian interglacial reconstructed from a Greenland folded ice core. *Nature* 493: 489–494.

De Ré FC, Gustani EC, Oliveira APF, Machado LPB, Mateus RP, Loreto ELS, Robe LJ. 2014a. Brazilian populations of Drosophila maculifrons (Diptera: Drosophilidae): low diversity levels and signals of a population expansion after the Last Glacial Maximum. *Biological Journal of the Linnean Society* 112: 55–66.

De Ré FC, Wallau GL, Robe LJ, Loreto ELS. 2014b. Characterization of the complete mitochondrial genome of flower-breeding Drosophila incompta (Diptera, Drosophilidae). *Genetica* 142: 525–535.

Dick CW, Roubik DW, Gruber KF, Bermingham E. 2004. Long-distance gene flow and cross-Andean dispersal of lowland rainforest bees (Apidae: Euglossini) revealed by comparative mitochondrial DNA phylogeography. *Molecular Ecology* 13: 3775–3785.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3: 294–299.

Franco FF, Manfrin MH. 2013. Recent demographic history of cactophilic Drosophila species can be related to Quaternary palaeoclimatic changes in South America. *Journal of Biogeography* 40: 142–154.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915–925.

Gibbard PL, Head MJ, Walker MJC. 2010. Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. *Journal of Quaternary Science* 25: 96–102.

Grimaldi DA. 1987. Phylogenetics and taxonomy of Zygothrica (Diptera: Drosophilidae). 186: 103–268.

Gustani EC, Oliveira APF, Santos MH, Machado LPB, Mateus RP. 2015. Demographic Structure and Evolutionary History of Drosophila ornatifrons (Diptera, Drosophilidae) from Atlantic Forest of Southern Brazil. *Zoological Science* 32: 141–150.

Hanski I, Gilpin M. 1991. Metapopulation dynamics: brief history and conceptual domain. *Biological Journal of the Linnean Society* 42: 3–16.

Hartl DL, Clark GC. 1997. *Principles of Population Genetics*. Sunderland, MA: Sinauer Associates.

Hofmann PRP. 1985. Variabilidade genética em espécies de nicho ecológico restrito. *Ciência e Cultura* 37: 579–581.

Iturralde-Vinent MA, MacPhee RDE. 1996. Age and Paleogeographical Origin of Dominican Amber. *Science* 273: 1850–1852.

Jowett T. 1986. Preparation of nucleic acids. *Drosophila: a practical approach*. IRL Press, 275–286.

Kolasa J, Li BL. 2003. Removing the confounding effect of habitat specialization reveals the stabilizing contribution of diversity to species variability. Proceedings of the Royal Society of London. Series B: Biological Sciences 270: S198–S201.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* 33: 1870–1874.

Levins R. 1969. Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control. *Bulletin of the Entomological Society of America* 15: 237–240.

Machado S, Santos JPJ, Robe LJ, Loreto ELS. 2014. An efficient and cheap entomological aspirator to collect mycophylic and anthophilic adult Drosophilidae flies. *Drosophila Information Service* 97: 169–171.

Miller MP. 2005. Alleles In Space (AIS): Computer Software for the Joint Analysis of Interindividual Spatial and Genetic Information. *Journal of Heredity* 96: 722–724.

Mills CG, Allen RJ, Blythe RA. 2019. Resource spectrum engineering by specialist species can shift the specialist-generalist balance. *Theoretical Ecology* https://doi.org/10.1007/s12080-019-00436-8

Nylander JAA. 2004. *MrModeltest v2*. Uppsala University.

Pearce CM, Skelton NJ, Naylor S, Kanaan R, Kelland J, Oelrichs PB, Sanders JKM, Williams DH. 1992. Parquin and carboxyparquin, toxic kaurene glycosides from the shrub Cestrum parqui. *Journal of the Chemical Society, Perkin Transactions* 1: 593–600.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* 67: 901–904.

Robe LJ, De Ré FC, Ludwig A, Loreto ELS. 2013. The Drosophila flavopilosa species group (Diptera, Drosophilidae): an array of exciting questions. *Fly* 7: 59–69.

de Rojas CB, D'Arcy WG. 1998. The Genera Cestrum and Sessea (Solanaceae: Cestreae) in Venezuela. *Annals of the Missouri Botanical Garden* 85: 273–351.

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

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Molecular Biology and Evolution* 34: 3299–3302.

Rull V. 2008. Speciation timing and neotropical biodiversity: the Tertiary-Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology* 17: 2722–2729.

Santos R de CO dos, Vilela CR. 2005. Breeding sites of Neotropical Drosophilidae (Diptera): IV. living and fallen flowers of Sessea brasiliensis and Cestrum spp. (Solanaceae). *Revista Brasileira de Entomologia* 49: 544–551.

Schmitz HJ, Valente VL da S. 2019. The flower flies and the unknown diversity of Drosophilidae (Diptera): a biodiversity inventory in the Brazilian fauna. *Papéis Avulsos de Zoologia* 59.

Sepel LMN, Golombieski RM, Napp M, Loreto ELS. 2000. Seasonal fluctuations of D. cestri and D. incompta, two species of the flavopilosa group. *Drosophila Information Service* 83: 122–126.

Staden R. The Staden sequence analysis package. Molecular Biotechnology 5: 233.

Sunnucks null. 2000. Efficient genetic markers for population biology. *Trends in Ecology & Evolution* 15: 199–203.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.

Wheeler MR, Takada H, Brncic D. 1962. The flavopilosa species group of Drosophila. *Studies in Genetics II*. University of Texas Publication, 395–413.

Figures

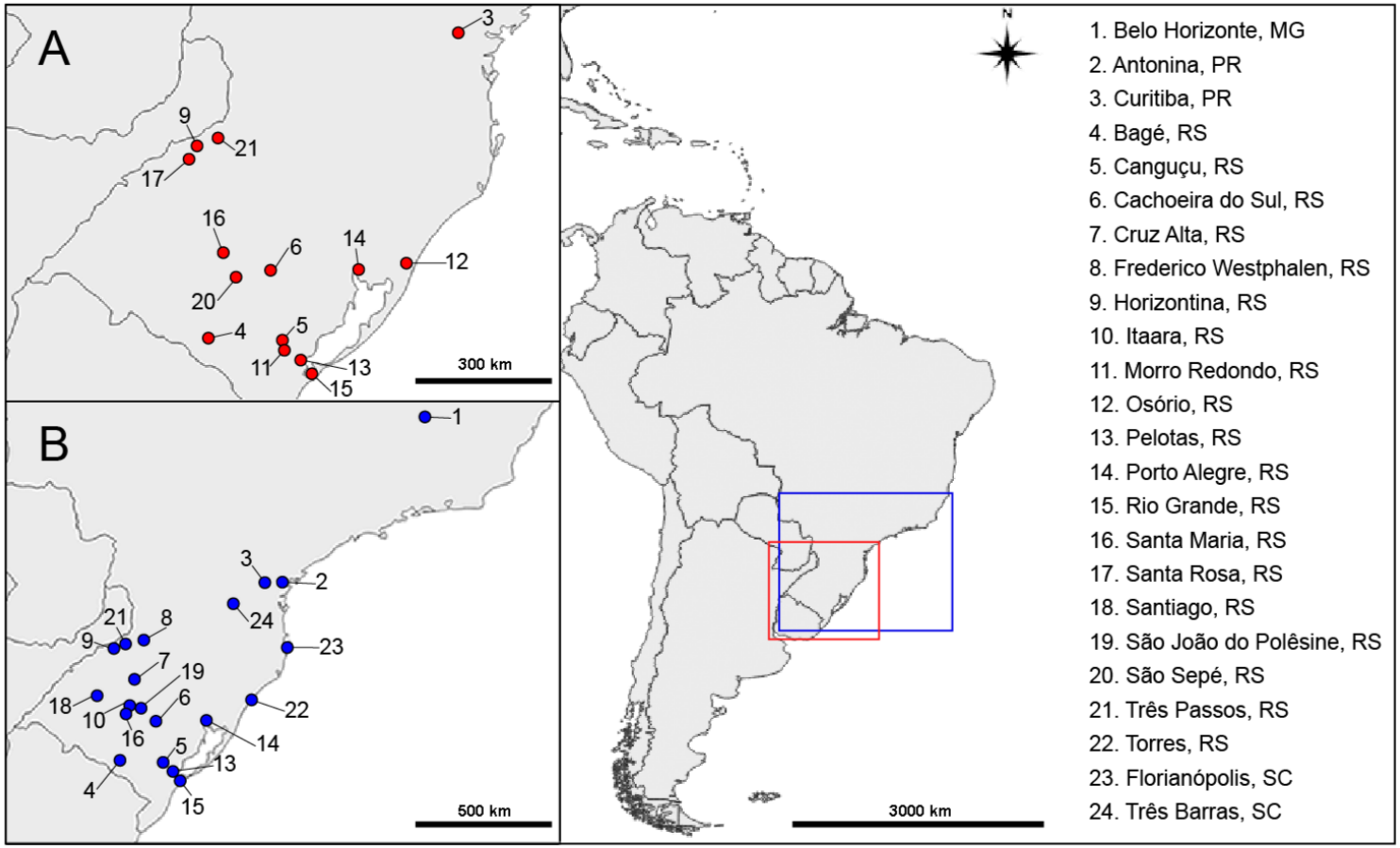


Figure 1. Map of the sample points for (A) *Drosophila cestri* and (B) *Drosophila incompta* across the Brazilian states of Minas Gerais (MG), Paraná (PR), Rio Grande do Sul (RS) and Santa Catarina (SC).

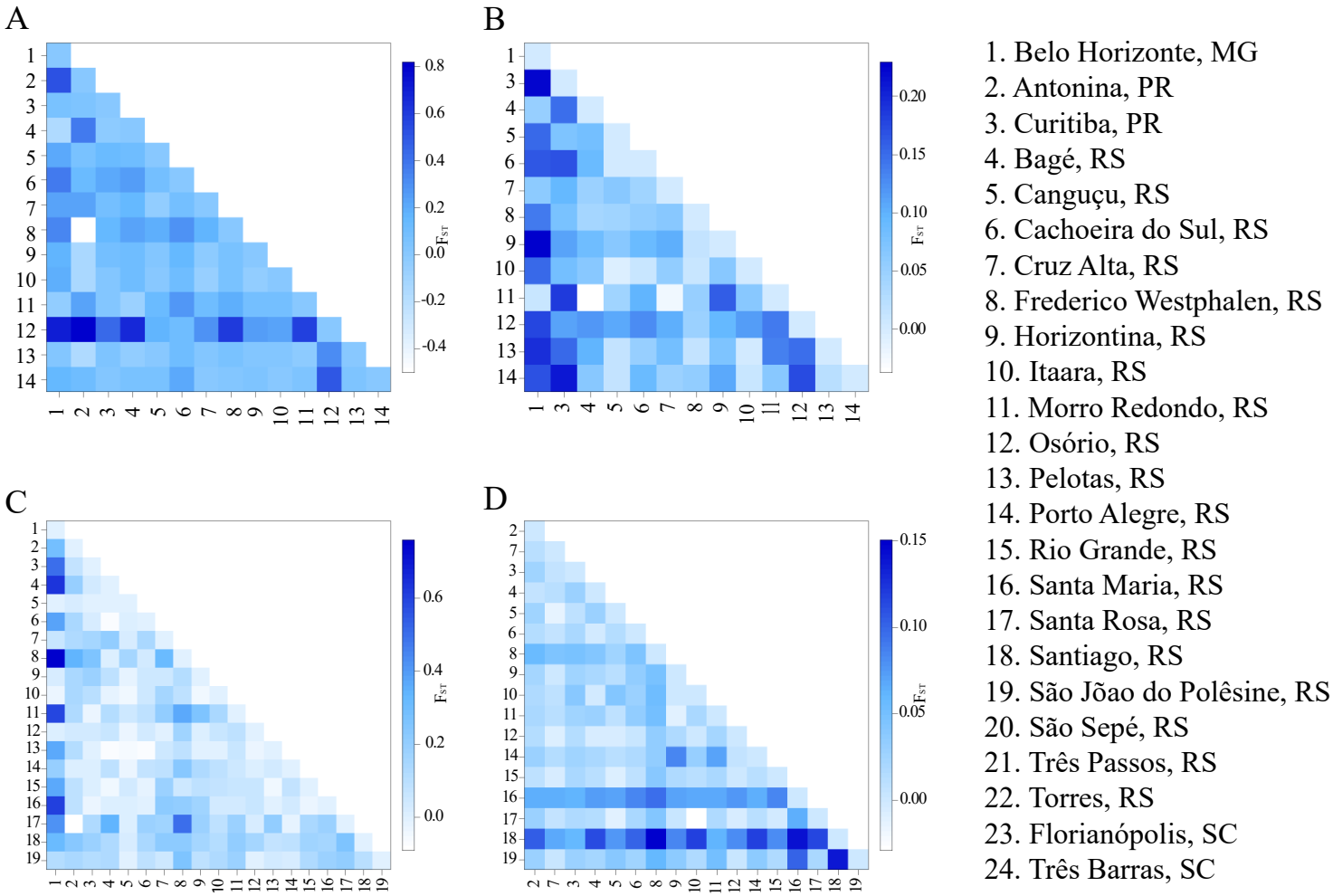


Figure 2. Graphics of F_{ST} values for *COI* (A and C) and *E* (B and D) genes of *Drosophila cestri* (A and B) and *Drosophila incompta* (C and D). Colors correspond to the F_{ST} values range.

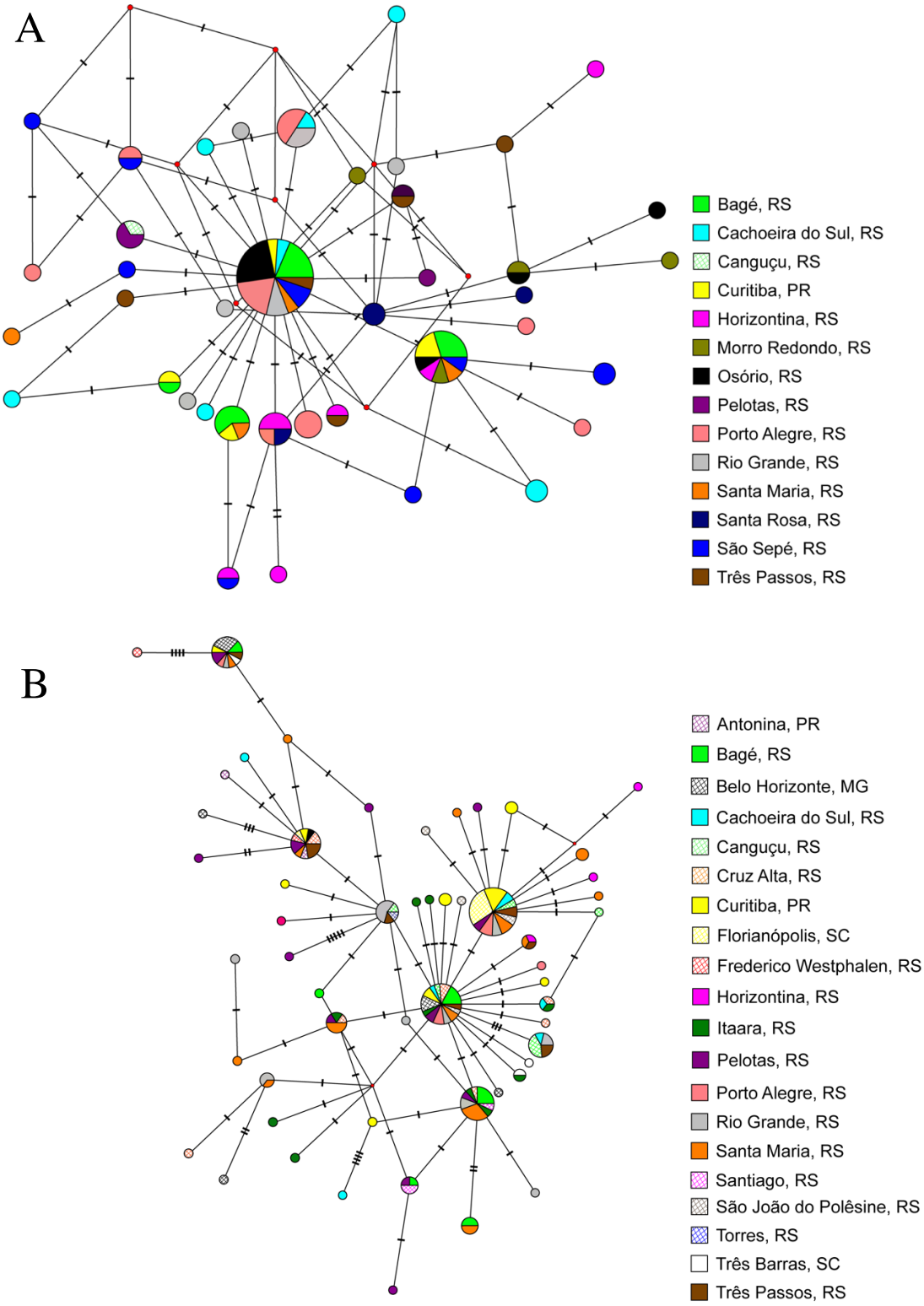


Figure 3. Median-joining network for the *COI* haplotypes of *D. cestri* (A) and *D. incompta* (B). Perpendicular lines correspond to the number of mutational steps between two haplotypes, represented by circles. Red circles correspond to median vectors. The haplotype size corresponds to its frequency.

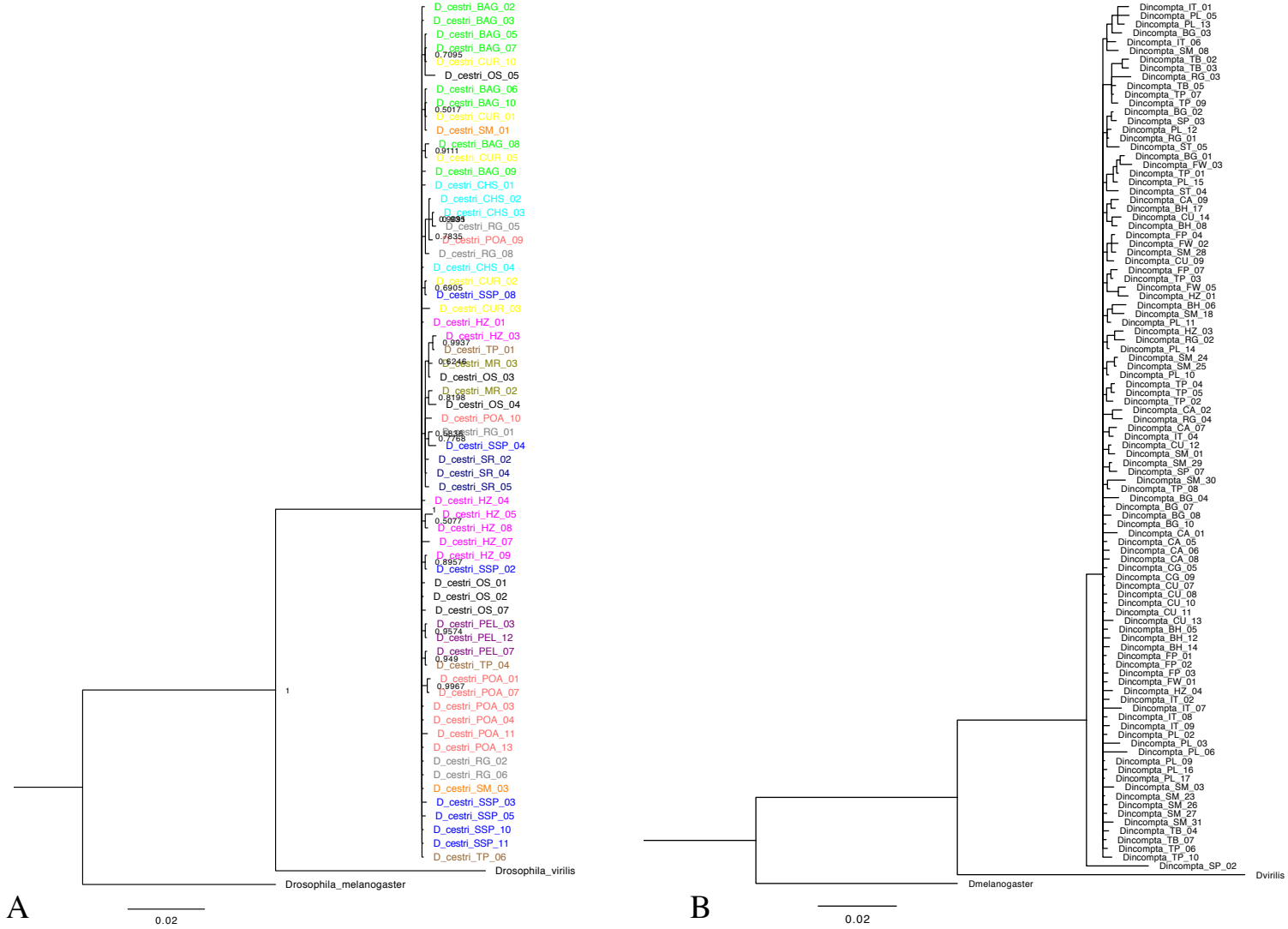


Figure 4. Majority rule consensus tree inferred through BI in MrBayes using the multilocus dataset of *Drosophila cestri* (A) and *D. incompta* (B). Numbers in front of each internal node reflect the posterior probability of each clade. Terminal node names were presented in different colors, according to their origin.

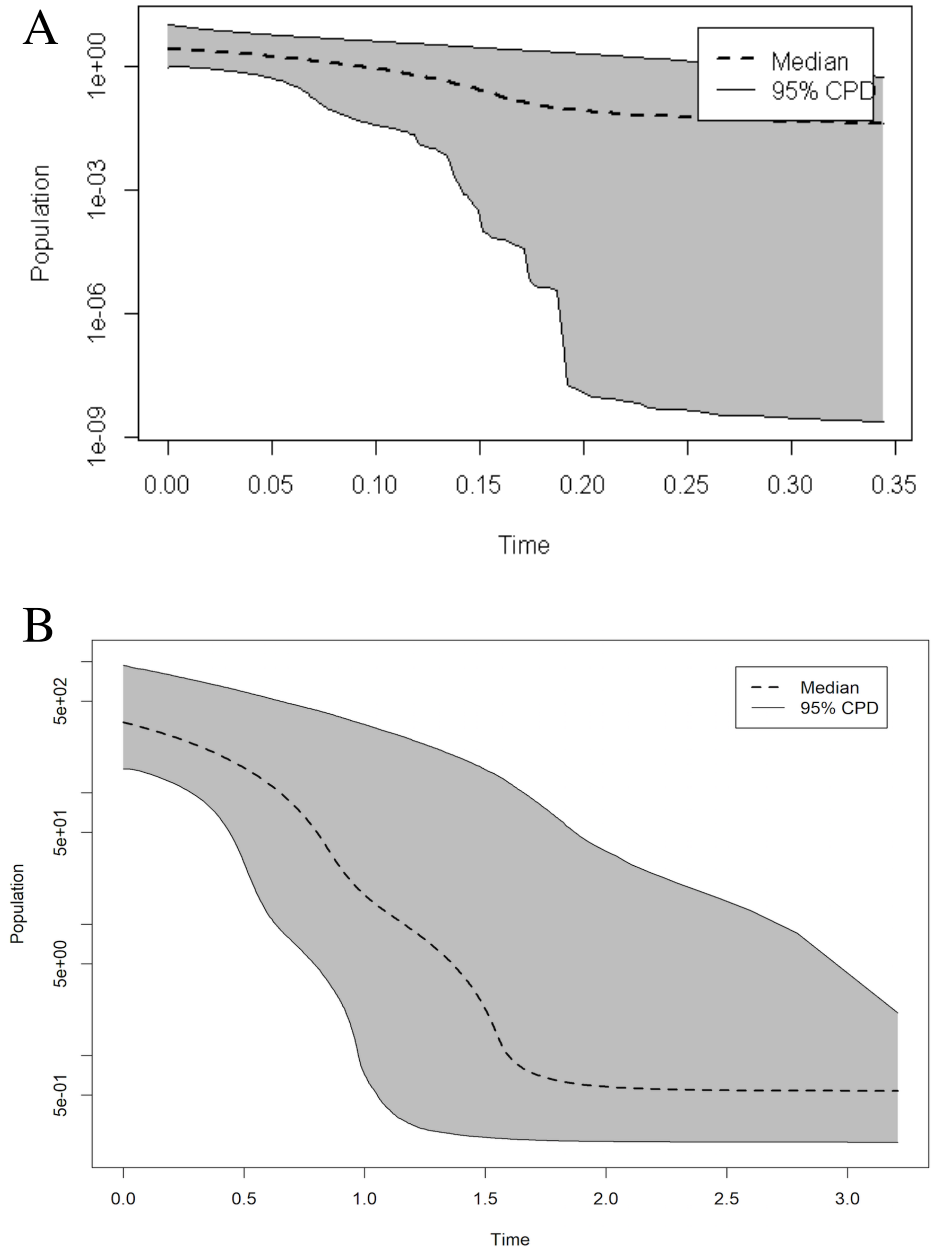


Figure 5. Variation in effective population size inferred for *D. cestri* (A) and *D. incompta* (B) through an Extended Bayesian Skyline Plot (EBSP) reconstructed with the total set of sequences of *COI* and *E.*

Tables

Table 1. List of sampling points and their respective geographical coordinates and number of sequenced individuals for *D. cestri* and *D. incompta*. "Multi-locus" indicates the number of specimens with both *COI* and *E* genes characterized.

		Number of sequenced individuals						Geographical coordinates	
		Drosophila cestri			*Drosophila incompta*				
Locality	**State**	**COI**	**E**	**Multi-locus**	**COI**	**E**	**Multi-locus**	**South (S)**	**West (S)**
Belo Horizonte	MG	0	0	0	11	9	6	19°49′1″	43°57′21″
Antonina	PR	0	0	0	1	0	0	25°25′44″	48°42′43″
Curitiba		10	7	5	14	10	8	25°25′47″	49°16′19″
Bagé	RS	10	8	8	10	7	7	31°19'48.51"	54°6'1.66"
Canguçu		1	0	0	7	5	2	31°23′42″	52°40′33″
Cachoeira do Sul		5	5	4	7	5	0	30°2′20″	52°53′38″
Cruz Alta		0	0	0	11	7	7	28°34'06.95"	53°37'20.72"
Frederico Westphalen		0	0	0	5	5	4	27º21'32.84"	53º23'46.81"
Horizontina		7	11	7	4	3	3	27°37′33″	54°18′28″
Itaara		0	0	0	10	9	7	29º35'27.36"	53º45'31.00"
Morro Redondo		4	7	2	0	0	0	31°35'20.58"	52°38'34.98"
Osório		6	7	6	0	0	0	29°53'32.04"	50°15'43.32"
Pelotas		4	6	3	15	16	13	31°46′19″	52°20′34″
Porto Alegre		12	9	8	6	0	0	30°2'4.73″	51°13'3.57″
Rio Grande		7	8	5	14	4	4	32°2′6″	52°5′56″
Santa Maria		7	2	2	29	16	13	29°41′2″	53°48′25″
Santa Rosa		4	7	3	0	0	0	27°52'14.46"	54°28'48.56"
Santiago		0	0	0	4	3	2	29º11'29.26"	54º51.'59.51"
São João do Polêsine		0	0	0	7	4	3	29°38'59.73"	53°31'00.33"
São Sepé		12	8	7	0	0	0	30° 9' 50.36"	53°34'18.56"
Três Passos		5	3	3	10	10	10	27°27'21.36"	53°55'48.57"
Torres		0	0	0	2	0	0	29°20'18.23"S	49°43'41.61"W
Florianópolis	SC	0	0	0	7	8	5	27°35′49″	48°32′56″
Três Barras		0	0	0	9	7	5	26°6′21″	50°19′19″
Total		**94**	**88**	**63**	**183**	**128**	**99**		

COI = cytochrome oxidase subunit I; E = Ebony; MG = Minas Gerais; PR = Paraná; RS = Rio Grande do Sul; SC = Santa Catarina.

Table 2. Estimations of genetic diversity and neutrality tests performed for *D. cestri* and *D. incompta*.

	Drosophila cestri			*Drosophila incompta*		
	COI	**Ebony**	**Multi-locus**	**COI**	**Ebony**	**Multi-locus**
N	94	176	126	182	256	194
S	21	100	192	31	254	257
h or a	31	112	116	30	252	191
Hd or He (sd)	0.889 (0.023)	0.985 (0,004)	0.998 (0.002)	0.801 (0.020)	1.000 (0.000)	1.000 (0.001)
π (sd)	0.00305 (0.00023)	0.00523 (0,00036)	0.00964 (0.0022)	0.00343 (0.00024)	0.02551 (0.00058)	0.01690 (0.0004)
Tajima's D	-1.89508*	-2.42245**	-2.53419***	-2.08262*	-2.14384**	-2.0882*
Fu & Li's D	-2.41065*	-6.16539**	-3.82486**	-5.83215**	-3.83953**	-3.13511*
Fu & Li's F	-2.64835*	-5.31249**	-3.87368**	-5.13704**	-3.52494**	-3.1226**

COI = Cytochrome oxidase subunit I; h or a = number of haplotypes or alleles, for COI and E, respectively; Hd or He = haplotype diversity or expected heterozygosity, for *COI* and *E*, respectively; N = number of sequences; S = number of polymorphic sites; π = nucleotide diversity; * $P < 0.05$; ** $P < 0.02$; *** $P < 0.001$.

Table 3. Pairwise FST values recovered in the comparisons between populations of *Drosophila cestri* based on *COI* (below diagonal) and *E* sequences (above diagonal).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Bagé, RS		--	**0.22396***	**0.05387***	**0.15036***	**0.16499***	**0.06134***	**0.13951***	**0.22941***	**0.15136***	0.00826	**0.17601***	**0.19413***	**0.16881***
2. Canguçu, RS	0.51111	--	--	--	--	--	--	--	--	--	--	--	--	--
3. Cachoeira do Sul, RS	0.06433	0.04040		**0.14970***	**0.07255***	**0.17086***	**0.09456***	**0.07794***	**0.11124***	**0.08034***	**0.18367***	**0.10924***	**0.15030***	**0.20891***
4. Curitiba, PR	-0.15974	0.37500	-0.00877		**0.08090***	**0.08956***	**0.04780***	0.03911	**0.08529***	**0.06459***	-0.03787	**0.11980***	0.09783	0.06609
5. Horizontina, RS	**0.20079***	0.05882	**0.13775***	0.11537		0.00271	**0.06647***	**0.04385***	**0.06587***	-0.00575	0.04246	**0.10799***	0.01528	0.01226
6. Morro Redondo, RS	**0.36945***	0.12821	**0.20376***	0.24699	0.08289		**0.04219***	**0.05805***	**0.08989***	0.01090	0.09777	**0.12560***	**0.05232***	0.08791
7. Osório, RS	**0.23019***	0.23636	0.09091	0.15601	-0.03783	0.09379		**0.06290***	**0.10272***	**0.05696***	-0.02450	**0.10046***	**0.08180***	0.04634
8. Pelotas, RS	**0.32513***	-0.50000	0.14037	**0.23077***	**0.18297***	0.29252	0.17489		0.01477	0.01112	0.05390	**0.04721***	0.01640	0.05742
9. Porto Alegre, RS	**0.16703***	-0.13636	**0.07950***	0.10124	0.02338	0.14740	-0.02128	0.07451		**0.06540***	**0.16159***	**0.09337***	**0.07488***	**0.10451***
10. Rio Grande, RS	**0.18708***	-0.14286	0.01542	0.10256	0.00885	0.10660	-0.06580	0.06114	-0.04035		0.06397	0.11418	**0.00108***	0.06326
11. Santa Maria, RS	-0.03609	0.22222	0.03226	-0.08247	0.12211	**0.27187***	0.11765	**0.18129***	0.09471	0.09524		**0.14020***	**0.13635***	0.06326
12. Santa Rosa, RS	**0.70516***	0.81818	**0.44388***	**0.66867***	0.16012	0.11111	0.28546	**0.60976***	**0.24734***	0.24093	**0.58664***		**0.14619***	**0.17510***
13. São Sepé, RS	0.02130	-0.15152	0.04983	-0.03489	0.00632	0.10078	0.02553	0.07494	0.01948	0.01645	-0.00792	**0.31852***		0.02463
14. Três Passos, RS	0.15254	0.11111	0.01855	0.07143	0.06935	0.21081	0.00248	0.02939	0.04054	0.00348	0.03743	**0.50871***	0.04717	

Asterisks indicate statistically significant values (P<0.05).

Table 4. Pairwise FST values recovered in the comparisons between populations of *Drosophila incompta* based on *COI* (below diagonal) and *E* sequences (above diagonal).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Antonina, PR		--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
2. Bagé, RS	0.28775		0.02450	0.00481	0.02654	0.01393	0.01083	**0.05717***	0.02147	0.01768	0.02050	0.01286	--	0.02693	0.00755	0.06385	0.02706	**0.10366***	**0.02908***
3. Cruz Alta, RS	0.49744	0.07723		0.02857	0.00979	0.01975	0.01975	**0.04767***	0.02444	0.04043	0.02549	0.01448	--	0.02517	0.02100	**0.06325***	0.01763	**0.06059***	**0.02883***
4. Canguçu, RS	0.63810	**0.19932***	0.02769		0.02844	-0.00351	-0.00351	**0.04127***	0.02018	-0.00162	0.02875	-0.00654	--	0.01947	0.01038	**0.07398***	0.01431	**0.11223***	0.02205
5. Cachoeira do Sul, RS	-0.00000	0.02549	-0.00614	-0.00202		0.01470	0.01470	0.02467	0.01265	0.04212	0.00964	-0.00642	--	0.01175	-0.00863	**0.06936***	-0.01181	**0.07698***	**0.03822***
6. Curitiba, PR	0.38462	**0.13496***	0.02015	-0.07445	0.01285		0.00502	**0.04495***	**0.03395***	0.02839	**0.03511***	0.00200	--	0.01656	0.00828	**0.08517***	0.01971	**0.10024***	**0.02428***
7. Belo Horizonte, MG	0.05946	0.13005	**0.15955***	**0.20372***	0.01983	**0.14358***		**0.04504***	-0.00029	0.00800	0.01110	-0.00875	--	0.01351	-0.00557	**0.06379***	0.00665	**0.06833***	0.00389
8. Florianópolis, SC	0.76000	**0.34074***	**0.25807***	0.01403	**0.15436***	0.03309	**0.31739***		**0.04765***	0.05380	**0.04955***	**0.03352***	--	0.03385	**0.04090***	**0.09712***	0.04134	**0.15061***	**0.06003***
9. Frederico Westphalen, RS	0.02632	0.13816	**0.18670***	0.08864	-0.01717	0.05231	0.04188	0.11843		-0.00118	-0.01189	0.01701	--	**0.08480***	0.01791	**0.06960***	0.00258	**0.08479***	0.03542
10. Horizontina, RS	-0.03030	0.13846	0.09781	-0.01818	-0.04755	0.04584	0.15052	0.08843	-0.04058		0.02092	0.00091	--	0.03154	0.01140	0.06770	-0.02935	**0.11853***	0.01749
11. Itaara, RS	0.59064	**0.13333***	-0.03178	**0.10949***	0.04100	**0.08576***	0.20612	**0.36086***	**0.26718***	**0.14068***		**0.02160***	--	**0.07082***	**0.01956***	**0.06679***	0.01600	**0.06137***	**0.05065***
12. Pelotas, RS	0.00809	0.03189	0.03821	0.05640	-0.03077	0.03177	-0.00335	**0.16412***	0.00319	0.06213	**0.09143***		--	0.00452	-0.00016	**0.07712***	-0.00418	**0.08112***	0.00296
13. Porto Alegre, RS	0.37143	0.11184	-0.01024	-0.07586	-0.05524	-0.07291	0.04786	0.08237	-0.00781	-0.00599	0.05897	-0.02854		--	--	--	--	--	--
14. Rio Grande, RS	0.19161	0.01036	0.00272	0.08391	-0.03761	**0.07019***	0.08754	**0.23811***	**0.12451***	0.08373	**0.04717***	0.00618	0.02826		0.01114	0.06381	0.01166	**0.11783***	0.03727
15. Santa Maria, RS	0.36716	**0.09339***	0.00323	-0.04131	0.03046	-0.01827	**0.17489***	0.05236	0.08409	0.07749	0.06643	**0.05594***	-0.04990	**0.06940***		**0.08437***	-0.00288	0.08077	0.01622
16. São João do Polêsine, RS	0.60494	0.09352	-0.03029	0.00438	0.00466	-0.01105	**0.21594***	0.21171	**0.16846***	0.03829	0.02853	0.06624	-0.01093	0.03927	-0.03694		0.06585	**0.14140***	**0.10128***
17. Santiago, RS	0.42222	-0.09233	0.12338	**0.33976***	0.01120	**0.17882***	0.16753	**0.50429***	0.18108	0.09179	**0.18530***	0.02517	0.20530	-0.01796	**0.13986***	0.17614		**0.11560***	0.02067
18. Três Barras, SC	0.31395	**0.23814***	**0.17429***	0.12348	0.02025	**0.13531***	0.21522	**0.21717***	0.15359	0.11574	**0.20649***	**0.15302***	0.07953	0.09332	**0.16289***	0.15356	**0.26175***		**0.13780***
19. Três Passos, RS	0.11688	**0.14667***	**0.11846***	0.13092	-0.01004	**0.09386***	0.05398	**0.26705***	0.08933	0.12610	**0.18341***	-0.00305	0.03739	0.02450	**0.12332***	**0.17776***	0.20305	0.09885	

Asterisks indicate statistically significant values (P<0.05).

Supplementary material

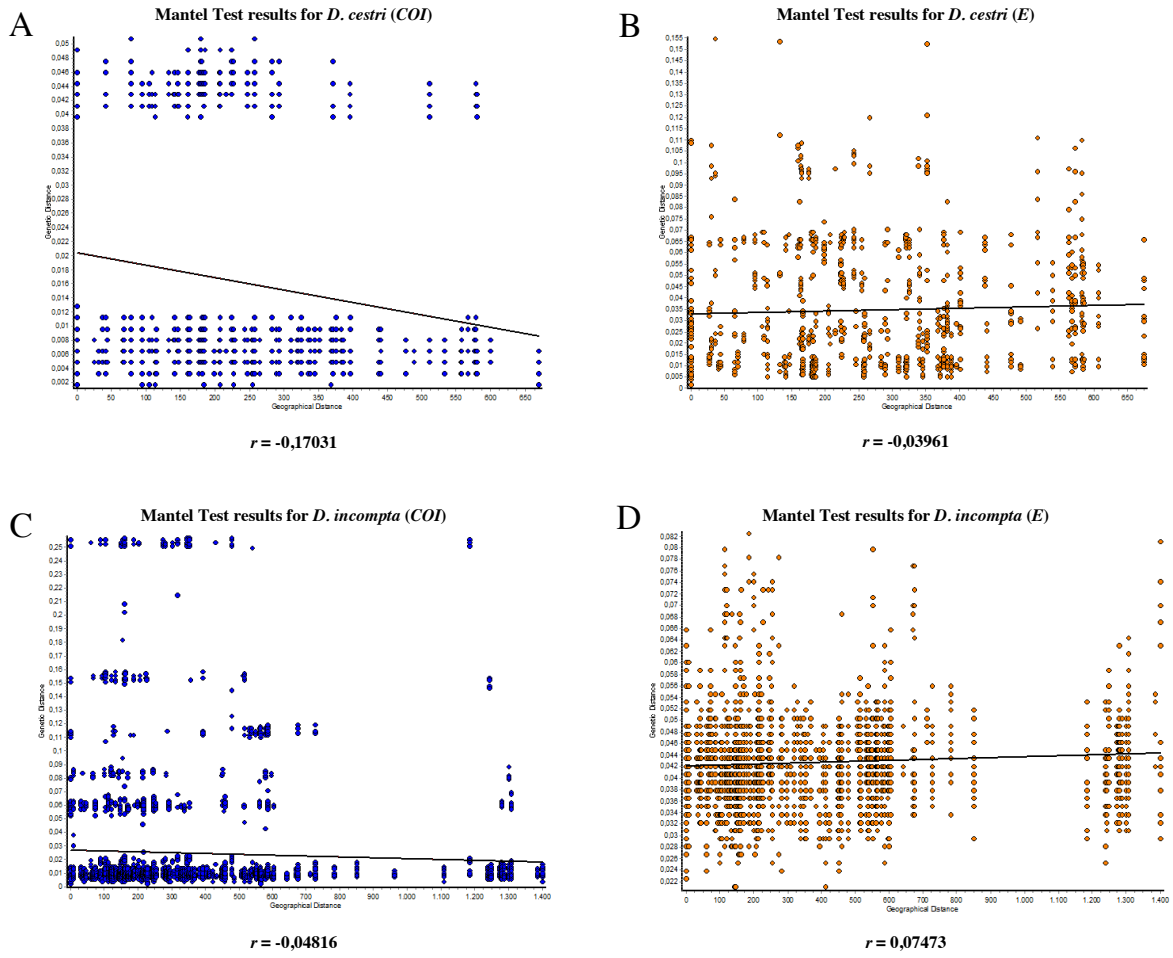


Figure S 1. Graphics of the Mantel Test results, addressing the isolation-by-distance hypothesis, for *D. cestri* (A and B, for *COI* and *E* genes, respectively) and *D. incompta* (C and D, for *COI* and *E* genes, respectively). Each circle corresponds to a pairwise estimation value between genetic (Y axis) and geographical (X axis) distance.

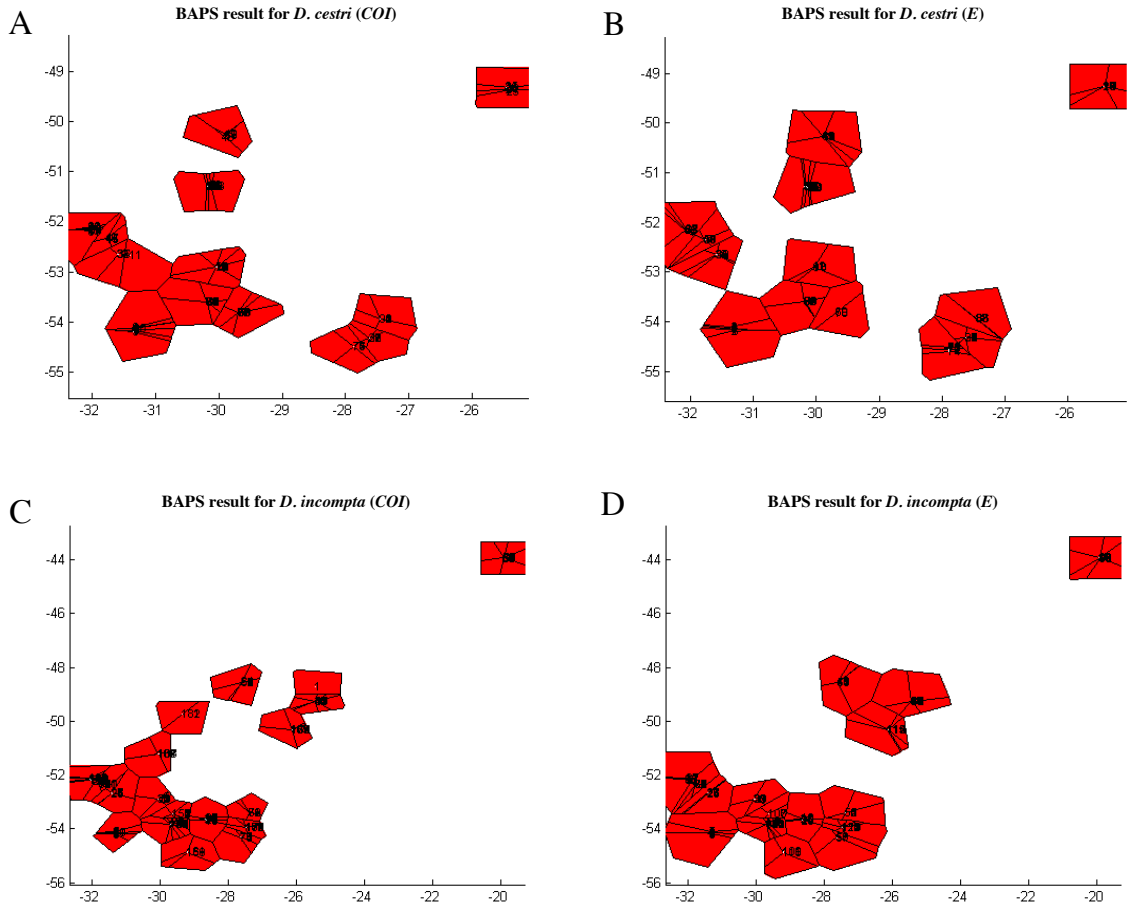


Figure S 2. BAPS results from the *COI* (A and C) or *E* (B and D) genes of *Drosophila cestri* (A and B) or *D. incompta* (C and D). The geographical distribution of each population is presented as a function of its geographical coordinates [latitude (x-axis) and longitude (y-scores)], with clusters distinguished by color.

CAPÍTULO IV – CONCLUSÕES E CONSIDERAÇÕES FINAIS

Esta Dissertação focou em estudar a história evolutiva e a influência da especialização ecológica na diversificação e processos de estruturação populacional do grupo *flavopilosa* de *Drosophila*. Inicialmente, no **Capítulo II**, são avaliadas as relações filogenéticas dentro do grupo, bem como a eficácia da técnica de DNA *barcode* nessas espécies e a datação de eventos de divergência. Este capítulo também auxilia na resolução de questões relacionadas ao posicionamento taxonômico de algumas espécies do grupo. Os resultados obtidos suportam e validam estudos anteriores, confirmando a técnica de DNA *barcode* como uma alternativa no reconhecimento das espécies desse grupo, em especial para as fêmeas – as quais não possuem chaves de identificação.

A subdivisão nos subgrupos *flavopilosa* e *nesiota* também foi recuperada e, como hipotetizado por estudos anteriores (ROBE et al., 2013), os resultados indicam que *D. cestri* e *D. cordeiroi* devem ser alocadas dentro do subgrupo *flavopilosa*. Além disso, os dados moleculares corroboram o posicionamento de *Drosophila mariaehelenae* dentro do subgrupo *nesiota* (VILELA, 1983). A análise cronofilogenética, por sua vez, datou a origem do grupo para aproximadamente 10 milhões de anos atrás e indicou que as diversificações subsequentes provavelmente ocorreram nos últimos 5 milhões de anos – antes, contudo, da formação do istmo do Panamá (cerca de 2,8 milhões de anos atrás) (O'DEA et al., 2016). Nesse cenário, a origem geográfica do grupo parece ser a América do Sul; no entanto, análises biogeográficas que incluam OTU's distribuídas na América Central e do Norte ainda devem ser realizadas para confirmar essa hipótese.

As análises populacionais executadas no **Capítulo III** indicam que *D. cestri* e *D. incompta* compartilham padrões semelhantes de distribuição da diversidade genética. Ambas espécies apresentaram valores baixos de diversidade nucleotídica e valores altos de diversidade haplotípica ou de heterozigosidade esperada, os quais, junto aos valores negativos significativos para os testes de neutralidade, sugerem expansão populacional recente. As redes de haplótipos, valores de F_{ST} e as análises de BAPS corroboraram os resultados anteriormente descritos e demonstraram baixa estruturação em suas populações; além disso, os testes de Mantel descartaram a hipótese de isolamento por distância em ambas espécies, uma vez que a distância genética em suas populações não é explicada pela distância geográfica.

As estimativas das dinâmicas populacionais acessadas por meio de *skyline plots* indicam que essas expansões ocorreram em momentos distintos para cada espécie. Enquanto *D. cestri*

apresentou uma pequena expansão populacional durante o Eemiano (o último período interglacial), as populações de *D. incompta* aparentemente começaram a expansão na metade do Pleistoceno. Nesse sentido, a hipótese levantada por DE RÉ et al. (2014) – a qual considera os altos níveis de polimorfismo encontrados no genoma mitocondrial de *D. incompta* uma consequência de expansões/retrações populacionais cíclicas e migrações em busca de indivíduos de *Cestrum* com flores – parece ser a que melhor explica os resultados obtidos e a sobrevivência de espécimes do grupo *flavopilosa* durante os períodos sem recurso disponível.

Referências Bibliográficas

DE RÉ, F. C. et al. Characterization of the complete mitochondrial genome of flower-breeding Drosophila incompta (Diptera, Drosophilidae). **Genetica**, v. 142, n. 6, p. 525–535, 1 dez. 2014.

O'DEA, A. et al. Formation of the Isthmus of Panama. **Science Advances**, v. 2, n. 8, p. e1600883, 1 ago. 2016.

ROBE, L. J. et al. The Drosophila flavopilosa species group (Diptera, Drosophilidae): an array of exciting questions. **Fly**, v. 7, n. 2, p. 59–69, 21 abr. 2013.

VILELA, C. R. Occurrence of the Drosophila flavopilosa species group (Diptera, Drosophilidae) in the state of São Paulo (Brazil) with description of one new species. **Revista Brasileira de Zoologia**, v. 2, n. 2, p. 63–69, 1983.

APÊNDICE I

Manuscrito submetido ao periódico *Genetics and Molecular Biology*[§]

Bias, gaps and global progress among phylogeographical studies with Drosophilidae (Insecta: Diptera)

Henrique da Rocha Moreira[1], Lizandra Jaqueline Robe[2] and Maríndia Deprá[1,3*]

[1]Programa de Pós-Graduação em Biologia Animal, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil;
[2]Programa de Pós-Graduação em Biodiversidade Animal, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brazil;
[3]Programa de Pós-Graduação em Genética e Biologia Molecular, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil;

*Corresponding author: Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul
Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil
E-mail: marindiadepra@gmail.com

[§] Normas e regras de submissão disponíveis em https://www.gmb.org.br/submission-of-papers

Abstract

Since its establishment some 30 years ago, phylogeographical approaches have been used to enlighten questions related to the historical distribution of current lineages. For over a century, Drosophilidae species have been used as model organisms in several fields of study, from urban environmental health to ecological genetics. Given this, the aim of this work was to (i) evaluate the bias, gaps and progress related to scientific production of phylogeographical studies with drosophilids; (ii) compile the results and conclusions reached by those papers; and (iii) identify similarities or incongruities within the phylogeographical history recovered for different species or regions. Papers about phylogeography with Drosophilidae published in peer-reviewed journals from 1987 to 2018 were compiled. A total of 184 papers were analyzed, from which 32 addressed phylogeographical questions with drosophilids. The results showed no linear correlation between the number of papers published and year, and several bias concerning both biogeographical regions and species or species groups were detected. The obtained evidences show that several Drosophilidae species were affected by the Pleistocene climatic oscillations, even in the Neotropics. This emphasizes the importance of using these species as model organisms to shed some light into biogeographical patterns related to the entire biodiversity of a region.

Keywords: *Drosophila*; model species; molecular clock; molecular markers; scientometric research.

Introduction

About 30 years ago, Avise et al. (1987) discussed the influence and the extension of microevolutionary processes in the macroevolutionary differences between species. In this sense, they pointed a lack of a field of study that would connect population genetics with phylogenetic systematics. The term "phylogeography" was then established to designate studies that integrate genealogies and biogeography in temporal and spatial dimensions, with an emphasis on historical aspects that are able to explain the distributions of current lineages (Avise, 2009).

In this sense, phylogeography has increasingly been used to solve questions related to the origin of the current biodiversity patterns, showing an exponential growth since its establishment (Beheregaray, 2008). Among those studies, phylogeography has been remarkably employed in the context of speciation and conservation biology of several taxa, e.g. animals (da Silva and Patton, 1998; Terranova et al., 2007) and plants (Hodel et al., 2016; Tembrock et al., 2017). Moreover, phylogeographical studies may also help to understand general patterns of species diversity and distribution (Hewitt, 2004). Concerning conservation genetics, phylogeographical evaluations have provided important insights for the establishment of effective conservation strategies aiming to recover endangered species (Dool et al., 2016; Garcez et al., 2018; Rosauer et al., 2016).

Phylogeography also helps in current assessments of diversity and conservation by enabling the discovery of new species or lineages (Alfaro et al., 2018). In the last field, phylogeography may contribute to understand the ecological (Bonebrake et al., 2010) and evolutionary relationships (Rosauer et al., 2016) between populations or species, which are essential not only to address speciation or diversification patterns but to comprehend the real dimensions affecting survival or extinction. Moreover, phylogeography may help in the evaluation of general patterns of genetic variability [as intra- and interpopulation structure (Eizirik et al., 2001) and demographic history (Zhang et al., 2017)] and assist in the discovery of historical processes affecting biodiversity [such as vicariance or split into refugia (Song et al., 2018), expansion and secondary contact (Gum et al., 2005), inference of introgression (Jeratthitikul et al., 2013) and coevolution (Cuthill and Charleston, 2015)].

For over a century, drosophilids have been greatly used as model organisms in several fields of study (Kohler, 1993; Markow and O'Grady, 2006). In the field of phylogeography, these species have not only contributed to the understanding of common (Izumitani et al., 2016) or peculiar evolutionary histories (De Ré et al., 2014; Fonseca et al., 2017), but to assess general

patterns of diversification (Markow, 2019). Drosophilidae is a diverse and broadly distributed dipteran family, that comprises more than 4,000 species occurring from tundra to tropics (Bächli, 2020; Throckmorton, 1975), that may explore the most distinct substrates (Carson, 1971; Powell, 1997; Tosi et al., 1990). The taxonomists commonly divide the family into radiations (or lineages), which represent multiple speciation events and their consecutive diversifications (Throckmorton, 1975). *Drosophila* radiations were recently ranked to the category of subgenera (Yassin, 2013) and are further divided into species groups, which cluster morphologically related species (Grimaldi, 1990, 1987).

Given the relevance of Drosophilidae as an excellent and well known group of model organisms, and the role of phylogeography in understanding general patterns and processes that affect species evolution, the aim of this work was to verify and evaluate the progress, bias and gaps related to scientific production of phylogeographical studies with drosophilids. We also aimed to compile the results and conclusions reached by those papers, identifying similarities or incongruities among the phylogeographical history recovered for different species of Drosophilidae.

Methods

The literature surveys were conducted on January 2019, in the Web of Science (Clarivate Analytics) and Scopus (Elsevier) databases, using the words "phylogeograph*" and "drosophil*" to search in the titles, abstracts and key-words. The period analyzed (1987-2018) starts in the second half of 1980, when the term phylogeography was first coined by Avise et al. (1987). The resulting papers had their titles and abstracts (and full text, when necessary) inspected to filter out only phylogeographical studies with Drosophilidae.

The selected papers were manually listed to a Microsoft Excel spreadsheet, and basic information was extracted from each one, such as (i) year of publication, (ii) journal of publication, (iii) number of species and species group(s) included [according to Bächli (2019)], (iv) country and institution of the first author, (v) total number and inheritance patterns of genetic markers used in the study, (vi) whether morphological data were included in the methodology or not, (vii) whether molecular clock was used to estimate divergence/diversification times or not, and (viii) total number and name of the biogeographical region(s) of the species included. We obtained the journal impact factors reported in the 2017 Edition of the Journal Citation Reports (Clarivate Analytics) and biogeographic regions of the sampled species followed the regionalization proposed by

Morrone (2015). Geological periods followed the scale proposed by Walker et al. (2013). The historical trend in the number of publications per year was evaluated using Pearson correlation in R software (RStudio Team, 2015).

Results

The surveys resulted in 184 papers retrieved from the Web of Science database and 73 papers retrieved from the Scopus database. After the first inspection, the resulting papers were manually reexamined to remove duplicates. From these, 152 papers were further excluded because they were based in other works with Drosophilidae. Our analyses, therefore, counted with 32 papers which are listed on Table S1.

A peak in the number of publications can be noticed in 2007 and 2008, followed by a decrease and subsequent stabilization in relation to the number of publications per year (Figure 1). Congruent with this, the Pearson correlation coefficient scored $r = 0.1196$, hence indicating no linear correlation between the number of papers published and year. Although the papers were published in a few journals, most of them have high impact factors, with a weighted average of 4.25 (Table 1). Most of the journals have published only one paper along the period analyzed, which represents 3.13% of the total number of publications. Nonetheless, the two preferred journals ("Molecular Ecology" and "Molecular Phylogenetics and Evolution") published more than 1/3 of the total number of papers.

In most of the papers (56.25%), a single mitochondrial locus was chosen as molecular marker (Figure 2A, Table S1). About 28% of the articles combined mitochondrial (mtDNA) and nuclear DNA (nDNA) in the phylogeographical analyses. Wilder and Hollocher (2003) were the first to use this perspective, choosing the Cytochrome c Oxidase subunit II (COII) gene as mtDNA locus and the nuclear genes Cu/Zn-*superoxide dismutase*, *engrailed* and *male fertility factor kl-5* (*sod*, *en* and *kl-5*, respectively) as nDNA markers (Table S1). Studies investigating phylogeographical patterns through solely nuclear DNA correspond to nearly 16% (Figure 2A, Table S1). Kopp et al. (2006) were the first to choose only nuclear loci (*kl-2* and *kl-3* genes) instead of a mtDNA marker or a nuclear versus mitochondrial gene approach (Table S1). Nonetheless, three authors used microsatellite data, either combined with mtDNA (Mirol et al., 2007; Schiffer et al., 2007) or as the only molecular marker (Fraimout et al., 2017). In addition to the molecular data, seven articles combined morphological analyses into the phylogeographical methodologies (Figure 2B).

The sole inference of inter- or/and intrapopulation variability levels, population structure and demographic history compose the most addressed question (ten papers) (Figure 3). Inference of historical processes affecting other biodiversity scales were performed solely or in combination with other aims in other 11 papers (Figure 3). In total, molecular clocks were employed to estimate divergence or diversification times between lineages in ten articles (31.25%), from which seven and two studies dated the events to Pleistocene and Pliocene Epochs, respectively. Izumitani et al. (2016) was excluded from this analysis due to the large *Drosophila* taxonomic sampling and will be discussed in particular in the next sections.

The species studied belong to the *Drosophila*, *Idiomyia*, *Lordiphosa*, *Scaptodrosophila*, *Scaptomyza* and *Zaprionus* genera. Regarding species groups, the *melanogaster* group of *Drosophila* appears as the most studied (considered in 12 articles), followed by the *repleta* and *immigrans/virilis* groups of *Drosophila* (nine and four articles, respectively) (Table 2). Most of the articles studied only one species, followed by three or more species (20 and 10 articles, respectively).

The Nearctic region was explored in 15 of the 32 studies, and comprised the most studied biogeographical region among the papers (Figure 4, Table S1). The Palearctic and Neotropical follow as the second most studied regions, each representing 13 of the papers. In contrast, the Ethiopian region appears only in three articles, less than 10% of the total. When a single biogeographical region was accessed, the Neotropical region stands as the most solo region studied, in a total of eight papers (Figure 4), followed by the Nearctic region (five papers). These findings may be related to the affiliations of the authors, as the United States, Brazil and Japan encompassed the three countries with more publications, scoring nine, six and four published articles, respectively (Figure 5, Table S2).

Discussion

Although phylogeography started as a new discipline with the publication of Avise et al. (1987), phylogeographical studies with Drosophilidae began only 12 years later, with the publication of Kelemen and Moritz (1999) in the *Evolution* journal. In this paper, a comparative approach was performed in a sibling pair of *Drosophila* species (*D. serrata* and *D. birchii*) belonging to the *melanogaster* group and distributed in the Australian region. A new article was published only about three years later by Brito et al. (2002), in which a single *Drosophila* species (*D. buzzatii*, *repleta* group) was studied, with a focus on Brazilian populations.

Drosophilidae currently counts with 76 genera (Bächli, 2020), though only six of them were studied on a phylogeographical context. Furthermore, although *Drosophila* appears in most of the retrieved articles, only 19 of the 58 species groups of the genus (Bächli 2019) have been used to study phylogeographical patterns. There are, therefore, at least 39 *Drosophila* species groups yet to be phylogeographically studied, in addition to hundreds of species belonging to other genera. In many cases, this bias is possibly an outcome of the difficulty of achieving a sufficient number of samples. Nonetheless, we also highlight a bias that is strengthened in the last two years towards studying the *melanogaster* group. In fact, this species group accounts for 37.5% of the published studies, and two of the four papers published since 2017 focus on *Drosophila suzukii*, a pest and invasive species, native to Asia (Asplen et al., 2015), that feeds and/or breeds in healthy fruits (Goodhue et al., 2011). In this case, Fraimout et al. (2017) and Choi et al. (2018) used a phylogeographical approach to clarify how this species spread around the globe until its recent arrival in South America (Deprá et al., 2014).

According to estimations performed with molecular clocks, most of the populations of Drosophilidae experienced diversifications within the Neogene and Quaternary periods (Pliocene and Pleistocene Epochs, in particular). The Quaternary period, that ranges from ±2.6 Mya until present (Walker et al., 2018), has been characterized by intermittent glaciations, in a cycle of approximate 100,000 years (Hewitt, 1996), which seem to have impacted the dynamics and population sizes of different *Drosophila* species (Barrios-Leal et al., 2018; Brehm et al., 2004; Franco and Manfrin, 2013; Mirol et al., 2008, 2007; Robert Liu et al., 2015; Smith et al., 2012). Nonetheless, Izumitani et al. (2016) performed a large *Drosophila* subgenus sampling to test Throckmorton's hypothesis on its ancestral distributions. Their findings corroborate the hypothesis of migrations from the Old to the New World via Bering Land Bridge, but suggest these events occurred in a wider period of time, ranging from the Oligocene to the Pliocene.

Concerning the biogeographical regions, besides the general pattern of concentration of studies on Nearctic, Neotropical and Palearctic regions, we bring a summary of the essential findings in the following sections.

Palearctic and Nearctic regions

Gao et al. (2011) revised the phylogenetic relationships between the *Lordiphosa* genus and the *Sophophora* subgenus, proposing their divergence as a vicariant event during the Eocene Epoch. According to these authors, this event was possibly related to the breach of the

North Atlantic Land Bridge between Europe and North America. Sillero et al. (2014) detected an expansion of *D. americana* (*virilis* group) during the last interglacial period, in addition to a bottleneck during the last glacial maximum (LGM) followed by a recent expansion. Within the same species group, the mtDNA of *Drosophila virilis* also suggests a population expansion associated with post-glacial colonization, but in this case microsatellites indicate a bottleneck occurring after this event (Mirol et al., 2008). In addition, Mirol et al. (2007) detected the presence of two populations of *Drosophila montana* (*virilis* group) separated during Pleistocene. The contrasting patterns found by these authors for these lineages based on historical demography suggest isolation into two refugia during the LGMs.

Concerning other species of *Drosophila,* Brehm et al. (2004) detected the absence of separation between Moroccan and Iberian haplotypes of *Drosophila subobscura* (*obscura* group), refuting the hypothesis that the strait of Gibraltar acted as a barrier to gene flow. More recently, Novković and Kimura (2015) performed a comparative phylogeographical study between three species of *Drosophila* [*D. albomicans* (*immigrans* group), *D. bipectinata* and *D. takahashi* (*melanogaster* group)] and one species of a *Drosophila*-parasitoid wasp (*Leptopilina ryukyuensis*). Specific phylogeographical patterns were found for each species, thus suggesting no correlation between host and parasite evolutionary histories.

Neotropical region

For species of the *repleta* group, Brito et al. (2002) found low nucleotide diversity and little population structure for *Drosophila buzzatii* (*buzzatii* species complex), and supported the hypothesis that this species has actively migrated to Brazil some time ago. Within the same species complex, Moraes et al. (2009) revealed a subdivision of *D. gouveai* in different lineages, with footprints of range expansion before the LGM triggered by the climatic changes of the Pleistocene. The population structure seen on *D. meridionalis* (*mulleri* subgroup) by Barrios-Leal et al. (2018) also evidenced the presence of different lineages, whose divergence was dated to the end of a glacial period in Pleistocene – around 840 kya. Finally, Smith et al. (2012) found that the split of populations of *D. mojavensis* (*mojavensis* species complex) coincides with interglacial periods, suggesting their history is intimately linked to Pleistocene climatic oscillations. The divergence estimations coincides with the global sea level rise of near 100m during the Holstein interglacial, which could have separated the Baja peninsula from mainland Mexico (Siddall et al., 2007).

In regard to other species groups, De Ré et al. (2014) reported for *D. maculifrons* (*guarani* group) a population expansion near the transition of Pleistocene to Holocene (around 12,000ya). According to this study, the sampled region did not present the environmental requirements suitable for the target species during the LGM, which would imply populations of *D. maculifrons* expanded and settled only after this event. Within the same group, Gustani et al. (2015) inferred that *D. ornatifrons* was most likely affected by the same historical events as *D. maculifrons*, since the authors detected demographic expansions after the LGM, intensified near the transition of Pleistocene to Holocene. Nonetheless, Brisson et al. (2005) found a lack of evidence for historical events in *D. polymorpha* (*cardini* group). Within this species, the restricted gene flow between populations appears to have occurred through patterns of isolation by distance. Such a pattern was not detected for phenotypic differences related to abdominal pigmentation, which seems to be related to habitat type and desiccation resistance.

Australian and Oriental regions

According to He et al. (2007), *Drosophila lacertosa* (*robusta* group) showed no correlation between geographical and genetic distances, suggesting a recent population expansion, putatively associated to post-glacial habitat changes. Otherwise, Kelemen and Moritz (1999) showed that populations of *D. serrata* and *D. birchii* (*melanogaster* group, *serrata* species complex) contracted during Pleistocene glacial periods. *Drosophila birchii* presented a recent range expansion, as suggested by the low diversity and the lack of geographical structure, a pattern later confirmed by Schiffer et al. (2007). *Drosophila serrata*, however, presented a phylogeographical break into two distinct lineages, which may represent a pair of cryptic species resulting from a split into two distinct refugia during the glacial period (Kelemen and Moritz, 1999).

Ethiopian region

Kopp et al. (2006) estimated the origin of *D. simulans* (*melanogaster* group, *simulans* species complex) in Madagascar or East Africa. Regarding a species located in another genus of Drosophilidae, *Zaprionus indianus*, Yassin et al. (2008) indicated subdivision in two lineages whose reproductive isolation supports independent species status, and whose phylogenetic patterns reinforces the African origin of the complex.

Conclusions

Despite the wide use of Drosophilidae species as model organisms in science, we unexpectedly recovered only a few phylogeographical studies with its species. The gap on the chosen taxa reinforces the demand on the less studied species or species groups. Likewise, there is an underrepresentation concerning some biogeographical regions, like the sub-Saharan Africa (Ethiopian region), which appears in much less articles than the Palearctic or Nearctic regions, for example. These gaps may hide answers to phylogeographical patterns or questions that are yet poorly understood or even totally unsolved, regarding not only the focused species, but the biodiversity of a region taken as a whole. In the first case, phylogeography can help, for example, to elucidate questions related to species of economic interest, as it was recently demonstrated for the pest and invasive species *Drosophila suzukii* (Choi et al., 2018; Fraimout et al., 2017). In the last case, Drosophilidae species can be used as a model to infer geographical barriers or historical events affecting the biodiversity of a region.

In fact, the obtained evidences support that several Drosophilidae species were highly affected by the climatic oscillations within the Pleistocene. Among the studied species, those presenting higher dependences on milder climate and humid habits [such as *D. americana*, *D. birchii*, *D. maculifrons*, *D. ornatifrons*, *D. serrata* (De Ré et al., 2014; Gustani et al., 2015; Kelemen and Moritz, 1999; Sillero et al., 2014)] present signals of bottlenecks or population contractions during glacial periods, followed by population expansions during interglacial periods. Nevertheless, species inhabiting xeric environments [such as the *buzzatii* species complex – *D. buzzatii* and *D. gouveai* (Brito et al., 2002; Moraes et al., 2009)] usually present contrasting signals, with range expansions dated to glacial periods. All these findings emphasizes the importance of using Drosophilidae species as model organisms, not only on the impact of past or future climate changes, but also on environmental perturbations, as recently shown by Valente-Gaiesky (2019).

Acknowledgments

The authors acknowledge Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Finance Code 001 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for providing scholarship and financial support.

Conflict of Interest

The authors declare that there is no conflict of interest that could be perceived as prejudicial to the impartiality of the reported research.

Authors Contributions

HRM designed and performed the study and wrote the first draft, LJR and MD supervised, reviewed and edited the final draft, all authors read and approved the final version.

References

Alfaro, F.M., Muñoz-Ramírez, C.P., Zúñiga-Reinoso, Á., Trewick, S.A., Méndez, M.A., 2018. Phylogeography of the Chilean red cricket Cratomelus armatus (Orthoptera: Anostostomatidae) reveals high cryptic diversity in central Chile. Biol J Linn Soc 123, 712–727. https://doi.org/10.1093/biolinnean/bly019

Asplen, M.K., Anfora, G., Biondi, A., Choi, D.-S., Chu, D., Daane, K.M., Gibert, P., Gutierrez, A.P., Hoelmer, K.A., Hutchison, W.D., Isaacs, R., Jiang, Z.-L., Kárpáti, Z., Kimura, M.T., Pascual, M., Philips, C.R., Plantamp, C., Ponti, L., Vétek, G., Vogt, H., Walton, V.M., Yu, Y., Zappalà, L., Desneux, N., 2015. Invasion biology of spotted wing Drosophila (Drosophila suzukii): a global perspective and future priorities. J Pest Sci 88, 469–494. https://doi.org/10.1007/s10340-015-0681-z

Avise, J.C., 2009. Phylogeography: retrospect and prospect. Journal of Biogeography 36, 3–15. https://doi.org/10.1111/j.1365-2699.2008.02032.x

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C., 1987. INTRASPECIFIC PHYLOGEOGRAPHY: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annu. Rev. Ecol. Syst. 18, 489–522. https://doi.org/10.1146/annurev.es.18.110187.002421

Bächli, G., 2020. TaxoDros: The Database on Taxonomy of Drosophilidae [WWW Document]. URL https://www.taxodros.uzh.ch/ (accessed 1.7.20).

Barrios-Leal, D.Y., Franco, F.F., Silva, E.C.C., Santos, C.K.B., Sene, F.M., Manfrin, M.H., 2018. Deep intraspecific divergence in Drosophila meridionalis, a cactophilic member of the New World Drosophila repleta group. Biol J Linn Soc 123, 163–178. https://doi.org/10.1093/biolinnean/blx134

Beheregaray, L.B., 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Mol. Ecol. 17, 3754–3774. https://doi.org/10.1111/j.1365-294X.2008.03857.x

Bonebrake, T.C., Ponisio, L.C., Boggs, C.L., Ehrlich, P.R., 2010. More than just indicators: A review of tropical butterfly ecology and conservation. Biological Conservation 143, 1831–1841. https://doi.org/10.1016/j.biocon.2010.04.044

Brehm, A., Harris, D.J., Hernández, M., Perez, J.A., Larruga, J.M., Pinto, F.M., González, A.M., 2004. Phylogeography of Drosophila subobscura from North Atlantic islands inferred from mtDNA A+T rich region sequences. Mol. Phylogenet. Evol. 30, 829–834. https://doi.org/10.1016/j.ympev.2003.10.018

Brisson, J.A., De Toni, D.C., Duncan, I., Templeton, A.R., 2005. Abdominal pigmentation variation in drosophila polymorpha: geographic variation in the trait, and underlying phylogeography. Evolution 59, 1046–1059.

Brito, R.A. de, Manfrin, M.H., Sene, F.M., 2002. Mitochondrial DNA phylogeography of Brazilian populations of Drosophila buzzatii. Genetics and Molecular Biology 25, 161–171. https://doi.org/10.1590/S1415-47572002000200009

Carson, H.L., 1971. The ecology of Drosophila breeding sites. University of Hawaii Foundation Lyon Arboretum Fund.

Choi, D.-S., Park, J.S., Kim, M.J., Kim, J.S., Jeong, S.Y., Jeong, J.S., Park, J., Kim, I., 2018. Geographic variation in the spotted-wing drosophila, Drosophila suzukii (Diptera: Drosophilidae), based on mitochondrial DNA sequences. Mitochondrial DNA A DNA Mapp Seq Anal 29, 312–322. https://doi.org/10.1080/24701394.2016.1278534

Cuthill, J.F.H., Charleston, M., 2015. Wing patterning genes and coevolution of Müllerian mimicry in Heliconius butterflies: Support from phylogeography, cophylogeny, and divergence times. Evolution 69, 3082–3096. https://doi.org/10.1111/evo.12812

da Silva, M.N., Patton, J.L., 1998. Molecular phylogeography and the evolution and conservation of Amazonian mammals. Mol. Ecol. 7, 475–486.

De Ré, F.C., Gustani, E.C., Oliveira, A.P.F., Machado, L.P.B., Mateus, R.P., Loreto, E.L.S., Robe, L.J., 2014. Brazilian populations of Drosophila maculifrons (Diptera: Drosophilidae): low diversity levels and signals of a population expansion after the Last Glacial Maximum. Biological Journal of the Linnean Society 112, 55–66. https://doi.org/10.1111/bij.12244

Deprá, M., Poppe, J.L., Schmitz, H.J., De Toni, D.C., Valente, V.L.S., 2014. The first records of the invasive pest Drosophila suzukii in the South American continent. J Pest Sci 87, 379–383. https://doi.org/10.1007/s10340-014-0591-5

Dool, S.E., O'Donnell, C.F.J., Monks, J.M., Puechmaille, S.J., Kerth, G., 2016. Phylogeographic-based conservation implications for the New Zealand long-tailed bat, (Chalinolobus tuberculatus): identification of a single ESU and a candidate population for genetic rescue. Conserv Genet 17, 1067–1079. https://doi.org/10.1007/s10592-016-0844-3

Eizirik, E., Kim, J.-H., Menotti-Raymond, M., Jr, P.G.C., O'Brien, S.J., Johnson, W.E., 2001. Phylogeography, population history and conservation genetics of jaguars (Panthera onca, Mammalia, Felidae). Molecular Ecology 10, 65–79. https://doi.org/10.1046/j.1365-294X.2001.01144.x

Fonseca, P.M., Loreto, E.L.S., Gottschalk, M.S., Robe, L.J., 2017. Cryptic diversity and speciation in the Zygothrica genus group (Diptera, Drosophilidae): the case of Z. vittimaculosa Wiedemann. Insect Systematics & Evolution 48, 285–313. https://doi.org/10.1163/1876312X-48022158

Fraimout, A., Debat, V., Fellous, S., Hufbauer, R.A., Foucaud, J., Pudlo, P., Marin, J.-M., Price, D.K., Cattel, J., Chen, X., Deprá, M., François Duyck, P., Guedot, C., Kenis, M., Kimura, M.T., Loeb, G., Loiseau, A., Martinez-Sañudo, I., Pascual, M., Polihronakis Richmond, M., Shearer, P., Singh, N., Tamura, K., Xuéreb, A., Zhang, J., Estoup, A., 2017. Deciphering the Routes of invasion of Drosophila suzukii by Means of ABC Random Forest. Mol. Biol. Evol. 34, 980–996. https://doi.org/10.1093/molbev/msx050

Franco, F.F., Manfrin, M.H., 2013. Recent demographic history of cactophilic Drosophila species can be related to Quaternary palaeoclimatic changes in South America. Journal of Biogeography 40, 142–154. https://doi.org/10.1111/j.1365-2699.2012.02777.x

Gao, J., Hu, Y., Toda, M.J., Katoh, T., Tamura, K., 2011. Phylogenetic relationships between Sophophora and Lordiphosa, with proposition of a hypothesis on the vicariant divergences of tropical lineages between the Old and New Worlds in the family Drosophilidae. Molecular Phylogenetics and Evolution 60, 98–107. https://doi.org/10.1016/j.ympev.2011.04.012

Garcez, D.K., Barbosa, C., Loureiro, M., Volcan, M.V., Loebmann, D., Quintela, F.M., Robe, L.J., 2018. Phylogeography of the critically endangered neotropical annual fish, Austrolebias wolterstorffi (Cyprinodontiformes: Aplocheilidae): genetic and morphometric evidence of a new species complex. Environ Biol Fish 101, 1503–1515. https://doi.org/10.1007/s10641-018-0795-2

Goodhue, R.E., Bolda, M., Farnsworth, D., Williams, J.C., Zalom, F.G., 2011. Spotted wing drosophila infestation of California strawberries and raspberries: economic analysis of potential revenue losses and control costs. Pest Manag. Sci. 67, 1396–1402. https://doi.org/10.1002/ps.2259

Grimaldi, D.A., 1990. A phylogenetic, revised classification of genera in the Drosophilidae (Diptera). Bull Am Mus Nat Hist 197, 1–139.

Grimaldi, D.A., 1987. Phylogenetics and taxonomy of Zygothrica (Diptera: Drosophilidae). Bull Am Mus Nat Hist 186, 103–268.

Gum, B., Gross, R., Kuehn, R., 2005. Mitochondrial and nuclear DNA phylogeography of European grayling (Thymallus thymallus): evidence for secondary contact zones in central Europe. Molecular Ecology 14, 1707–1725. https://doi.org/10.1111/j.1365-294X.2005.02520.x

Gustani, E.C., Oliveira, A.P.F., Santos, M.H., Machado, L.P.B., Mateus, R.P., 2015. Demographic Structure and Evolutionary History of Drosophila ornatifrons (Diptera, Drosophilidae) from Atlantic Forest of Southern Brazil. Zool. Sci. 32, 141–150. https://doi.org/10.2108/zs140062

He, L., Watabe, H., Xiangyu, J., Gao, J., Liang, X., Aotsuka, T., Zhang, Y., 2007. Genetic differentiation and cryptic speciation in natural populations of Drosophila lacertosa. Mol. Phylogenet. Evol. 43, 24–31. https://doi.org/10.1016/j.ympev.2006.08.010

Hewitt, G.M., 2004. The structure of biodiversity – insights from molecular phylogeography. Frontiers in Zoology 1, 4. https://doi.org/10.1186/1742-9994-1-4

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc 58, 247–276. https://doi.org/10.1111/j.1095-8312.1996.tb01434.x

Hodel, R.G.J., Cortez, M.B. de S., Soltis, P.S., Soltis, D.E., 2016. Comparative phylogeography of black mangroves (Avicennia germinans) and red mangroves (Rhizophora mangle) in Florida: Testing the maritime discontinuity in coastal plants. Am. J. Bot. 103, 730–739. https://doi.org/10.3732/ajb.1500260

Izumitani, H.F., Kusaka, Y., Koshikawa, S., Toda, M.J., Katoh, T., 2016. Phylogeography of the Subgenus Drosophila (Diptera: Drosophilidae): Evolutionary History of Faunal Divergence between the Old and the New Worlds. PLoS ONE 11, e0160051. https://doi.org/10.1371/journal.pone.0160051

Jeratthitikul, E., Hara, T., Yago, M., Itoh, T., Wang, M., Usami, S., Hikida, T., 2013. Phylogeography of Fischer's blue, Tongeia fischeri, in Japan: Evidence for introgressive hybridization. Molecular Phylogenetics and Evolution 66, 316–326. https://doi.org/10.1016/j.ympev.2012.10.009

Kelemen, L., Moritz, C., 1999. Comparative Phylogeography of a Sibling Pair of Rainforest Drosophila Species (drosophila Serrata and D. Birchii). Evolution 53, 1306–1311. https://doi.org/10.1111/j.1558-5646.1999.tb04545.x

Kohler, R.E., 1993. Drosophila: A life in the laboratory. J Hist Biol 26, 281–310. https://doi.org/10.1007/BF01061971

Kopp, A., Frank, A., Fu, J., 2006. Historical biogeography of Drosophila simulans based on Y-chromosomal sequences. Mol. Phylogenet. Evol. 38, 355–362. https://doi.org/10.1016/j.ympev.2005.06.006

Markow, T.A., 2019. Ecological and Evolutionary Genomics: The Cactophilic Drosophila Model System. J Hered 110, 1–3. https://doi.org/10.1093/jhered/esy062

Markow, T.A., O'Grady, P.M., 2006. Phylogenetic relationships of Drosophilidae, in: Drosophila: A Guide to Species Identification and Use. Academic Press, Cambridge, MA, pp. 3–64.

Mirol, P.M., Routtu, J., Hoikkala, A., Butlin, R.K., 2008. Signals of demographic expansion in Drosophila virilis. BMC Evolutionary Biology 8, 59. https://doi.org/10.1186/1471-2148-8-59

Mirol, P.M., Schäfer, M.A., Orsini, L., Routtu, J., Schlötterer, C., Hoikkala, A., Butlin, R.K., 2007. Phylogeographic patterns in Drosophila montana. Mol. Ecol. 16, 1085–1097. https://doi.org/10.1111/j.1365-294X.2006.03215.x

Moraes, E.M., Yotoko, K.S.C., Manfrin, M.H., Solferini, V.N., Sene, F.M., 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. https://doi.org/10.1111/j.1365-2699.2009.02145.x

Morrone, J.J., 2015. Biogeographical regionalisation of the world: a reappraisal. Aust. Systematic Bot. 28, 81–90. https://doi.org/10.1071/SB14042

Novković, B., Kimura, M.T., 2015. Phylogeography, Interaction Patterns and the Evolution of Host Choice in Drosophila-Parasitoid Systems in Ryukyu Archipelago and Taiwan. PLoS ONE 10, e0129132. https://doi.org/10.1371/journal.pone.0129132

Powell, J.R., 1997. Progress and prospects in evolutionary biology: The Drosophila model. Oxford University Press, New York, NY.

Robe, L.J., De Ré, F.C., Ludwig, A., Loreto, E.L.S., 2013. The Drosophila flavopilosa species group (Diptera, Drosophilidae): an array of exciting questions. Fly 7, 59–69. https://doi.org/10.4161/fly.23923

Robert Liu, F.-G., Tsaur, S.-C., Huang, H.-T., 2015. Biogeography of Drosophila (Diptera: Drosophilidae) in East and Southeast Asia. J Insect Sci 15. https://doi.org/10.1093/jisesa/iev056

Rosauer, D.F., Blom, M.P.K., Bourke, G., Catalano, S., Donnellan, S., Gillespie, G., Mulder, E., Oliver, P.M., Potter, S., Pratt, R.C., Rabosky, D.L., Skipwith, P.L., Moritz, C., 2016. Phylogeography, hotspots and conservation priorities: an example from the Top End of Australia. Biological Conservation, Advancing reptile conservation: Addressing knowledge gaps and mitigating key drivers of extinction risk 204, 83–93. https://doi.org/10.1016/j.biocon.2016.05.002

RStudio Team, 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.

Schiffer, M., Kennington, W.J., Hoffmann, A.A., Blacket, M.J., 2007. Lack of genetic structure among ecologically adapted populations of an Australian rainforest Drosophila species as indicated by microsatellite markers and mitochondrial DNA sequences. Mol. Ecol. 16, 1687–1700. https://doi.org/10.1111/j.1365-294X.2006.03200.x

Siddall, M., Chappell, J., Potter, E.-K., 2007. Eustatic sea level during past interglacials, in: Sirocko, F., Claussen, M., Sánchez Goñi, M.F., Litt, T. (Eds.), Developments in Quaternary Sciences, The Climate of Past Interglacials. Elsevier, pp. 75–92. https://doi.org/10.1016/S1571-0866(07)80032-7

Sillero, N., Reis, M., Vieira, C.P., Vieira, J., Morales-Hojas, R., 2014. Niche evolution and thermal adaptation in the temperate species Drosophila americana. J. Evol. Biol. 27, 1549–1561. https://doi.org/10.1111/jeb.12400

Smith, G., Lohse, K., Etges, W.J., Ritchie, M.G., 2012. Model-based comparisons of phylogeographic scenarios resolve the intraspecific divergence of cactophilic Drosophila mojavensis. Mol. Ecol. 21, 3293–3307. https://doi.org/10.1111/j.1365-294X.2012.05604.x

Song, W., Cao, L.-J., Li, B.-Y., Gong, Y.-J., Hoffmann, A.A., Wei, S.-J., 2018. Multiple refugia from penultimate glaciations in East Asia demonstrated by phylogeography and ecological modelling of an insect pest. BMC Evolutionary Biology 18, 152. https://doi.org/10.1186/s12862-018-1269-z

Tembrock, L.R., Simmons, M.P., Richards, C.M., Reeves, P.A., Reilley, A., Curto, M.A., Meimberg, H., Ngugi, G., Demissew, S., Al Khulaidi, A.W., Al-Thobhani, M., Simpson, S., Varisco, D.M., 2017. Phylogeography of the wild and cultivated stimulant plant qat (Catha edulis, Celastraceae) in areas of historical cultivation. Am. J. Bot. 104, 538–549. https://doi.org/10.3732/ajb.1600437

Terranova, M.S., Brutto, S.L., Arculeo, M., Mitton, J.B., 2007. A mitochondrial phylogeography of Brachidontes variabilis (Bivalvia: Mytilidae) reveals three cryptic species. Journal of Zoological Systematics and Evolutionary Research 45, 289–298. https://doi.org/10.1111/j.1439-0469.2007.00421.x

Throckmorton, L.H., 1975. The Phylogeny, Ecology, and Geography of Drosophila, in: Handbook of Genetics. Plenum Publishing, New York, NY, pp. 421–469.

Tosi, D., Martins, M., Vilela, C.R., Pereira, M.A.Q.R., 1990. On a new cave-dwelling species of bat-guano-breeding drosophila closely related to D. repleta Wollaston (Diptera, Drosophilidae). Brazilian Journal of Genetics 13, 19–31.

Valente-Gaiesky, V.L. da S., 2019. Can insect assemblages tell us something about the urban environment health? Anais da Academia Brasileira de Ciências 91. https://doi.org/10.1590/0001-3765201920190445

Walker, J.D., Geissman, J.W., Bowring, S.A., Babcock, L.E., 2018. The Geological Society of America Geologic Time Scale v. 5.0. Geological Society of America. https://doi.org/10.1130/2018.CTS005R3C

Wilder, J.A., Hollocher, H., 2003. Recent radiation of endemic Caribbean Drosophila of the dunni subgroup inferred from multilocus DNA sequence variation. Evolution 57, 2566–2579.

Yassin, A., Capy, P., Madi-Ravazzi, L., Ogereau, D., David, J.R., 2008. DNA barcode discovers two cryptic species and two geographical radiations in the invasive drosophilid Zaprionus indianus. Mol Ecol Resour 8, 491–501. https://doi.org/10.1111/j.1471-8286.2007.02020.x

Zhang, L.-J., Cai, W.-Z., Luo, J.-Y., Zhang, S., Wang, C.-Y., Lv, L.-M., Zhu, X.-Z., Wang, L., Cui, J.-J., 2017. Phylogeographic patterns of Lygus pratensis (Hemiptera: Miridae): Evidence for weak genetic structure and recent expansion in northwest China. PLOS ONE 12, e0174712. https://doi.org/10.1371/journal.pone.0174712

Internet resources

Journal Citation Reports (Clarivate Analytics), https://jcr.clarivate.com/ (January 3, 2019)

Scopus (Elsevier), http://www.scopus.com/ (January 3, 2019)

Web of Science (Clarivate Analytics), http://www.webofknowledge.com/ (January 3, 2019)

Figures

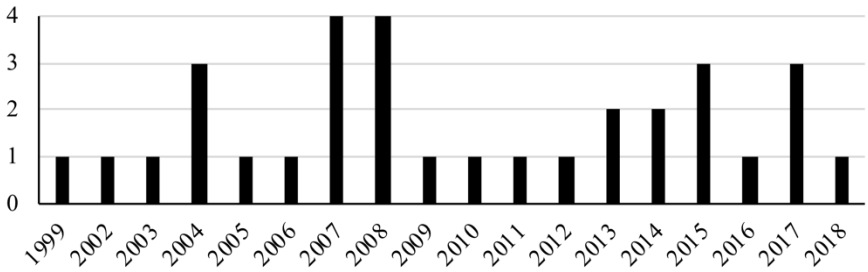


Figure 1. Number of phylogeographic studies with drosophilids published during the period of 1999-2018.

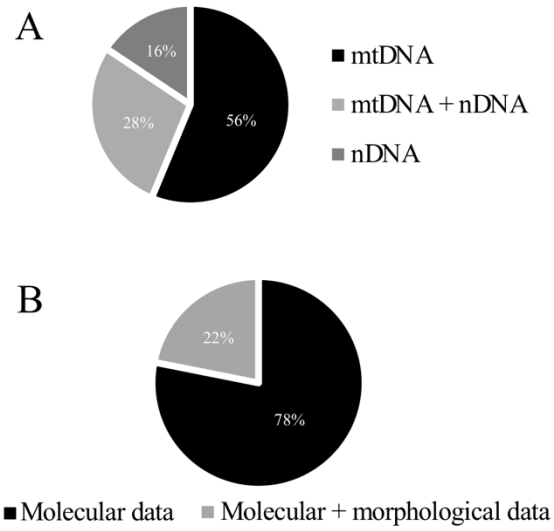


Figure 2. Proportion of phylogeographic studies with drosophilids published during 1999-2018 sorted by (A) Type of molecular marker. (B) Inclusion or not of morphological data. mtDNA, mitochondrial DNA; nDNA, nuclear DNA.

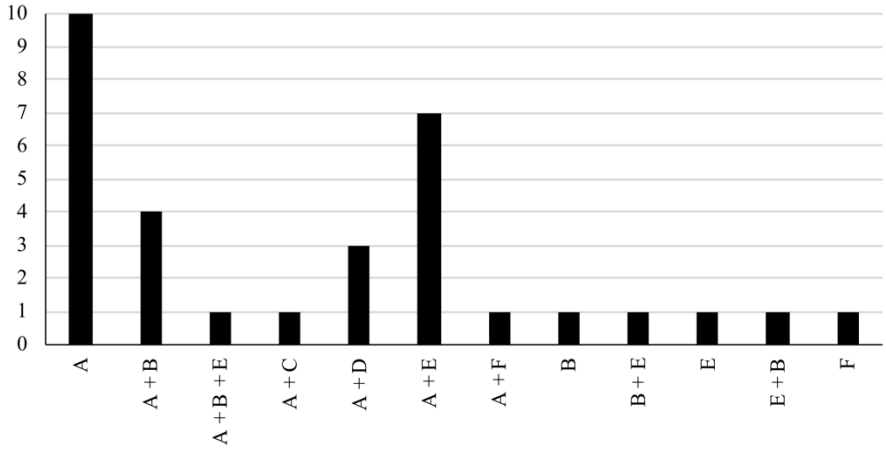


Figure 3. Type of questions or phylogeographical hypothesis addressed by the evaluated papers. (A) Inference of inter- or/and intrapopulation variability levels, population structure and demographic history; (B) inference of evolutionary factors (selection, migration, mutation, genetic drift); (C) inference of introgression; (D) inference of cryptic diversity (species or lineages); (E) inference of historical processes affecting other biodiversity scales (vicariance due to a river, strait, refugia, etc); (F) co-evolution.

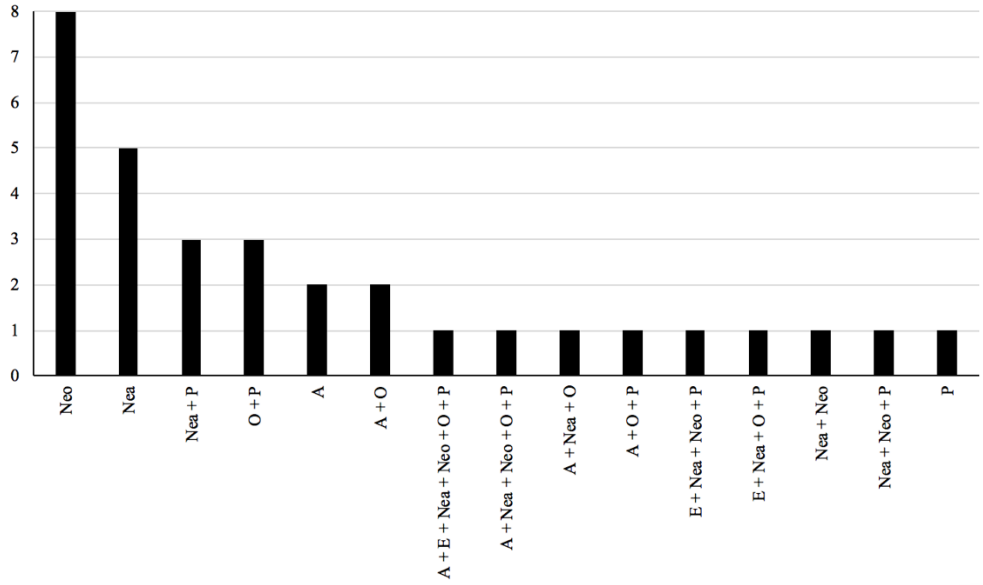


Figure 4. Number of phylogeographic studies with drosophilids published during 1999-2018 sorted by biogeographical regions. A = Australian; E = Ethiopian; Nea = Nearctic; Neo = Neotropical; O = Oceanian; P = Palearctic.

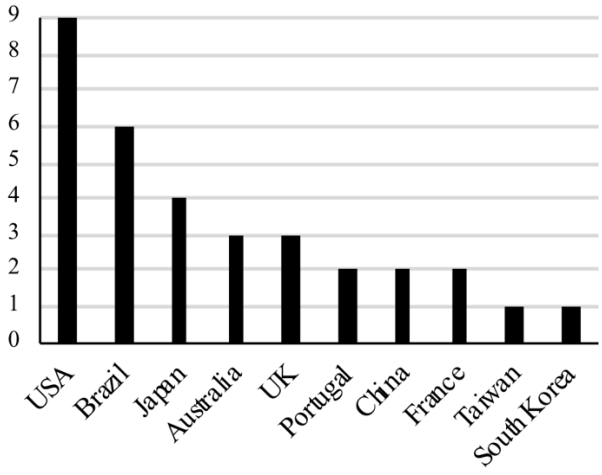


Figure 5. Number phylogeographic studies with drosophilids published by each country during the period of 1999-2018.

Tables

Table 1. List of Journals that published phylogeographic studies with Drosophilidae during the period of 1999-2018, with their respective Impact Factor (IF) and absolute and relative number of publications.

	Journal title	Number of publications	IF	Percentage
1	Molecular Ecology	7	6.131	21.88%
2	Molecular Phylogenetics and Evolution	5	4.412	15.63%
3	Evolution	3	3.818	9.38%
4	Biological Journal of the Linnean Society	3	2.532	9.38%
5	Molecular Biology and Evolution	2	10,217	6.25%
6	Journal of Biogeography	2	4.154	6.25%
7	Zoological Science	2	0.906	6.25%
8	PLOS ONE	2	2.766	6.25%
9	Genetics and Molecular Biology	1	1.493	3.13%
10	BMC Evolutionary Biology	1	3.027	3.13%
11	Molecular Ecology Resources	1	7.059	3.13%
12	Journal of Evolutionary Biology	1	2.538	3.13%
13	Journal of Insect Science	1	1.324	3.13%
14	Mitochondrial DNA Part A	1	0.575	3.13%

Table 2. List of genera and species groups of Drosophilidae included in the evaluated phylogeographic studies, with their respective number of articles (for more details, see Table S1).

Genus	Subgenus	species group	Number of articles
Drosophila	*Drosophila*	*repleta*	9
		immigrans	4
		virilis	4
		cardini	3
		guarani	3
		funebris	2
		bizonata	1
		guttifera	1
		histrio	1
		nannoptera	1
		polychaeta	1
		quinaria	1
		robusta	1
		testacea	1
		tripunctata	1
	Sophophora	*melanogaster*	12
		obscura	2
		saltans	1
		willistoni	1
Idiomyia		*orphnopeza*	1
Lordiphosa		*denticeps*	1
		fenestrarum	1
		miki	1
		nigricolor	1
Scaptodrosophila		*rufifrons*	1
Scaptomyza			1
Zaprionus	*Zaprionus*	*vittiger*	1

Table S1. List of the 32 papers retrieved in our search, with their respective assorting about the evaluation criteria.

Publication #	Authors	Year	Journal	DOI	Country	Institution
1	Kelemen & Moritz	1999	Evolution	10.1111/j.1558-5646.1999.tb04545.x	Australia	University of Queesland, Australia
2	Brito et al.	2002	Genetics and Molecular Biology	10.1590/S1415-47572002000200009	USA	Washington University, USA
3	Wilder & Hollocher	2003	Evolution	10.1111/j.0014-3820.2003.tb01500.x	USA	Princeton University, USA
4	Ballard	2004	Molecular Biology and Evolution	10.1093/molbev/msh028	USA	University of Iowa, USA
5	Brehm et al.	2004	Molecular Phylogenetics and Evolution	10.1016/j.ympev.2003.10.018	Portugal	University of Madeira, Portugal
6	Hurtado et al.	2004	Molecular Ecology	10.1111/j.1365-294X.2004.02169.x	USA	University of Arizona, USA
7	Brisson et al.	2005	Evolution	10.1111/j.0014-3820.2005.tb01043.x	USA	Washington University, USA
8	Kopp et al.	2006	Molecular Phylogenetics and Evolution	10.1016/j.ympev.2005.06.006	USA	University of California, USA
9	Mirol et al.	2007	Molecular Ecology	10.1111/j.1365-294X.2006.03215.x	UK	The University of Leeds, UK
10	Reed et al.	2007	Molecular Ecology	10.1111/j.1365-294X.2006.02941.x	USA	University of Arizona, USA
11	Schiffer et al.	2007	Molecular Ecology	10.1111/j.1365-294X.2006.03200.x	Australia	University of Melbourne, Australia
12	He	2007	Molecular Phylogenetics and Evolution	10.1016/j.ympev.2006.08.010	China	Kunming Institute of Zoology, China
13	Mirol et al.	2008	BMC Evolutionary Biology	10.1186/1471-2148-8-59	UK	The University of Leeds, UK
14	Sawamura et al.	2008	Molecular Phylogenetics and Evolution	10.1016/j.ympev.2008.06.007	Japan	University of Tsukuba, Japan
15	Schug	2008	Molecular Ecology	10.1111/j.1365-294X.2008.03770.x	USA	University of North Carolina, USA
16	Yassin et al.	2008	Molecular Ecology Resources	10.1111/j.1471-8286.2007.02020.x	France	Centre National de la Recherche Scientifique, France
17	Moraes et al.	2009	Journal of Biogeography	10.1111/j.1365-2699.2009.02145.x	Brazil	Universidade Federal de São Carlos, Brazil
18	Sawamura et al.	2010	Zoological Science	10.2108/zsj.27.303	Japan	University of Tsukuba, Japan
19	Gao et al.	2011	Molecular Phylogenetics and Evolution	10.1016/j.ympev.2011.04.012	China	Yunnan University, China
20	Smith et al.	2012	Molecular Ecology	10.1111/j.1365-294X.2012.05604.x	UK	University of St. Andrews, UK
21	Franco & Manfrin	2013	Journal of Biogeography	10.1111/j.1365-2699.2012.02777.x	Brazil	Universidade Federal de São Carlos, Brazil
22	Eldon et al.	2013	Molecular Ecology	10.1111/mec.12326	USA	University of Hawaii, USA
23	De Ré et al.	2014	Biological Journal of the Linnean Society	10.1111/bij.12244	Brazil	Universidade Federal de Santa Maria, Brazil
24	Sillero et al.	2014	Journal of Evolutionary Biology	10.1111/jeb.12400	Portugal	Centro de Investigação em Ciências Geo-Espaciais, Portugal
25	Liu et al.	2015	Journal of Insect Science	10.1093/jisesa/iev056	Taiwan	National Central University, Taiwan
26	Novkovic & Kamura	2015	PLOS ONE	10.1371/journal. pone.0129132	Japan	Hokkaido University, Japan
27	Gustani	2015	Zoological Science	10.2108/zs140062	Brazil	Universidadade Estadual do Centro-Oeste, Brazil
28	Izumitani et al.	2016	PLOS ONE	10.1371/journal.pone.0160051	Japan	Hokkaido University, Japan
29	Rampasso et al.	2017	Biological Journal of the Linnean Society	10.1093/biolinnean/blx073	Brazil	Universidade de São Paulo, Brazil
30	Fraimout	2017	Molecular Biology and Evolution	10.1093/molbev/msx050	France	Sorbonne Universités, France
31	Choi	2017	Mitochondrial DNA Part A	10.1080/24701394.2016.1278534	South Korea	Animal and Plant Quarantine Agency, South Korea
32	Barrios-Leal et al.	2018	Biological Journal of the Linnean Society	10.1093/biolinnean/blx134	Brazil	Universidade de São Paulo, Brazil

Table S1. Continued.

Publication #	N sp	sp groups	N sp groups	Biogeographical regions
1	2	melanogaster	1	Australian
2	1	repleta	1	Neotropical
3	9	cardini	1	Neotropical
4	1	melanogaster	1	Ethiopian; Neartic; Neotropical; Paleartic; Oriental; Australian
5	1	obscura	1	Paleartic
6	3	nannoptera; repleta	2	Neartic
7	1	cardini	1	Neotropical
8	1	melanogaster	1	Oriental; Paleartic; Neartic; Australian
9	1	virilis	1	Neartic; Paleartic
10	4	repleta	1	Neartic
11	1	melanogaster	1	Australian
12	1	robusta	1	Oriental; Paleartic
13	1	virilis	1	Paleartic; Neartic
14	6	melanogaster	1	Australian; Oriental
15	2	melanogaster	1	Paleartic; Oriental; Australian
16	1	vittiger	1	Neotropical; Paleartic; Ethiopian; Neartic
17	1	repleta	1	Neotropical
18	6	melanogaster	1	Australian; Oriental
19	27	melanogaster; obscura; saltans; willistoni; repleta; immigrans; funebris; polychaeta; virilis; fenestrarum; miki; nigricolor; denticeps; rufifrons	13	Oriental; Ethiopian; Neartic; Paleartic
20	1	repleta	1	Neartic
21	4	repleta	1	Neotropical
22	1	grimshawi	1	Neartic
23	1	guarani	1	Neotropical
24	1	virilis	1	Neartic
25	5	immigrans; melanogaster	2	Oriental; Paleartic
26	3	immigrans; melanogaster	2	Paleartic; Oriental
27	1	guarani	1	Neotropical
28	45	bizonata; cardini; funebris; guarani; guttifera; histrio; immigrans; quinaria; testacea; tripunctata	10	Neartic; Neotropical; Paleartic; Oriental; Australian
29	1	repleta	1	Neartic; Neotropical
30	1	melanogaster	1	Neartic; Neotropical; Paleartic
31	1	melanogaster	1	Neartic; Paleartic
32	1	repleta	1	Neotropical

Table S1. Continued.

Publication #	N markers	Type of markers	mtDNA marker	nDNA marker	Morphological data	Molecular clock	Period
1	1	mtDNA	COI	--	No	No	--
2	1	mtDNA	COI	--	No	No	--
3	4	mtDNA; nDNA	COII	sod; en; kl-5	No	Yes	Pliocene
4		mtDNA	Mitogenome	--	No	No	--
5	1	mtDNA	COI	--	No	Yes	Pleistocene
6	1	mtDNA	COI	--	No	No	--
7	2	mtDNA; nDNA	cytB	pgd	Yes	No	--
8	2	nDNA	--	kl-2; kl-3	No	No	--
9	18	mtDNA; nDNA	COI; COII	microsatellites	Yes	Yes	Pleistocene
10	1	mtDNA	COI	--	No	Yes	Pliocene
11	11	mtDNA; nDNA	ND5	microsatellites	No	No	--
12	1	mtDNA	ND2	--	No	No	--
13	2	mtDNA	COI; COII	--	No	Yes	Pleistocene
14	1	mtDNA	pseudo-COI	--	No	No	--
15	2	mtDNA	cytB; ssA	--	Yes	No	--
16	2	mtDNA	COI; COII	--	Yes	No	--
17	1	mtDNA	COII	--	No	No	--
18	4	nDNA	--	Actn; white; CG7785; ZnT63C	No	No	--
19	4	mtDNA; nDNA	ND2; COII	Adh; 28SrRNA	No	No	--
20	15	nDNA	--	multiple loci	No	Yes	Pleistocene
21	1	mtDNA	COI	--	No	Yes	Pleistocene
22	1	mtDNA	COI	--	No	No	--
23	2	mtDNA	COI; COII	--	No	No	--
24	16	nDNA	--	multiple loci	No	No	--
25	8	mtDNA; nDNA	ND4; ND4L; tRNA-His; tRNA-Pro; tRNA-Thr; ND5; ND6	esc	No	Yes	Pleistocene
26	2	mtDNA; nDNA	COI	gpdh	Yes	No	--
27	2	mtDNA	COI; COII	--	No	No	--
28	5	mtDNA; nDNA	COI; COII	gpdh; Adh; 28S	No	Yes	--
29	1	mtDNA	COI	--	Yes	No	--
30	25	nDNA	--	microsatellites	No	No	--
31	2	mtDNA	COI; ND4	--	No	No	--
32	2	mtDNA; nDNA	COI	period	Yes	Yes	Pleistocene

Table S2. List of the Institutions of origin of the first authors of the evaluated papers.

Institution	Number of papers
Hokkaido University, Japan	2
Washington University, USA	2
University of Arizona, USA	2
The University of Leeds, UK	2
University of Tsukuba, Japan	2
Universidade Federal de São Carlos, Brazil	2
Universidade de São Paulo, Brazil	2
Princeton University, USA	1
University of Iowa, USA	1
University of Madeira, Portugal	1
University of California, USA	1
University of Melbourne, Australia	1
Kunming Institute of Zoology, China	1
University of North Carolina, USA	1
Centre National de la Recherche Scientifique, France	1
Yunnan University, China	1
University of St. Andrews, UK	1
University of Hawaii, USA	1
Universidade Federal de Santa Maria, Brazil	1
Centro de Investigação em Ciências Geo-Espaciais, Portugal	1
National Central University, Taiwan	1
Universidade Estadual do Centro-Oeste, Brazil	1
Sorbonne Universités, France	1
Animal and Plant Quarantine Agency, South Korea	1
University of Queesland, Australia	1