UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIODIVERSIDADE ANIMAL

ASPECTOS TAXONÔMICOS, ECOLÓGICOS E EVOLUTIVOS EM DROSOFILÍDEOS NEOTROPICAIS

TESE DE DOUTORADO

Stela Machado

Santa Maria, RS, Brasil

2015

ASPECTOS TAXONÔMICOS, ECOLÓGICOS E EVOLUTIVOS EM DROSOFILÍDEOS NEOTROPICAIS

Stela Machado

Tese apresentada ao Curso de Doutorado do Programa de Pós-graduação em Biodiversidade Animal, área de concentração em Sistemática e Biologia Evolutiva, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Doutor em Biodiversidade Animal**.

Orientadora: Dra. Lizandra Jaqueline Robe

Co-orientador: Dr. Marco Silva Gottschalk

Santa Maria, RS, Brasil

2015

Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Programa de Pós-graduação em Biodiversidade Animal

A comissão examinadora, abaixo assinada,
aprova a Tese de Doutorado

ASPECTOS TAXONÔMICOS, ECOLÓGICOS E EVOLUTIVOS EM DROSOFILÍDEOS NEOTROPICAIS

elaborada por
Stela Machado

como requisito parcial para obtenção do grau de
Doutor em Biodiversidade Animal

COMISSÃO EXAMINADORA:

Lizandra Jaqueline Robe, Dra.
(Presidente/Orientador)

Hermes José Schmitz, Dr. (UNILA)

Juliana Cordeiro, Dra. (UFPEL)

Vera Lúcia da Silva Valente Gaiesky, Dra. (UFRGS)

Daniel Ângelo Sganzerla Graichen, Dr. (UFSM)

Santa Maria, 27 de maio de 2015.

Este trabalho é dedicado ao meu marido, à minha família e a todos aqueles que sempre me apoiaram ao longo desta jornada.

AGRADECIMENTOS

Primeiramente, gostaria de agradecer a minha orientadora e grande amiga Lizandra J. Robe. Muito obrigada pela excelente orientação, pelos inúmeros ensinamentos e por me proporcionar essa experiência que, apesar de seus percalços e momentos desesperadores, me ajudou a crescer como profissional e como pessoa. Sei que sempre poderei contar com teus conselhos e ajuda no decorrer de minha jornada profissional e de tua amizade na vida pessoal. Gostaria de agradecer ao meu co-orientador, Dr. Marco S. Gottschalk, pelos ensinamentos e amizade e ao Dr. Elgion L.S. Loreto pelo apoio, paciência e colaboração. Também gostaria de agradecer à minha orientadora da iniciação científica, Dra. Marlise L. B. Santos, e ao meu orientador do mestrado, Dr. Sandro L. Bonatto, pelos ensinamentos que muito contribuíram na minha formação.

Ao meu marido, Jardel, agradeço pelo amor incondicional, apoio, carinho e amizade. Sem você seria impossível finalizar esse trabalho; foi você que me manteve forte e não me deixou desanimar quando tudo dava errado e desistir parecia a melhor opção. Obrigada também por sempre estar disponível para me escutar e discutir assuntos que você nem compreendia, corrigir meus textos e dar uns “puxões de orelha” quando necessário. E agora mais que nunca, obrigada por estar ao meu lado curtindo esse momento tão especial da chegada da nossa princesinha Helena. Amo muito você, meu amor.

À minha família, meus pais, João Pedro e Gislaine, e minhas irmãs, Scheila e Simone, pelo amor, carinho e incentivo ao estudo desde de sempre. Obrigada por sempre tentarem compreender o que eu estava fazendo, mesmo que difícil na maioria das vezes, e por se preocuparem por que as coisas não davam certo e como poderiam ajudar. Obrigada pela ajuda nas coletas, afinal família participativa é tudo. Amo muito todos vocês.

Agradeço aos colegas do Labdros e Labeem pelos diversos ensinamentos, pela convivência e pela amizade. Agradeço em especial ao João Pedro J. dos Santos pelos momentos de aprendizado tentando me fazer ver nos desenhos de descrição o que estávamos vendo nas lâminas e tentando me convencer porque uma foto não era a melhor opção, pelas intermináveis discussões e companheirismo

de coleta. A Francine C. de Ré pelas incontáveis ajudas e apoio, ao Pedro M. Fonseca pelo companheirismo nas viagens de coleta, principalmente nos momentos de teimosia com o GPS. Agradeço também aos técnicos dos laboratórios Bruna e Diego, assim como ao biólogo Ronaldo.

Agradeço também à toda equipe das diversas reservas pela receptividade e auxílio, sem o qual não seria possível realizar este trabalho.

Além de todas estas e outras pessoas que me apoiaram e que de alguma forma contribuíram com o meu trabalho, gostaria de agradecer à Universidade Federal de Santa Maria, à Universidade Federal de Rio Grande e ao Programa de Pós-graduação em Biodiversidade Animal pela oportunidade de desenvolver este trabalho.

À CAPES e ao CNPq, instituições que contribuíram financeiramente para a realização deste trabalho.

“Seen in the light of evolution, biology is, perhaps, intellectually the most satisfying and inspiring science. Without that light it becomes a pile of sundry facts-some of them interesting or curious but making no meaningful picture as a whole”.

Theodosius Dobzhansky
(Am. Biol. Teach, March 1973)

RESUMO

Tese de Doutorado
Programa de Pós-Graduação em Biodiversidade Animal
Universidade Federal de Santa Maria, RS, Brasil

ASPECTOS TAXONÔMICOS, ECOLÓGICOS E EVOLUTIVOS EM DROSOFILÍDEOS NEOTROPICAIS

Autora: Stela Machado
Orientadora: Lizandra Jaqueline Robe
Co-orientador: Marco Silva Gosttschalk
Local e data da Defesa: Santa Maria, 27 de maio de 2015.

Drosophila melanogaster é um conhecido organismo modelo que tem incitado estudos em diversos campos do conhecimento e auxiliado no avanço científico em várias áreas. Contudo, Drosophilidae, a família na qual essa espécie está inserida, é extremamente especiosa, sendo composta por mais de 4.200 espécies, a maioria das quais com conhecimento deficitário ou inexistente, além de muitas outras espécies desconhecidas. Mesmo hipóteses acerca da evolução do grupo formuladas em 1975 não foram ainda devidamente testadas. Dessa forma, a presente tese pretende preencher parte dessas lacunas testando a hipótese das radiações proposta por THROCKMORTON (1975) em diferentes grupos taxonômicos de drosofilídeos, a fim de definir se os mesmos evoluíram através de radiações adaptativas ou são meramente grupos taxonômicos de difícil resolução filogenética. Para isso, no capítulo 1, foram utilizadas espécies pertencentes às linhagens *virilis-repleta* e *tripunctata* do subgênero *Drosophila*, e os resultados mostram que a evolução do grupo parece ser mais complexa do que previamente proposto, com a primeira linhagem apresentando sinais de conservadorismo de nicho e a segunda sendo compatível com a ocorrência de uma radiação adaptativa. No capítulo 2 também foram encontrados sinais de conservação de nicho abiótico entre diferentes espécies micofílicas pertencentes ao grupo genérico *Zygothrica*, que utilizam corpos de frutificação de fungos macroscópicos como recurso para alimentação, oviposição e/ou cópula. Em outra frente, também no capítulo 2, visando contribuir para o conhecimento da diversidade da família, expandimos o número de registros e modelamos a distribuição potencial de várias espécies de *Hirtodrosophila*, *Paraliodrosophila*, *Mycodrosophila* e *Zygothrica* registradas no Brasil, e mostramos através da modelagem de nicho que algumas espécies desses gêneros parecem ser suscetíveis às mudanças climáticas globais devido aos fatores abióticos limitantes de sua distribuição e às curvas de resposta apresentadas por cada espécie. Já no capítulo 3, exploramos a diversidade de espécies do grupo genérico *Zygothrica* e a efetividade da metodologia do DNA *barcode* na exploração taxonômica do mesmo. Neste caso, demonstramos a presença de uma grande diversidade críptica de representantes do grupo genérico *Zygothrica* nos Neotrópicos, com registro de pelo menos 24 potenciais espécies novas. Além disso, também demonstramos o enorme potencial da tecnologia de DNA *barcode* na identificação e descoberta destas espécies. No capítulo 4 exploramos o gênero *Mycodrosophila*, sobretudo o complexo de *M. projectans*, que se mostrou um interessante grupo para o estudo de especiação, devido à presença frequente de casos de diversidade críptica em simpatria e sintopia. Esses trabalhos ressaltam o quanto alguns aspectos e grupos interessantes da família têm sido negligenciados, e chamam a atenção para o fato de que estes aspectos são fundamentais para a compreensão da taxonomia, da ecologia e da evolução não apenas dos grupos em foco, mas da própria espécie modelo premier da biologia evolutiva, *D. melanogaster*.

Palavras-chave: Conservadorismo de nicho. DNA *barcode*. Diversidade críptica. Drosofilídeos micofílicos. Especiação. *Hirtodrosophila*. Modelagem de nicho. *Mycodrosophila*. *Paraliodrosophila*. Radiação. *tripunctata*. *virilis-repleta*. *Zygothrica*.

ABSTRACT

PhD thesis
Programa de Pós-graduação em Biodiversidade Animal
Universidade Federal de Santa Maria, RS, Brazil

TAXONOMIC, ECOLOGICAL AND EVOLUTIONARY ASPECTS IN NEOTROPICAL DROSOPHILIDS

Author: Stela Machado
Advisor: Lizandra Jaqueline Robe
Co-advisor: Marco Silva Gosttschalk
Place and date: Santa Maria, May, 27th, 2015.

Drosophila melanogaster is a known model organism that has prompted studies in diverse fields of knowledge and helped in the scientific advance of several fields. However, Drosophilidae, the family in which this species is inserted, is extremely speciose, being composed by more than 4,200 species, most of which with deficient or absent knowledge, besides many unknown species. Even hypothesis concerning the evolution of the group which were formulated in 1975 were not yet properly tested. Therefore, the present thesis aims to fill some of these gaps by testing the hypothesis of radiation proposed by THROCKMORTON (1975) in different taxonomic groups of drosophilids, in order to determine if they evolved through adaptative radiations or are merely taxonomic groups with difficult phylogenetic resolution. So, in chapter 1, species encompassing the *virilis-repleta* and *tripunctata* lineages of the *Drosophila* subgenus were used, and the results have shown that the evolution of this group is more complex than previously proposed, with the first lineage presenting signs of niche conservatism and the second being compatible with the occurrence of an adaptive radiation. In chapter 2, we also found signs of abiotic niche conservatism between mycophylic species belonging to *Zygothrica* genus group that use fungal fruitification bodies as resource for feeding, oviposition and/or breeding. Also in this chapter, we contributed to the knowledge of the family diversity by expanding the records and modeling the potential distribution of several species of *Hirtodrosophila, Paraliodrosophila, Mycodrosophila* and *Zygothrica* recorded in Brazil. Besides, through the niche modeling approach, we showed that some species of these genera seem to be susceptible to the world climatic changes due to the abiotic factors which limit their distribution and the response curves presented by each species. In chapter 3, we explored the diversity of species of the *Zygothrica* genus group and the effectiveness of the DNA barcode methodology in their taxonomic exploitation. In this case, we demonstrated the presence of a great cryptic diversity of representatives of the *Zygothrica* genus group in the Neotropics, with the record of at least 24 putative new species. Besides, we also showed the huge potential of this methodology in the identification and species discovery in this group. In chapter 4, we explored the *Mycodrosophila* genus, principally the *M. projectans* complex, which showed to be an interesting group to the study of speciation due to the frequent presence of the cryptic diversity in sympatry and syntopy. These works highlighted how much some interesting aspects and groups of Drosophilidae have been neglected, and call attention to the fact that these are fundamental aspects to the understanding of taxonomy, ecology and evolution not just of the focused groups, but also of the premier model species of evolutionary biology, *D. melanogaster*.

Keywords: DNA barcode. Cryptic diversity. Mycophylic drosophilids. *Hirtodrosophila*. *Mycodrosophila*. Niche conservatism. Niche modeling. *Paraliodrosophila*. Radiation. Speciation. *tripunctata*. *virilis-repleta*. *Zygothrica.*

LISTA DE ABREVIATURAS E SIGLAS

Amd	α-metildopa
BOLD	Barcode of Life Data system (BOLD)
COI	Citocromo Oxidase I
COII	Citocromo Oxidase II
Ddc	Dopa descarboxilase
GMYC	*Generalized Mixed Yule Coalescent*
Hb	Hunchback
ITS	*internal transcribed spacer*
matK	maturase K
NJ	*Neighbor Joining*
PCR	reação de polimerização em cadeia
rbcL	ribulose–bifosfato carboxilase

SUMÁRIO

1. INTRODUÇÃO

A região Neotropical abriga uma grande diversidade de espécies (GIAM *et al*., 2012; LEWINSOHN E PRADO, 2005), sendo que cinco países neotropicais (Brasil, Colômbia, Equador, México, Peru e Venezuela) são reconhecidos mundialmente como megadiversos (MITTERMEIER *et al*., 1997; MITTERMEIER, 1988). Ao mesmo tempo, essa é a região biogeográfica menos estudada do planeta (LEWINSOHN E PRADO, 2005). Dentre os países megadiversos, o Brasil certamente merece destaque, uma vez que possui de 10% a 13% da diversidade biológica mundial (LEWINSOHN E PRADO, 2005). Contudo, a maior parte dela permanece desconhecida, sendo estimado que para cada espécie descrita dez permanecem esperando a descrição formal ou são completamente desconhecidas (LANDIM E HINGST-ZAHER, 2010). Além disso, dois dos biomas brasileiros, o Cerrado e a Mata Atlântica, são caracterizados como *hotspots* para a conservação da biodiversidade uma vez que abrigam uma grande riqueza de espécies endêmicas e apresentam uma significativa perda de habitat (MYERS *et al*., 2000). Esse cenário assume um caráter mais crítico em face da crise da biodiversidade (BUTCHART *et al*., 2010) e das mudanças climáticas globais (INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE, 2007).

A crise da biodiversidade, relacionada a elevadas taxas de extinção de espécies, sobretudo devido a pressões antrópicas (BUTCHART *et al*., 2010), ressaltou o problema do impedimento taxonômico, uma vez que o número de taxonomistas é reduzido quando comparado com o elevado número de espécies a serem descritas. No ritmo de extinção atual, muitas espécies serão perdidas antes mesmo que sejam conhecidas e que medidas para protegê-las sejam adotadas. No Brasil, por exemplo, apesar de iniciativas governamentais que incentivam a formação de novos taxonomistas, a taxa de adição de espécies novas é de aproximadamente 0,6% por ano. Com essa taxa, serão necessários mais de 10 séculos para que toda a biodiversidade brasileira seja conhecida (LANDIM E HINGST-ZAHER, 2010).

A delimitação e descrição das espécies é um ponto crucial para compreensão da biodiversidade, contudo é um processo demorado e minucioso. Idealmente, busca-se que as descrições sejam realizadas em uma abordagem integrativa, incorporando múltiplos aspectos como filogeografia, morfologia comparativa, ecologia e comportamento (DAYRAT, 2005; WILL *et al.*, 2005). Na prática, entretanto, a maioria dos taxonomistas utiliza principalmente aspectos morfológicos

nesse processo. Na verdade, para a maioria das espécies de invertebrados formalmente descritas, observamos na literatura apenas suas descrições originais, sem qualquer caracterização complementar.

Dessa forma, para permitir que avanços significativos no conhecimento desses grupos ocorram, é importante o emprego de metodologias versáteis, capazes de responder questões com diferentes enfoques, mesmo quando os dados disponíveis são limitados. Nesse contexto, metodologias como a modelagem de nicho ecológico (também chamada de modelagem de distribuição potencial das espécies) e o DNA *barcode* permitem explorar diferentes aspectos, desde a distribuição e o reconhecimento das espécies, até aspectos ecológicos e evolutivos, como o entendimento de seus nichos ecológicos e dos processos associados à sua especiação.

A modelagem de nicho ecológico utiliza informações sobre a ocorrência das espécies (presença e ausência, quando disponível) e dados ambientais para gerar modelos de adequabilidade ambiental para as mesmas (ELITH E LEATHWICK, 2009). Os registros de ocorrência utilizados para gerar os modelos de distribuição são facilmente obtidos de bancos de dados, revisões na literatura e em museus. Esses modelos gerados podem ser utilizados para: (i) guiar esforços de coleta das espécies-alvo, a fim de expandir o conhecimento sobre a sua ocorrência, e permitir que os investimentos de tempo e financeiros sejam melhor aplicados pelo direcionamento das amostragens para áreas com maior probabilidade de ocorrência (GUISAN *et al.*, 2006); (ii) auxiliar na resolução de questões de âmbito filogeográfico, buscando compreender os padrões de distribuição ou dinâmica populacional das espécies modeladas no passado, uma vez que dados ambientais do Quaternário estão disponíveis (DE RÉ *et al.*, 2014; KIMURA *et al.*, 2014; PLANAS *et al.*, 2014; SVENNING *et al.*, 2011); (iii) identificar áreas onde espécies invasoras possam ocorrer e assim subsidiar a criação de políticas públicas que possam auxiliar no controle ou mesmo evitar as invasões (BARBOSA *et al.*, 2012; DA MATA *et al.*, 2010; SHATZ *et al.*, 2013); (iv) analisar aspectos ecológico-evolutivos, como o processo de especiação, uma vez que os modelos gerados podem ser utilizados para avaliar a sobreposição do nicho abiótico entre as espécies modeladas e, de posse de avaliações filogenéticas, inferir as prováveis formas de especiação, principalmente entre espécies-irmãs (GLENNON *et al.*, 2012; GRAHAM *et al.*, 2004; KOZAK E WIENS, 2006; PETERSON, 1999; PYRON E BURBRINK, 2009). Esses

são apenas alguns exemplos das potenciais aplicações da modelagem de nicho ecológico, uma vez que os modelos gerados têm sido aplicados em uma ampla gama de áreas.

Outra metodologia que vem sendo amplamente utilizada tanto em grupos com grande conhecimento taxonômico e biológico como naqueles com conhecimento deficitário é o DNA *barcode*. Essa técnica foi formalmente estabelecida por HEBERT *et al*. (2003a) com o objetivo de reconhecer espécies e identificar espécimes, além de auxiliar na descoberta de espécies novas. Segundo os autores, as sequências de DNA funcionariam da mesma forma que um código de barras, uma vez que sequências com apenas 15 pares de bases criariam um bilhão de códigos, cerca de 100 vezes mais do que o necessário para catalogar toda a diversidade de espécies do planeta. Obviamente, devido a restrições funcionais dos genes, o número real de pares de bases necessários para este fim é bem maior, e, tipicamente, fragmentos com mais de 500 pares de bases são utilizados. HEBERT *et al*. (2003a, b) sugeriram que, em animais, o fragmento de DNA que possuía maior potencial para funcionar como "código de barras" era o gene mitocondrial Citocromo Oxidase I (COI), mais especificamente, cerca de 658 pares de bases da porção 5'. Esse fragmento foi escolhido por: (i) apresentar *primers* universais descritos, facilitando assim a amplificação de espécies nunca estudadas com essa metodologia, assim como de espécimes e vestígios biológicos desconhecidos; (ii) possuir uma amplitude maior de sinal filogenético que outros genes mitocondriais (HEBERT *et al.*, 2003a), permitindo distinguir até mesmo espécies proximamente relacionadas e crípticas (HEBERT *et al.*, 2004). Contudo, cabe ressaltar que esse gene tem se mostrado adequado para resolução de questões em níveis taxonômicos menos inclusivos, sendo inadequado para resolver questões filogenéticas em níveis taxonômicos maiores (HAJIBABAEI *et al.*, 2006).

Após sua proposição, o DNA *barcode* foi testado em uma ampla gama de organismos e, hoje, sequências de mais de 230.000 espécies (cerca de 154.900 de animais, 58.700 de plantas e 16.760 de fungos e outros organismos) estão depositadas apenas no sistema *The Barcode of Life Data system* (BOLD) (RATNASINGHAM E HEBERT, 2007). Contudo, o COI não funciona bem em plantas e fungos, de forma que, para o primeiro grupo, duas regiões do DNA do cloroplasto [ribulose–bifosfato carboxilase (rbcL) e maturase K (matK)] têm sido utilizadas, (CBOL PLANT WORKING GROUP, 2009; CONSORTIUM FOR THE BARCODE OF

LIFE, 2009; JINBO *et al.*, 2011) e, no caso de fungos, a região predominantemente empregada é a do *internal transcribed spacer* (ITS) (JINBO *et al.*, 2011; SEIFERT *et al.*, 2007; SEIFERT, 2009). Mesmo para animais, alguns autores testaram outras regiões mitocondriais para comparar seus desempenhos (VENCES *et al.*, 2005; XIA *et al.*, 2012). Contudo, para a maioria das espécies, o COI tem se mostrado de fato o marcador mais adequado para a proposta do DNA *barcode* (LUO *et al.*, 2011; XIA *et al.*, 2012).

Assim como a modelagem de nicho, as aplicações do DNA *barcode* são inúmeras. Após as espécies serem cuidadosamente identificadas por taxonomistas especializados no grupo em questão, terem seu "código de barras" gerado e depositado em banco de dados criado para esse fim, essas sequências podem ser utilizadas com vários propósitos. Dentre essas aplicações estão: (i) aquelas voltadas à conservação, como monitoramento de caça ou pesca ilegal (HOLMES *et al.*, 2009; JUN *et al.*, 2011; SANCHES *et al.*, 2012); (ii) estudos ecológicos, como análises de dietas, facilitando o reconhecimento dos itens alimentares de uma determinada espécie (DEAGLE E TOLLIT, 2007; GONZÁLEZ-VARO *et al.*, 2014; JOO E PARK, 2012; ZEALE *et al.*, 2011); e (iii) a elucidação do ciclo de vida das espécies, auxiliando na identificação de estágios imaturos de desenvolvimento e no reconhecimento de dimorfismo sexual (MILLER *et al.*, 2005; PRAMUAL E WONGPAKAM, 2014; RASO *et al.*, 2014). Com fins de conservação, uma das maiores contribuições do DNA *barcode* vem sendo evidenciar a existência de espécies crípticas (CLARE *et al*., 2011; CRAWFORD *et al*., 2012; HEBERT *et al*., 2004; MUTANEN *et al.*, 2013). Um exemplo desse potencial foi relatado por MUTANEN *et al.* (2013) estudando as mariposas do grupo *Elachista bifasciella* da região paleártica. Esse grupo é de difícil identificação por apresentar muitas similaridades morfológicas, possivelmente devido à especiação recente, embora o mesmo tenha sido alvo de exaustivas investigações taxonômicas. Nesse contexto, os autores esperavam que o DNA *barcode* apresentasse uma baixa resolução devido à presença de polimorfismos ancestrais compartilhados (*incomplete lineage sorting*, em inglês) e a possibilidade de introgressão. Contudo, ao analisar 40 das 49 espécies dessa região, os autores encontraram 25 casos de possíveis espécies novas e, após detalhado exame morfológico, 16 foram confirmadas, aumentando assim a diversidade deste grupo na região em 34%. Se essa alta taxa de espécies novas é encontrada em grupos onde a taxonomia é considerada adequada pelos

especialistas, o esperado para grupos e regiões pouco explorados, como para a maioria dos drosofilídeos neotropicais, é que a riqueza de espécies desconhecidas seja ainda maior.

1.1 Drosophilidae

Drosophilidae é uma família de moscas pequenas, cujos representantes dificilmente excedem 6-7 mm de comprimento (BRYAN, 1938; GRIMALDI, 2010). Com mais de 4.200 espécies distribuídas em 77 gêneros e duas subfamílias (Steganinae e Drosophilinae) (BÄCHLI, 2015), esta família se destaca pela alta variabilidade. Dentre seus representantes, existem espécies endêmicas de pequenas regiões até espécies cosmopolitas, com táxons habitando os mais diversos ambientes, desde o nível do mar até as altas montanhas, dos trópicos até as tundras, sendo os bosques e as florestas os locais de maior diversidade e abundância (THROCKMORTON, 1975). Esse grupo apresenta um papel crucial nos ecossistemas, pois os estágios imaturos são saprófitos, de forma que a eficiência na exploração deste nicho alimentar pode estar relacionada à sua ampla distribuição (THROCKMORTON, 1975). O grupo também possui papel fundamental na biologia, uma vez que uma de suas espécies, *Drosophila melanogaster*, é utilizada como organismo modelo para estudos em diversos campos do conhecimento (JENNINGS, 2011). Na genética e na biologia evolutiva, *D. melanogaster* tem fornecido importantes informações sobre hereditariedade, mutagênese, controle gênico no desenvolvimento embrionário, além de ser um importante modelo na medicina, farmacologia e fisiologia, apenas para citar alguns exemplos (JENNINGS, 2011; PANDEY E NICHOLS, 2011; ROBERTS, 2006). Apesar do vasto conhecimento sobre esse organismo modelo, o mesmo não se aplica aos demais representantes da família, que embora também apresentem grande potencial, não têm sido muito explorados.

Em um dos primeiros trabalhos visando explorar a evolução da família a partir da análise de dados fósseis e caracteres morfológicos internos, THROCKMORTON (1975) sugeriu que Drosophilidae surgiu, provavelmente, durante o período Eoceno, há aproximadamente 50 milhões de anos. Segundo este autor, a família evoluiu a

partir de ancestrais com hábitos alimentares oportunistas e versáteis, voltados principalmente à saprofagia de folhas em decomposição, e se diversificou a partir da especialização na utilização de organismos fermentadores como fungos e bactérias que exploram diferentes recursos. Para o autor, a grande diversidade de espécie que compõem a família é explicada pela evolução através de radiações, sendo elas as radiações Steganinae, Drosophilinae, *Scaptodrosophila*, *Sophophora* e *Drosophila* (esta última subdivida nas radiações *virilis-repleta* e *immigrans-Hirtodrosophila*). Entretanto, apesar das radiações propostas por THROCKMORTON (1975) serem consideradas "radiações taxonômicas" e não necessariamente radiações adaptativas (MORALES-HOJAS E VIEIRA, 2012), por representarem apenas ranqueamentos de espécies proximamente relacionadas, o próprio autor ressalta que para alguns grupos este seria o caso.

Embora THROCKMORTON (1975) tenha fornecido o primeiro cenário evolutivo para Drosophilidae, o mesmo é criticado pela ausência de métodos filogenéticos explícitos ou de medidas de suporte para as relações inferidas (REMSEN E O'GRADY, 2002). Entretanto, este trabalho forneceu importantes hipóteses para serem testadas, e muitas destas vem sendo suportadas por trabalhos posteriores (ROBE *et al.*, 2005; ROBE *et al*., 2010a, RUSSO *et al.*, 2013; VAN DER LINDE E HOULE, 2008; VAN DER LINDE *et al.*, 2010; YASSIN, 2013). Este é o caso, por exemplo, da ubiquidade de táxons parafiléticos dentro de Drosophilidae e, especialmente, dentro do gênero *Drosophila* (O'GRADY E MARKOW, 2009; ROBE *et al.*, 2005; RUSSO *et al.*, 2013; VAN DER LINDE E HOULE, 2008; VAN DER LINDE *et al.*, 2010; YASSIN, 2013).

Análises filogenéticas mais recentes confirmaram a monofilia da família (GRIMALDI, 1990a; REMSEN E O'GRADY, 2002; RUSSO *et al.*, 2013) e de suas subfamílias (GRIMALDI, 1990a; REMSEN E O'GRADY, 2002). Entretanto, as relações filogenéticas dentro e entre alguns grupos taxonômicos ainda não foram esclarecidas (MARKOW E O'GRADY, 2007; ROBE *et al.*, 2010b; VAN DER LINDE E HOULE, 2008). Isso ocorre porque a maior parte dos estudos que buscam esclarecer as relações filogenéticas dentro da família se concentram em *Drosophila*, sendo os demais gêneros pouco representados (KWIATOWSKI E AYALA, 1999; ROBE *et al*., 2005; ROBE *et al*., 2010b; VAN DER LINDE E HOULE, 2008; VAN DER LINDE *et al.*, 2010). Mesmo dentro desse gênero, poucas relações evolutivas estão realmente bem elucidadas, uma vez que *Drosophila* é o maior gênero da

família, compreendendo mais de um quarto de toda sua diversidade (BÄCHLI, 2015). Também existe um grande viés no que tange à área de distribuição das espécies melhor caracterizadas, com a região Neotropical apresentando uma carência de estudos (ROBE *et al.*, 2005). Entretanto, para alguns grupos há filogenias robustas que auxiliam no esclarecimento da evolução das espécies, como no caso das linhagens neotropicais *immigrans-tripunctata* e *virilis-repleta*, ambas pertencentes ao subgênero *Drosophila* (DURANDO *et al.*, 2000; HATADANI *et al.*, 2009; MORALES-HOJAS E VIEIRA, 2012; OLIVEIRA *et al.*, 2012; ROBE *et al.*, 2005; ROBE *et al.*, 2010a, b; TATARENKOV E AYALA, 2001; YOTOKO *et al.*, 2003).

Nesse contexto, ROBE *et al.* (2010b), utilizaram os genes α-metildopa (*Amd*) e Dopa descarboxilase (*Ddc*) para reconstruir as relações filogenéticas entre 72 espécies, a maior parte das quais pertencentes ao subgênero *Drosophila*, um dos oito subgêneros de *Drosophila* (BÄCHLI, 2015). Os autores encontraram que o gênero *Drosophila* é parafilético em relação a *Hirtodrosophila*, *Liodrosophila*, *Scaptomyza* e *Zaprionus*. Segundo eles, as radiações *immigrans-tripunctata* e *virilis-repleta* parecem ter surgido no início do Oligoceno, sendo as linhagens *tripunctata* e *repleta*, possivelmente, derivadas e monofiléticas. Em outro trabalho, ao analisar o cenário evolutivo da linhagem *tripunctata* utilizando uma filogenia gerada com os genes *Amd*, *Ddc, hunchback* (*Hb*) e Citocromo Oxidase II (COII) e as variáveis climáticas registradas nos pontos de ocorrência das espécies, ROBE *et al.* (2010a) encontraram evidências de que essa linhagem surgiu na América do Norte ou Central, invadindo e diversificando-se na América do Sul. Os autores também testaram a hipótese de radiações sugerida por THROCKMORTON (1975) ao analisar a presença de sinal filogenético nas médias climáticas encontradas nos pontos de distribuição de cada espécie, encontrando evidências que suportam padrões de divergência de nicho. De fato, a ausência de sinal filogenético para estes caracteres é compatível com o esperado em radiações adaptativas, mediante o relaxamento da seleção natural sobre caracteres ecológicos pela invasão de novos ambientes (MORALES-HOJAS E VIEIRA, 2012; YODER *et al.*, 2010).

Por outro lado, ainda sob essa perspectiva de análise de variáveis ambientais, KELLERMAN *et al.* (2012a, b) testaram a presença de sinal filogenético para a resistência fisiológica à dissecação, ao frio e ao calor em espécies de *Drosophila*, e encontraram diferentes forças de sinal filogenético, compatíveis com a presença de uma inércia evolutiva (conservadorismo de nicho). Desta forma, contrariamente a

ROBE *et al.* (2010a), este trabalho contradiz a hipótese de radiações (adaptativas) de THROCKMORTON (1975). Na mesma linha, mas com uma perspectiva diferente, MORALES-HOJAS E VIEIRA (2012) utilizaram 218 espécies pertencentes ao subgênero *Drosophila* para testar a hipótese de radiação adaptativa através de testes que analisam modificações nas taxas de especiação ao longo da árvore filogenética. Os autores não encontraram suporte para a hipótese de radiações adaptativas, uma vez que o modo de evolução constante não pôde ser rejeitado, o que reacendeu o debate acerca da ocorrência ou não das radiações adaptativas.

No que diz respeito ao nicho biótico de Drosophilidae, um dos trabalhos mais completos feitos nos neotrópicos até o momento é o de GOTTSCHALK (2008), que analisou a utilização de frutos como sítios de oviposição por Drosophilidae na Mata Atlântica, tanto em ambientes naturais como urbanizados. Realizando uma ampla amostragem com mais de 38.700 indivíduos identificados, o autor encontrou 66 espécies emergindo de 80 espécies vegetais. Segundo ele, a linhagem *virilis-repleta* parece ser composta tanto por espécies polífagas (ex. *D. hydei*, *D. mercatorum*, *D. repleta*, *D. onca* e *D. zottii*), como monófagas (*D. annulimana*, *D. senei* e *D. zottii* affinis). Já na linhagem *immigrans-tripunctata*, a polifagia parece ser o padrão mais comum para uma ampla variedade de espécies analisadas.

Juntos os trabalhos de GOTTSCHALK (2008) e ROBE *et al.* (2010b) fornecem subsídio para avaliar interessantes questões de interface ecológico-evolutivas em espécies Neotropicais do subgênero *Drosophila*. A utilização destes estudos juntamente com dados de distribuição das espécies disponibilizados em bancos de dados especializados permite, por exemplo, avaliar questões sobre as forças envolvidas na especiação do grupo e os padrões de evolução de nicho envolvidos na sua diversificação.

1.1.1 Diversidade de drosofilídeos no Brasil

No Brasil, há o registro de mais de 300 espécies de drosofilídeos, das quais aproximadamente 60% pertencem ao gênero *Drosophila* (DÖGE *et al.*, 2008; GOTTSCHALK *et al.*, 2008; POPPE *et al.*, 2012, 2014). Entretanto, devido ao forte viés, tanto em metodologias de coleta como em regiões amostradas (GOTTSCHALK

et al., 2008), certamente essa diversidade é muito maior. De fato, em diversos estudos realizados na região (DÖGE *et al.*, 2008; GARCIA *et al*., 2012; GOTTSCHALK *et al*., 2007; GOTTSCHALK *et al.*, 2008; GOTTSCHALK *et al.*, 2009; MEDEIROS E KLACZKO, 2004; VALADÃO *et al*., 2010; POPPE *et al.*, 2014), um grande número de espécimes coletados não puderam ser devidamente atribuídos a nenhuma das espécies até então reconhecidas.

Segundo GOTTSCHALK *et al.* (2008), o método tradicional de coleta de drosofilídeos no Brasil (armadilhas com iscas de banana) é um dos principais responsáveis pelo conhecimento deficitário da diversidade de espécies de Drosophilidae, uma vez que essa metodologia atrai principalmente espécies frugívoras ou generalistas, sendo as espécies com outras preferências alimentares apenas esporadicamente coletadas. Além disso, este autor também evidencia a presença de um forte viés geográfico, uma vez que historicamente o maior esforço amostral tem se concentrado nas regiões Sul e Sudeste, abrangendo especialmente a Mata Atlântica. Entretanto, mesmo nessa região, registros de espécies não determinadas de *Drosophila* são comuns (GARCIA *et al*., 2012; GOTTSCHALK *et al*., 2007; GOTTSCHALK *et al.*, 2008; GOTTSCHALK *et al.*, 2009). De acordo com GOTTSCHALK *et al.* (2008), para termos um conhecimento maior acerca das espécies de drosofílideos, é necessário explorar outras técnicas de amostragem, ampliando as metodologias de coleta, o que pode incluir o uso de armadilhas com outras iscas atrativas, a coleta em recursos nativos como flores, fungos, frutos e cactos, e o próprio direcionamento do esforço de amostragem para outras regiões/biomas como Amazônia, Caatinga, Cerrado, Pantanal e Pampa.

1.1.2 Drosofilídeos micofílicos

Dentre os grupos de drosofilídeos com grande potencial para estudos ecológicos, mas ainda pouco explorados, estão aqueles com hábitos micófagos[1] ou

[1]Apesar do termo “micófago” ser amplamente utilizado na literatura, a associação entre algumas espécies de drosofilídeos com os corpos de frutificação de fungos macroscópicos parece ser mais complexa do que sua única utilização como recurso alimentar. Dessa forma, apesar de termos utilizado o termo “micófagos” no capítulo 2, julgamos que o termo “micofílico” pode descrever melhor

micofílicos (COURTNEY *et al.*,1990). Dentre os drosofilídeos micofílicos (Figura 1) encontram-se espécies de Drosophilinae dos gêneros *Drosophila* e *Scaptomyza* e do grupo genérico *Zygothrica* (sendo este representado pelos gêneros *Hirtodrosophila*, *Mycodrosophila*, *Paramycodrosophila*, *Paraliodrosophila* e *Zygothrica*) (GRIMALDI, 1990a), e de Steganinae do gênero *Leucophenga* (COURTNEY *et al.*, 1990). Apesar de se acreditar que a micofagia/micofilia surgiu independentemente diversas vezes dentro de Drosophilidae, possivelmente a partir de espécies detritívoras (THROCKMORTON, 1975), GRIMALDI (1990a) aponta esta característica como uma sinapomorfia do grupo genérico *Zygothrica*.

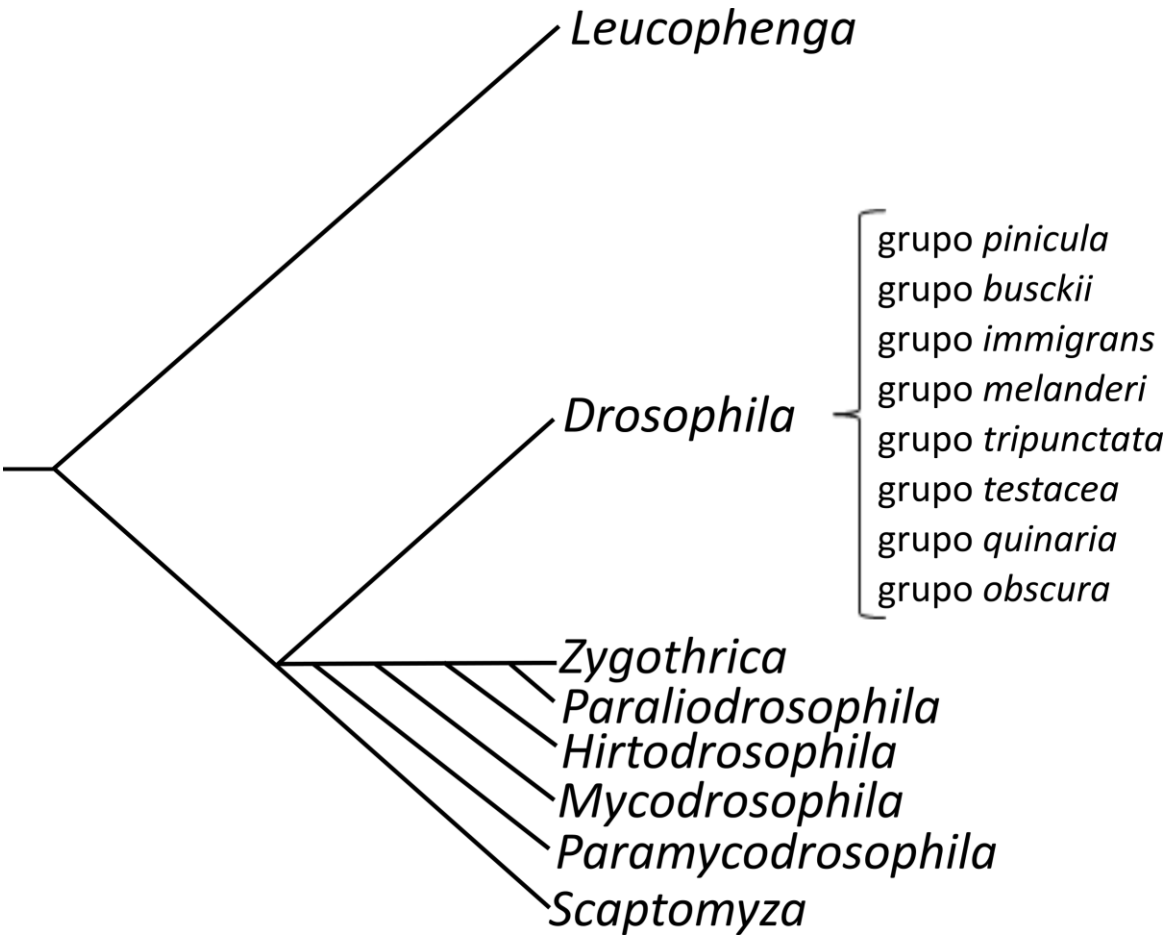

Figura 1. Principais grupos de drosofilídeos micófagos/micofílicos. Modificado de COURTNEY *et al.* (1990).

a associação de algumas espécies alvo deste trabalho, e por isso ele será adotado nos demais capítulos e no restante do texto, quando for o caso.

Apesar do fato de que todos esses grupos possuem associações ecológicas com fungos, o conhecimento a respeito dos mesmos é bem heterogêneo. Há diversos trabalhos realizados com espécies do grupo *quinaria* e *tripunctata* de *Drosophila* (COURTNEY *et al.*, 1990; GRIMALDI E JAENIKE, 1984; GRIMADI, 1985; JAENIKE, 1978a, b, 1995; JAENIKE E JAMES, 1991; ROUQUETTE E DAVIS, 2003; STUMP *et al.*, 2011; TUNO *et al.*, 2007), mas aqueles voltados, por exemplo, para a caracterização de espécies micofílicas do Havaí ou àquelas pertencentes ao grupo genérico *Zygothrica* são escassos (COURTNEY *et al.*, 1990). Na verdade, a maioria dos estudos com drosofilídeos micofílicos tem sido realizada principalmente com espécies de *Drosophila* na América do Norte, Europa e Japão (BURLA *et al.*, 1991; COURTNEY *et al.*, 1990; KASUYA *et al.*, 2013; TODA *et al.*, 1999; TUNO *et al.*, 2007; YAMASHIDA E HIJII, 2007; YOROZUYA, 2006, 2009). Na região Neotropical, há poucos estudos realizados com espécies micofílicas, e até 2007 estavam restritos principalmente a trabalhos taxonômicos de descrições ou redescrições de espécies (BURLA, 1956; FROTA-PESSOA, 1945; GRIMALDI, 1987, 1990a, b; VILELA E BÄCHLI, 1990, 2004, 2005, 2007; WHEELER E TAKADA, 1963). Uma exceção é a revisão da sistemática do gênero *Zygothrica* realizada por GRIMALDI (1987) sob a ótica da Sistemática Filogenética, onde são apresentadas também informações sobre a ecologia e comportamento das espécies. Entretanto, somente com os trabalhos de ROQUE E TIDON (2008) e GOTTSCHALK *et al.* (2009) que aspectos ecológicos de algumas destas espécies começaram a ser explorados.

1.1.2.1 Grupo genérico *Zygothrica*

Este grupo foi estabelecido por GRIMALDI (1990a), englobando cinco gêneros com diferentes graus de associação com fungos macroscópicos. Devido a essa associação e ao desconhecimento de muitos aspectos de sua biologia, as espécies desse grupo são difíceis de serem mantidas em culturas de laboratório (THROCKMORTON, 1975), principalmente as espécies de *Zygothrica* que ovipositam em flores (GRIMALDI, 1987). A maior diversidade de espécies deste grupo genérico se concentra na região circuntropical, contudo não está restrita a ela. *Hirtodrosophila* e *Mycodrosophila* são gêneros cosmopolitas, já *Paraliodrosophila* e

Zygothrica são essencialmente Neotropicais e *Paramycodrosophila* está presente no Oriente, com poucas espécies chegando a América do Norte e Central (GRIMALDI, 2010; VILELA E BÄCHLI, 2004; WHEELER E TAKADA, 1963).

As relações filogenéticas dos gêneros de drosofilídeos micofílicos precisam ser melhor elucidadas, uma vez que os trabalhos que existem incorporaram poucos representantes dos mesmos, principalmente do grupo genérico *Zygothrica* (DA LAGE *et al.*, 2007; ROBE *et al.*, 2005; RUSSO *et al.*, 2013; VAN DER LINDE *et al.*, 2010). THROCKMORTON (1975) sugeriu que o grupo surgiu a partir da radiação *immigrans-Hirtodrosophila* do gênero *Drosophila*, que se diversificou inicialmente na região tropical do Velho Mundo e se separou nas linhagens *tripunctata* e *Hirtodrosophila* que posteriormente colonizaram o Novo Mundo. Entretanto, ao realizar análises filogenéticas com base em caracteres morfológicos, GRIMALDI (1990a) sugeriu que o grupo genérico *Zygothrica* é monofilético e basal em relação ao gênero *Drosophila* (Fig. 2).

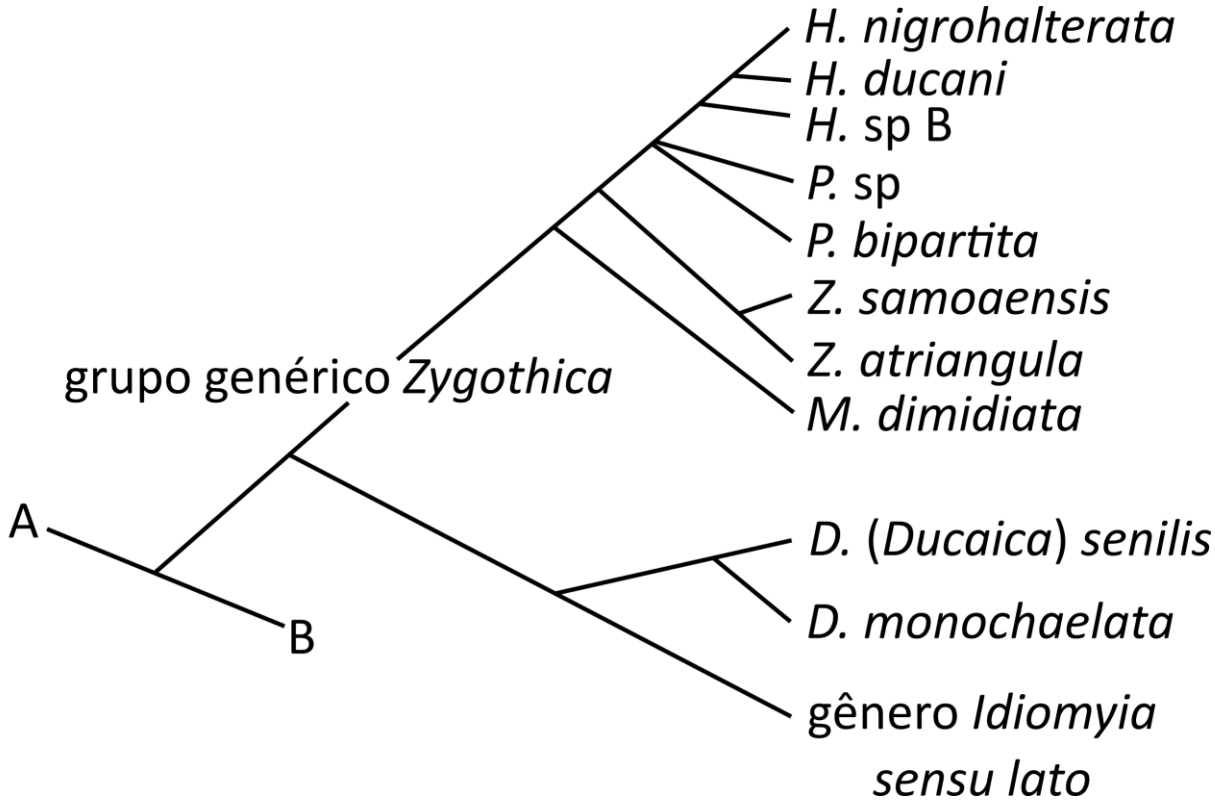

Figura 2. Cladograma das relações filogenéticas do grupo genérico *Zygothrica*. Modificado de Grimaldi (1990a).

Mais recentemente, uma série de trabalhos vem apresentando suporte para o posicionamento de diferentes representantes do grupo genérico *Zygothrica* dentro do gênero *Drosophila*, de modo a torná-lo parafilético (ROBE *et al.*, 2005; RUSSO *et*

al., 2013; VAN DER LINDE *et al.*, 2010; YASSIN, 2013). Além disso, VAN DER LINDE *et al.* (2010) encontraram evidências do não monofiletismo de *Hirtodrosophila*, uma vez que uma de suas espécies, *H. duncani*, foi agrupada com o subgênero *Sophophora*, distante do clado formado por *H. pictiventris*, *H. thoracis* e *H.* sp. e três espécies de *Mycodrosophila*. Na mesma linha, RUSSO *et al.* (2013) utilizaram seis genes nucleares para reconstruir a filogenia da família, e ao incluir sete espécies de *Hirtodrosophila* e três de *Mycodrosophila* (sendo apenas *H. pictiventris* originaria do Novo Mundo e as demais da Eurásia), não encontraram suporte para a monofilia de *Hirtodrosophila*, cujas espécies foram divididas em dois clados, um agrupado com *Dichaetophora* e outro com *Mycodrosophila*. Entretanto, devido ao baixo suporte dos clados (*bootstrap* inferior a 50), politomias caracterizariam melhor as relações dos grupos, e os clados sugeridos pelos autores não se sustentam. Além disso, exceto por *H. pictiventris* onde os seis genes foram analisados, apenas um gene foi utilizado na reconstrução das relações das demais espécies, em geral *Adh*. Para corroborar o monofiletismo ou não do grupo é necessário realizar a inclusão de mais espécies, representando tanto as linhagens do Velho como do Novo Mundo, além da inserção de mais genes na análise.

1.1.2.1.1 *Hirtodrosophila*

Este gênero cosmopolita (GRIMALDI, 2010; VILELA E BÄCHLI, 2004) é composto por 160 espécies (BÄCHLI, 2015), 28 das quais são encontradas na região Neotropical (VILELA E BÄCHLI, 2004) e 16 registradas no Brasil (GOTTSCHALK *et al.*, 2008). Contudo, de acordo com GRIMALDI (2010) muitas espécies ainda esperam para serem descobertas/descritas. O gênero é estritamente micofílico, com raras exceções, como por exemplo, *H. batracida,* que ocorre na América Central e preda ovos de rãs (COURTNEY *et al.*, 1990; GRIMALDI, 1994). Ao contrário de *Mycodrosophila* e *Zygothrica*, o gênero *Hirtodrosophila* não parece ser territorialista (GRIMALDI, 1987). Análises de emergência de adultos a partir de substrato têm encontrado espécies de *Hirtodrosophila* emergindo de fungos das famílias Tricholomataceae (GRIMALDI, 1987), Pleurotaceae (ROQUE E TIDON,

2008), Agaricaceae, Auriculariaceae, Marasmiaceae e Polyporaceae (GOTTSCHALK *et al.*, 2009).

1.1.2.1.2 *Mycodrosophila*

Mycodrosophila é composto por cerca de 130 espécies (BÄCHLI, 2015) distribuídas mundialmente, principalmente nos trópicos (MCEVEY E POLAK, 2005; THROCKMORTON, 1975; WHEELER E TAKADA, 1963). Contudo, menos de 4% delas são encontradas na região Neotropical (MCEVEY E POLAK, 2005). De maneira geral, como o próprio nome sugere, as espécies desse gênero são micófagas obrigatórias (COURTNEY *et al.*, 1990), associadas principalmente com fungos Polyporaceae (GOTTSCHALK *et al.*, 2009; WHEELER E TAKADA, 1963), Tricholomataceae (GRIMALDI, 1987), Agaricaceae (GOTTSCHALK *et al.*, 2009) e Pleurotaceae (ROQUE E TIDON, 2008). No Brasil, WHEELER E TAKADA (1963) descreveram três espécies: *M. projectans*, *M. elegans* e *M. brunnescens*. Entretanto, como a maioria dos trabalhos tem sido realizada com os gêneros *Hirtodrosophila* e *Zygothrica*, a diversidade de *Mycodrosophila* pode ser maior.

ROQUE E TIDON (2008) e GOTTSCHALK *et al.* (2009) avaliaram espécies que emergiram de cogumelos, e ambos trabalhos registraram a emergência apenas de *M. projectans*. No caso de ROQUE E TIDON (2008), esta espécie foi encontrada nas emergências de *Pleurotus* sp. (Pleurotaceae), oriundo do Parque Nacional Chapada dos Guimarães/MT. Já GOTTSCHALK *et al.* (2009) registraram emergência a partir de *Coprinus comatus* e *Macrolepiota* sp. (Agaricaceae), além de Polyporaceae, oriundos de Porto Alegre/RS e Ribeirão Preto/SP. De fato, *M. projectans* parece ser a espécie Neotropical desse gênero com a distribuição mais ampla, sendo registrada na América do Norte, América Central, Ilhas Caribenhas e na América do Sul, onde ocorre no Brasil, na Bolívia, Colômbia e Equador (BÄCHLI, 2015; WHEELER E TAKADA, 1963).

1.1.2.1.3 *Paramycodrosophila*

Esse gênero é composto por 16 espécies, a maioria das quais se encontra distribuída na região da Ásia, Oceania e Austrália, com algumas poucas espécies ocorrendo na América do Norte e Central, de forma que o mesmo ainda não foi registrado na região Neotropical (BÄCHLI, 2015).

1.1.2.1.4 *Paraliodrosophila*

Esse gênero é composto por apenas cinco espécies (BÄCHLI, 2015). A maioria delas ocorre na América Central, exceto por *P. antennata* (= *H. thoracis* senso BURLA, 1956), que apresenta ampla distribuição no continente Americano, e *P. burlai*, que é registrada apenas no Brasil (BÄCHLI, 2015). Por serem muito parecidas com *Mycodrosophila*, às vezes podem ser confundidas com estas, mas se diferenciam pela ausência de *lappet* desenvolvido (VILELA E BÄCHLI, 2007). Confirmando a hipótese de VILELA E BÄCHLI (2007), a respeito da associação desse gênero com fungos macroscópicos, GOTTSCHALK *et al.* (2009) coletaram *P. antennata* emergindo de fungos Marasmiaceae e Polyporaceae.

1.1.2.1.5 *Zygothrica*

Este gênero é composto por 124 espécies (BÄCHLI, 2015) e apresenta uma distribuição predominantemente Neotropical (GRIMALDI, 1987; GRIMALDI E FENSTER, 1989; THROCKMORTON, 1975), com poucas espécies registradas na África e no Indo-Pacífico (GRIMALDI, 1990b). Possivelmente é o gênero Neotropical mais especioso de Drosophilidae (GRIMALDI, 1987), com 54 espécies registradas apenas no Brasil (GOTTSCHALK *et al.*, 2008). Os membros desse gênero são encontrados em alta densidade de indivíduos nos corpos de frutificação dos fungos (Figura 3), que são utilizados como sítios de alimentação, cortejo e arena para

disputa de fêmeas (COURTNEY *et al.*, 1990; GRIMALDI, 1987). Contudo, parece ocorrer uma segregação no uso de recursos entre adultos e larvas, de maneira que os fungos não parecem ser os sítios de oviposição para a maioria das espécies (GRIMALDI, 1987). De fato, enquanto nos outros gêneros do grupo genérico *Zygothrica* a micofagia das larvas parece ser frequente, em *Zygothrica* apenas cerca de 10% das espécies parecem utilizar esse recurso como fonte alimentar (COURTNEY *et al.*, 1990). Dessa forma, flores parecem ser o principal sítio de oviposição e alimentação das larvas para a maioria das espécies (GRIMALDI, 1987, 1990b; MALOGOLOWKIN, 1952; PIPKIN *et al.*, 1966). Segundo GRIMALDI (1987), a micofagia em *Zygothrica* parece refletir a retenção de um estado ancestral compartilhado com os grupos irmãos *Mycodrosophila* e *Hirtodrosophila*. Além disso, *Zygothrica* também parece ser pouco especializada na escolha dos fungos, apresentando registros em Agaricaceae, Marasmiaceae, Polyporaceae, Tricholomataceae e Auriculariaceae (GRIMALDI, 1987; GOTTSCHALK *et al.*, 2009).

Figura 3. Fungo colonizado predominantemente por indivíduos do gênero *Zygothrica*. Foto: Stela Machado, Pedro M. Fonseca.

Uma característica marcante em algumas espécies de *Zygothrica* é o dimorfismo sexual, caracterizado em alguns casos por uma extrema hipercefalia nos machos. Este é o caso, por exemplo, do grupo *dispar*, onde este traço assume proporções bizarras (GRIMALDI, 1987; GRIMALDI E FENSTER, 1989). Os machos

destas espécies são extremamente territorialistas (GRIMALDI, 1987) e as disputas se dão através de “brigas com cabeçadas” (GRIMALDI E FENSTER, 1989).

1.1.3 Potenciais do grupo genérico *Zygothrica*

KELLERMANN *et al.* (2012a, b) têm demonstrado experimentalmente como fatores abióticos como dissecação, calor e frio são fatores importantes na distribuição de Drosophilidae e o quanto este grupo pode ser influeciado pelo aquecimento global. A associação das espécies do grupo genérico *Zygothrica* com fungos, organismos altamente suscetíveis às alterações em termos de umidade, precipitação e temperatura, as torna ainda mais interessantes para estudos relacionados ao impacto das mudanças climáticas, tanto para mudanças atuais como no passado. Contudo, o primeiro passo para grupos onde o conhecimento é tão escasso, é o emprego de metodologias como a modelagem de nicho ecológico e o DNA *barcode*, que permitem explorar aspectos básicos sobre a diversidade e a biologia das espécies. Por meio da modelagem, utilizando os dados de ocorrência disponíveis principalmente nos trabalhos de descrição e eventuais registros adicionais é possível começar a inferir a potencial distribuição das espécies, assim como fatores abióticos que regem e limitam suas distribuições. Já o DNA *barcode* pode auxiliar a revelar a diversidade do grupo, além de facilitar a realização de trabalhos posteriores, uma vez que um grupo muito pequeno de taxonomistas é capaz de identificar estas espécies com eficiência. A partir disso, estudos voltados à compreensão da história evolutiva do grupo e de espécies dentro dele se tornam mais factuais.

2. OBJETIVOS

2.1. Objetivo Geral

O objetivo da presente tese é avaliar aspectos taxonômicos, ecológicos e evolutivos em drosofilídeos neotropicais, utilizando como modelos espécies do subgênero *Drosophila* e do grupo genérico *Zygothrica*, e como ferramentas as técnicas de DNA *barcode* e modelagem de nicho.

2.2. Objetivos Específicos

- Testar a hipótese de evolução de espécies Neotropicais do subgênero *Drosophila* por meio de radiações, utilizando a abordagem de modelagem de nicho ecológico para avaliar os padrões de sobreposição de nicho (Capítulo 1);

- Caracterizar os padrões de distribuição geográfica e as potenciais limitações climáticas de drosofilídeos micofílicos Neotropicais pertencentes ao grupo genérico *Zygothrica* (Capítulo 2), testando a hipótese de evolução por meio de radiação adaptativa;

- Avaliar a aplicabilidade e a eficiência da metodologia do DNA *barcode* como ferramenta para conhecer a diversidade de espécies do grupo genérico *Zygothrica* na região Neotropical e impulsionar estudos com o grupo (Capítulo 3);

- Caracterizar a diversidade críptica do gênero *Mycodrosophila* nos Neotrópicos, com base em marcadores moleculares e morfométricos (Capítulo 4), avaliando alguns de seus padrões evolutivos.

3. RESULTADOS

3.1. Capítulo 1

Título:

Historical patterns of niche dynamics in Neotropical species of the Drosophila subgenus (Drosophilidae, Diptera): insights about *tripunctata* and *virilis-repleta* radiations

Manuscrito submetido ao periódico:

Evolutionary Ecology

ISSN:

0269-7653 (versão impressa)
1573-8477 (versão online)

Historical patterns of niche dynamics in Neotropical species of the *Drosophila* subgenus (Drosophilidae, Diptera): insights about *tripunctata* and *virilis-repleta* radiations

Stela Machado[1] stelamchd@gmail.com

Marco Silva Gottschalk[3] gotts007@yahoo.com

Lizandra Jaqueline Robe[1,2] lizbiogen@gmail.com

[1]Universidade Federal de Santa Maria – UFSM, Programa de Pós Graduação em Biodiversidade Animal, Avenida Roraima 1000, CEP 97105-900 - Santa Maria, RS - Brazil, Phone: +55 55 3220-8912

[2]Universidade Federal do Rio Grande – FURG, Programa de Pós Graduação em Biologia de Ambientes Aquáticos Continentais, Av. Itália, km 08 Campus Carreiros, CEP 96203-900 - Rio Grande, RS – Brasil, Phone: +55 53 3293-5128

[3]Universidade Federal de Pelotas – UFPel, Instituto de Biologia, Departamento de Ecologia, Zoologia e Genética, Campus Capão do Leão, Mail Box 354, CEP 96010-900 - Capão do Leão, RS – Brazil, Phone: +55 53 3275-7340

Corresponding Author:

Lizandra Jaqueline Robe

Short Running Title: Niche dynamics in the *Drosophila* subgenus

Keywords: Neotropics; niche equivalency tests; niche modeling; niche overlap; phylogenetic signal; speciation.

Total word count: 7,066

Separate word counts for each section: Abstract (211), introduction (1,024), material and methods (1,987), results (1,567), discussion (2,077), conclusion (165), acknowledgements (65).

Number of cited references: 96

Number of figures and tables: three figures and three tables.

List of items to be published as online appendix: one table and two figures

Abstract

Niche conservatism (NC) presence is a controversial question in evolutionary ecology. In *Drosophila,* little is known about which is the preponderant evolutionary pattern, since the radiations hypothesis first proposed by Throckmorton assumed niche divergence according to a niche occupancy scenario. Nevertheless, this hypothesis was not yet straightforwardly tested. So, our aim here was to test the role of NC patterns across evolution of American drosophilids belonging to the *tripunctata* and *virilis-repleta* lineages of the *Drosophila* subgenus, through measures of geographical, abiotic and biotic niche overlap and evaluations regarding the presence of phylogenetic signal or niche identity. We were able to recover phylogenetic signal attributable to phylogenetic niche conservatism when all species were analyzed together, but not in more restricted groups. As concerns the performed identity tests, for the *tripunctata* lineage niche equivalency was seldom rejected, although contrasting results were obtained for the *virilis-repleta* lineage species. So, despite the fact that our results for the *Drosophila* subgenus do not support the hypothesis of radiation proposed for this group, analysis performed at more restricted levels suggest phylogenetic associations have arisen early in the phylogeny of these species, and that latter the two main lineages may have followed contrasting niche dynamics patterns, with the *virilis-repleta* lineage species evolution presenting signs of an actual radiation.

Introduction

The implementation of new approaches in the study of evolution has been adding to the knowledge regarding the role of the environment in this process (Wiens and Graham 2005). The use of ecological niche modeling (ENM), which uses occurrence information about species (presence and absence whenever available) and environmental layers to generate models of environmental suitability for a given species (Elith and Leathwick 2009) has been proved a useful tool in this field. ENM approaches have gained special prominence in studies aiming to identify the role of niche conservatism (NC) or phylogenetic niche conservatism (PNC) in the speciation process (Peterson et al 1999; Graham et al 2004; Kozak and Wiens 2006; Mcnyset 2009; Pyron and Burbrink 2009; Glennon et al 2012). Such patterns may arise through evolutionary constraints owing to the action of a stabilizing selection, due to high rates of gene flow or even due to genetic constraints associated with pleiotropy or with the lack of genetic variation (Wiens and Graham 2005). Some authors have demonstrated that NC could be associated to allopatric speciation because individuals would be unable to persist or disperse through unfavorable environmental conditions, driving to fragmentation (Wiens 2004; Wiens and Graham 2005; Kozak and Wiens 2006). Alternatively, niche divergence (ND) may lead to cladogenesis while populations adapt to new environments (Coyne and Orr 2004).

Diverse evidences have pointed NC as a common process, which presents an important role in the diversification of species and lineages (Peterson et al 1999; Kozak and Wiens 2006; Kozak and Wiens 2010; see Peterson 2011 supplementary material). However, despite the advances in this area, the role of NC for evolution has been questioned and general conclusions about the theme are conflicting, mainly due to the lack of a widely accepted test (Warren et al 2008; Godsoe 2010; Harmon and Glor 2010; Wiens et al 2010; Peterson 2011; Hardy and Pavoine 2012). The most basic test of PNC involves the searching for the presence of a phylogenetic signal (Münkemüller et al 2012), although alternative models for discovering the origins of the patterns (see Cooper et al 2010 and Crisp and Cook 2012 for revision), or even methodologies based on ENM perspectives (Warren et al 2008) have also been developed.

Drosophilidae is a highly diverse family, comprising more than 4,000 described species (Bächli 2014), which vary in terms of distribution from highly endemic to entirely cosmopolitan, being adapted to explore the more diverse array of resources (Throckmorton 1975; Val et al 1981; Tosi et al 1990; Chaves and Tidon 2008; Markow and O'Grady 2008; O'Grady and Markow 2009). Within this family, *Drosophila* encompasses the best-studied biological system. Even so, most of these studies are widely concentrated in a small number of species, and few attempts have been made in order to understand the evolution of the whole group (Throckmorton 1975; Grimaldi 1990; van der Linde and Houle 2008; Yassin 2013; Russo et al 2013). According to Throckmorton (1975), Drosophilidae diversification would have occurred through five successive adaptive radiations, which could explain the enormous diversity of species and habits within the family. One of this would have given rise to the *Drosophila* subgenus, which attained successful diversification levels on the Neotropics after independent expansions departing from the Old World (Throckmorton 1975; Robe et al 2010a).

Radiations occur when underutilized environments become available to an ancestral species allowing a rapid diversification into this new reality through a relaxation of natural selection on ecological traits (Yoder et al 2010; Morales-Hojas and Vieira 2012), so that ND would drive adaptive radiation. Despite the early suggestion of radiations involved in *Drosophila* evolution, the hypothesis of ND or absence of PNC was not yet straightforwardly tested for the group. An impediment to this is the fact that, to test PNC, phylogenetic relationships of the group under study should be well elucidated. This seems to be the case for at least part of the *tripunctata* and the *virilis-repleta* lineages (Durando et al 2000; Tatarenkov and Ayala 2001; Yotoko et al 2003; Robe et al 2005; Hatadani et al 2009; Robe et al 2010a, b; Morales-Hojas and Vieira 2012; Oliveira et al 2012), all of which belong to the Neotropical *Drosophila* subgenus. Robe et al (2010b) used their well-supported phylogeny of the *tripunctata* lineage together with climatic data associated with the occurrence points of each species to test PNC in an ENM devoid perspective. This study did not detect the presence of a phylogenetic signal for the evaluated climatic variables, suggesting that ND may be the predominant pattern in the diversification of the *tripunctata* lineage, in agreement with the hypothesis proposed by Throckmorton (1975). On the other hand, diversification analyses showed no support for adaptive radiation as a result of geographic dispersal or ecological resource shift in the

Drosophila subgenus, once an increased rate of speciation was not detected in a broad study with more than 200 species (Morales-Hojas and Vieira 2012).

Considering the evidences presented above, the aim of this study was to test the role of NC in the speciation process of American drosophilids, more specifically, for a subset of Neotropical species included in the *tripunctata* and *virilis-repleta* lineages of the *Drosophila* subgenus. For it, we used niche identity and phylogenetic signal measures to evaluate PNC patterns presented by (i) the abiotic niche properties, evaluated either through ENMs overlap patterns or through the mean value presented by the set of records compiled for each species' in regard to each climatic variable; (ii) the biotic niche overlap, presented by breeding sites resources. These analyses were also controlled in regard to their patterns of spatial autocorrelation in order to test if the detected phylogenetic signal could be an artifact of spatial proximity (phylogenetic structured adaptation) instead of a reflection of phylogenetic inertia (Freckleton and Jetz 2009). This integrative perspective differs from others already used for this group, where tests for presence of phylogenetic signal were performed using environmental data extracted from occurrence points (Robe et al 2010b) or through physiological constraints presented by each species (Kellerman 2012a, b). Thus, we hope to provide here subsidies to the discussion about the role adaptative radiations performed in the evolution of Neotropical species belonging to the *Drosophila* subgenus.

Material and methods

Selection of taxa and data compilation

As concerns the use of a straightforward PNC test encompassing the search for the presence of a phylogenetic signal, one point is imperative: the availability of a phylogenetic tree (Crisp and Cook 2012) to provide patristic distances. In this case, independent phylogenetic studies cannot be simultaneously considered since patristic distances needed to evaluate phylogenetic signal vary according the employed molecular marker. In this case, we consulted the resolved phylogenetic tree published by Robe et al (2010a) for the *Drosophila* subgenus to elect the species used in this study, starting with 87 species. Nevertheless, as some of these

are cosmopolitan (e.g. *D. immigrans* and *D. funebris*) or have only a few registers, this set was additionally reduced in order to add confidence to the obtained results.

So, we selected 25 entirely Neotropical species of the *Drosophila* subgenus (Table 1) that had at least 20 occurrence points recorded [the sole exception to this cut-off was *D. cuaso*, which was maintained despite its somewhat reduced number of registers (n = 13) because of its recent divergence in regard to its sister species, *D. paraguayensis*]. Despite the drawbacks related to the somewhat low sample size, this conservative criterion was necessary once ENM was demonstrated to be affected by the number of occurrence points (Stockwell and Peterson 2002; Hernandez et al 2006; Wisz et al 2008). In fact, Hernandez et al (2006) have shown that although useful ecological niche models may be produced with as few as 5-10 points, increased samples within this range are usually associated with smaller performance variances.

The *Drosophila* occurrence points were first compiled from Taxodros database (Bächli 2012). However considering that Neotropical registers may be biased in virtue of limitations related to the number of records and to imprecise geographical locations (Kamino et al 2012), all points were carefully revised. In this step, special attention was taken in cases of taxonomic conflicts [as with *D. cuaso* and *D. paraguayensis*, which were separated as independent cryptic species by Bächli et al (2000), so that registers previous to this date were disregarded].

Available literature for each of these species was also reviewed in order to compile information regarding the substrates used as breeding sites. For this, a main set of data was previously generated by Gottschalk (2008), who performed samplings looking for fruits in Atlantic Forest and urbanized areas around Florianópolis (Brazil) and monitoring for fly emergence. This work presents a large sampling effort, with records of breeding sites for 66 species of this group. Additionally, our data were complemented with information available in Frota-Pessoa (1952), Heed (1956), Pipkin et al (1966), Saavedra et al (1995a, b), De Toni et al (2001), Medina-Muñoz and Godoy-Herrera (2004), Santos and Vilela (2005), Blauth and Gottschalk (2007), Martins and Santos (2007), Roque and Tidon (2008), Gottschalk et al (2009), Valadão et al (2010), some of which also presented information regarding fungi and flowers substrates.

(Place to table 1)

Ecological Niche Modeling

For groups of invertebrates, as drosophilids, models that use presence-only data for generating ENM are commonly chosen over those using presence and absence records given the paucity and flaws of absence inferences regarding these taxa (Lobo and Tognelli 2011). In fact, absence inferences can introduce bias not only when the habitat is suitable but unoccupied or when the species is not at equilibrium, but also when the species remained undetected or was not correctly identified (Kamino et al 2011; Rocchini et al 2011). Since comparative studies between different ENM methods have demonstrated that Maxent (Phillips et al 2006) produces better results, especially when a limited number of records are available (Hernandez et al 2006; Wisz et al 2008), this algorithm was chosen to model the potential distribution of the species target of this study. Maxent is a machine learning method that estimates the probability distribution from incomplete information, using the principle of maximum entropy (Phillips et al 2006). Maxent estimates the most uniform distribution (maximum entropy) across a study area, given the constraint that the expected value of each environmental predictor variable under this estimated distribution corresponds to its empirical average (average values for the set of occurrence data) (Phillips et al 2004, 2006; Hernandez et al 2006). To fit the maximum entropy models, we used the software Maxent (version 3.3.3k, available in http://www.cs.princeton.edu/~schapire/maxent/). For all model runs, we used 25% of the data as test and 75% as training or calibration information, in each of 50 bootstrap replicates. For these, in addition to the presence records, 10,000 backgrounds points were randomly chosen for each species within the study area (America continent). The maximum number of iterations was set to 50,000 and a default regularization parameter was used (Phillips and Dudik 2008). The accuracy of the predictive distribution models was accessed through the AUC value (Area Under the Receiving Operating Curve), which represents the probability that a random chosen presence site is ranked above a random background site (Phillips et al 2006).

The environmental layers were obtained from WorldClim with resolution of 5 arc minutes (available in http://www.wordclim.org/; accessed in 12/21/2012) (Hijmans et al 2005). To avoid collinearity among variables, which can lead to the misidentification of the most relevant predictors (Kamino et al 2012), Pearson correlation coefficient between layers was measured in ENMTools 1.3 software (Warren et al 2010). This led to the choice of five variables showing small

correlations among each other (r < |0.75|), which were further used in the Maxent model projections: Bio6: Minimum temperature of coldest month, Bio 14: Precipitation of driest month, Bio 15: Precipitation seasonality (coefficient of variation), Bio18: Precipitation of warmest quarter, Bio19: Precipitation of coldest quarter. Other variables, as vegetation or soil properties, were not included, since it was empirically demonstrated by Kellerman et al (2009) that temperature and humidity are the limiting factors for the distribution of tropical drosophilids.

Abiotic Niche overlap

Abiotic pairwise niche overlap measures were obtained in ENMTools 1.3, with Schoener's D (D) and a measure derived from Hellinger's niche overlap (I), using the individual environmental suitability distribution models projected by Maxent. These measures of similarity range from 0 (niche models have no overlap) to 1 (niche models are identical) and are obtained by comparing the estimates of habitat suitability calculated for each grid cell in the study area. As suggested by Warren et al (2008), no threshold was adopted for these analyses because it could obscure biological details, leading to incorrect interpretation about patterns of niche overlap.

To test if the ENM generated for two species are more different than expected if they were drawn from the same underlying distribution, we used identity tests (Warren et al 2008, 2010) for the seven pairs of sister species included in our study (*D. bandeirantorum* and *D. pallidipennis, D. cuaso* and *D. paraguayensis*, *D. gasici* and *D. mesophragmatica*, *D. gaucha* and *D. pavani*, *D. griseolineata* and *D. maculifrons*, *D. mediopunctata* and *D. unipunctata*, *D. neocardini* and *D. polymorpha*). To test if niche identity also applies at higher divergence times, pairwise comparisons between species of clades 12, 22, 32 and 61 of Robe et al (2010a) were also accomplished, holding a total of 66 comparisons. In these tests, ENMs were created for each of 200 pseudoreplicates, in which empirical occurrences for the two species were pooled together and then randomized in order to produce two new samples with the same number of observations as the empirical data, and these were compared using D and I similarity measures. After, the original values for each pairwise comparison were compared with the null hypotheses data, and the hypothesis of niche similarity was rejected if empirical values were lower than the values obtained from the 5% left-handed distribution of the pseudoreplicated data sets (Warren et al 2008, 2010).

Range overlap measures

ENMTools was also used to evaluate pairwise range overlap (R), with the Maxent binary models achieved with the mean of the Minimum Training Presence threshold (MTP) presented by all species. Besides, distance in kilometers partitioning distribution between species was measured through the use of the respective midpoint distributions (mean geographical coordinate) (G_1) or through the mean of pairwise distances connecting the set of geographical coordinates between species (G_2).

Breeding sites overlap

Niche overlap regarding the type of substrate used by each species as breeding sites was evaluated in Ecosim 7.71 (Gotelli and Entsmiger 2001), using Czekanowski index (C), which ranges from 0 (without resource sharing) to 1 (identical resource use) (Gotelli and Entsmiger 2001). Significance of this measure was evaluated using the randomization combination RA3 (niche breadth retained/ zero states reshuffled), with 10,000 iterations. Despite the fact that resources are not usually equally available in environment, we used resource states as equiprobable, because we were unable to measure this difference.

Genetic distance measures

Phylogenetic patristic distance (P) was obtained in Passage 2 (Rosenberg and Anderson 2011) using the amd+ddc Bayesian tree presented by Robe et al (2010a). Patristic distance is calculated by summing the lengths of the branches that connect two nodes [typically terminal nodes] in a phylogenetic tree (Philippe et al 2011). In addition, divergence times (A) estimated by Robe et al (2010a) based on Tamura et al (2004) TMRA prior were also compiled in order to perform the phylogenetic signal evaluation (see below).

Phylogenetic signal

We employed different strategies for this measure:

(1) First, a Partial Mantel test, as performed in Passage 2 with 10,000 permutations, in order to measure the correlation between phylogenetic distances (P) or divergence time (A), and abiotic niche overlap (I and D) or biotic niche overlap (C),controlling for geographical distance (G_1 and G_2). The results of these tests were considered after a

Bonferroni correction for multiple tests, adapting the 5% level of significance to P <0.0042 and 0.00625 for the comparisons involving abiotic and biotic properties, respectively. Different strategies were employed in order to evaluate the presence of phylogenetic signal at different levels of phylogenetic affinity, based on the clades presented in Figure 4 of Robe et al (2010a): all species together (subgenus *Drosophila*); the *tripunctata* lineage and the *virilis-repleta* lineage.

(2) Second, a measure of Bloomberg's *K* (Blomberg et al 2003) and Pagel's ʎ (Pagel 1999) for all species or each lineage separately, as performed in R (R Development Core Team 2008) using Phytools package (Revell 2012) with a Bonferroni correction. Different from the Partial Mantel tests, these analyses were performed in an ENM devoid perspective, using the mean values presented by the set of registers compiled for each species in regard to each of the selected environmental WorldClim layers (Bio6, Bio14, Bio15, Bio18 and Bio19). Bloomberg's *K* statistics analyses if a given character presents the amount of phylogenetic signal expected if it evolved under a Brownian motion model along a branch in a topology (Blomberg et al 2003). A *K* value not significantly different from 0 reveals the absence of phylogenetic signal, whereas a value of 1 pinpoints its presence. Likewise, a Pagel's ʎ value of 0 suggests that the trait is evolving independently in the phylogeny, a value of 1 assumes that the trait is evolving according a Brownian motion, and intermediate values report the amount of the effect of Brownian motion (Freckleton et al 2002). The Brownian motion null model tested by both, Bloomberg's K and Pagel's ʎ, accepts that direction and degree of change vary randomly over time, so that among-species variance increases with the square root of elapsed time (Pearman et al 2008). Thus, this model encompasses a phylogenetic signal model of evolution under drift (Cooper et al 2010).

Evolutionary characters mapping

Ancestral state character reconstruction (ASCR) is a widely employed method to map morphological and ecological traits onto a molecular phylogeny. We used a reduced version of the Bayesian phylogeny generated with amd+ddc by Robe et al (2010a) to perform ASCR for breeding sites data under a parsimony model, as implemented in Mesquite 2.75 (Maddison and Maddison 2011). Furthermore, consistency and retention index presented by each resource in this tree were also evaluated.

Results

Ecological Niche Modeling and range overlap

Average AUC values for the 25 generated species distribution models was 0.979 (minimum of 0.942 for *D. canalinea* and maximum of 0.993 for *D. nappae*), showing that the models were reliable concerning the current registers. As the objective of this work is not the modeling *per se*, the maps with the environmental suitability distribution were generally omitted, unless its depiction was considered important to our aims (see below).

Range overlap measures generated from binary maps for all the set of studied species after applying the mean MTP threshold as cut-off, varied from almost complete allopatry (R = 0.002- 0.009) [as found for *D. pavani* in regard to some species of the *tripunctata* (*D. mediopunctata* and *D. nappae*), *guarani* (*D. ornatifrons*), *cardini* (*D. neocardi* and *D. polymorpha*) or *annulimana* (*D. annulimana*) group species], to complete sympatry (R = 1), [as found in the comparison involving *D. cuaso*/*D. annulimana* and *D. cuaso*/*D. canalinea*]. Moreover, 84.33% of the pairwise comparisons presented range overlap values higher than 0.64, showing that partial sympatry appears to be very common for the evaluated species (Fig. S1).

Sister species pattern

The patterns of potential range overlap between sister species can be evaluated in Figure 1, which presents the intersection between the maps of environmental suitability for each sister pair. As can be seen, several of these species presented wide potential distribution across the Neotropics, with areas of suitability extending from the south of USA to the south of South America. It is interesting to note that all analyzed sister species exhibited areas of environmental suitability in the region of Florida (USA), in some cases in a spot disjunct to the remaining of the potential distribution. Since this area is putatively unoccupied by the species in question, it could represent the distribution of other species with similar environmental requirements (Raxworthy et al 2007), although it most likely reflects an absence of equilibrium in the distribution patterns, or an absence of occupation due to biotic interactions or insurmountable barriers.

(Place to Figure 1)

Among the seven pairs of sister species, range overlap measures (R) varied, but were higher than 0.86 for five of the seven pairwise comparisons (Fig. S1), endorsing the wide overlap for most of the evaluated sister pairs (Fig. 1). In this case, higher R values were found in the comparisons involving *D. griseolineata/D. maculifrons* (pair A in Figure S1) and *D. cuaso*/*paraguayensis* (pair D in Figure S1), whereas the lower values were found for the pair *D. gaucha/D. pavani* (pair G in figure S1) (Table 2).

(Place to Table 2)

Abiotic niche overlap

Environmental niche overlap measures ranged from 0.42 (*D. pavani*/*D. unipunctata*) to 0.99 (*D. mediostriata*/*D. pallidipennis*) for Hellinger's I and from 0.20 (*D. pavani*/*D. unipunctata*) to 0.90 (*D. mediostriata*/*D. pallidipennis*) for Schoener's D. In general, the frequency distribution of the pairwise comparisons of I were biased towards higher values, with 85% of the comparisons ranging between 0.72 and 0.98 (Fig. S1). At contrast, D's values were generally lower, and the frequency distribution of this niche overlap metric was centered on intermediate values, with 86.33% of the comparisons ranging between 0.42 and 0.86 (Fig. S1). This behavior is probably linked to the difference between these two metrics: whereas I considers the two ENMs as probability distributions, D implicitly assumes that suitability scores produced by Maxent are proportional to species abundance (Warren et al 2010).

Despite the tendency in regard to higher or intermediate abiotic overlap values, identity tests performed at the higher divergence levels showed a low number of pairwise comparisons rejecting the null hypothesis of niche equivalency. So, within clades 12, 22 and 32 of Robe et al (2010a) phylogeny [encompassing the *mediostriata*, the *cardini* and the *tripunctata* lineages of the main *tripunctata* lineage (Robe et al 2010b)], 26.67%, 33.33% and 38.60% of the pairwise comparisons, respectively, rejected the null hypothesis. A different pattern was, nevertheless, encountered for the evaluated *virilis-repleta* lineage species [encompassing clade 61 in Robe et al (2010a)], for which 86.67% of the pairwise comparisons rejected niche equivalency.

Sister species pattern

Comparisons between sister species showed two groups in regard to I and D metrics: one with lower values for both indices (encompassing the pairwise comparisons involving *D. gaucha/D. pavani* and *D. mediopunctata*/*D. unipunctata*) and the other with values higher than 0.93 and 0.68, respectively (encompassing the other five comparisons) (Fig. S1). In fact, after adopting a Bonferroni correction and a conservative criterion in which rejection of niche equivalency requires simultaneous significance for both, D and I comparisons, among the seven identity tests performed between sister pairs, only the two presenting lower D and I values rejected the null hypothesis (Fig. S2). Thus, niche similarity signs were found in the comparisons involving *D. griseolineata*/*D. maculifrons*, *D gasici/D. mesophragmatica*, *D. cuaso*/*D. paraguayensis*, *D. neocardini*/*D.polymorpha* and *D. bandeirantorum*/*D. pallidipennis*. At contrast, the hypothesis of niche similarity was rejected for both environmental niche overlap measures in the comparisons involving *D. gaucha*/*D. pavani* and *D. mediopunctata*/*D. unipunctata*.

Biotic Niche overlap

Biotic niche overlap values (C) ranged from 0 (for several pairwise comparisons) to 0.69 (between *D. griseolineata*/*D.mediopunctata*), with 61.6% of the comparisons distributed between 0 and 0.2 (Fig. S1). The simulated RA3 model showed a mean observed C value of 0.18, but as the simulated C value was even lower, (0.10, p = 0.0) the results indicate that resource sharing is higher than expected by chance, suggesting that across evolutionary history, competition has not led to a partition of resources (Gotelli and Graves 1996).

Sister species pattern

The analysis concerning the use of breeding sites among sister species showed that the evaluated pairs present highly dispersed similarity values, as measured by the Czekanowski index (C) (Fig. S1), although all the comparisons presented biotic niche overlap values significantly higher than expected by chance (p = 0). In this case, among the four sister pairs for which this measure could be obtained (see above), *D. cuaso*/*D.paraguayensis* and *D. griseolineata*/*D. maculifrons* presented the lower values, *D. bandeirantorum*/*D. pallidipennis* presented

intermediate values, whereas *D. neocardini*/*D. polymorpha* were the sister pairs sharing the higher similarity as concerns resource use (Table 2).

Phylogenetic signal

The results of the different sets of Partial Mantel correlation tests are shown in Table 3, where it can be seen that phylogenetic signal was detected for all evaluated abiotic and biotic niche overlap measures when the entire set of species was considered. Contrasting with the pattern found for the *Drosophila* subgenus taken as a whole, none significant correlation was found for the subsets of species encompassing the *tripunctata* and *virilis-repleta* lineages. The advantage of Partial Mantel tests in regard to its simpler and more commonly used version is that it permits to control for spatial autocorrelation (G_1 and G_2), so that potential PNC patterns are not obscured (Smouse et al 1986).

Nevertheless, using mean climatic values in the ENM devoid perspective, the phylogenetic signal dispersed. Bloomberg's *K* statistic varied from 0.096 to 0.326, and none performed test showed significant values. Pagel's ʎ varied from 0.00007 to 0.671, and only 2 environmental variables (Bio6 and Bio18) presented significant values (ʎ = 0.671, p = 0.0032; and ʎ = 0.613, p = 0.0019, respectively; critical p value after Bonferroni correction p = 0.01) (Table S1). Individual analyzes performed with the *tripunctata* and *viris-repleta* lineages did not present any significant value in both, Bloomberg's *K* and Pagel's ʎ tests (Table S1).

(Place to Table 3)

Breeding sites

Ancestral state character reconstruction suggested several autapomorphies and homoplasies regarding breeding sites usage. Putative autapomorphies were found, for example, in flowers of Annonaceae, Apocynaceae, Araceae, Asclepiadaceae, Bignoniaceae, Fabaceae, Heliconeaceae and Marantaceae by *D. cardinoides*; fruits of Cecropiaceae, and fungus of Polyporaceae and Tricholomataceae by *D. paraguayensis*; fungus of Auriculariaceae by *D. cuaso*; and fruits of Malvaceae and Phytolaccaceae by *D. griseolineata*. Furthermore, putative homoplasies were also recovered, for example, between *D. cardini* and *D. cardinoides* (for Malvaceae flowers), *D. cardinoides*, *D. cardini*, *D. paramediostriata* and *D. canalinea* (for Passifloraceae fruits) and *D. cardinoides*, *D. mediostriata*, *D. mediopicta*, *D. pallidipennis* and *D. paraguayensis* (for Oxalidaceae fruits).

The consistence (CI) and retention (RI) indexes exhibited by the evaluated resources (Fig. 2, autapomorphies were ommited) also revealed moderate to high levels of homoplasy for all analyzed characters. Fruits of Rutaceae, Moraceae and Myrtaceae had the lowest values of CI (CI = 0.14 and 0.17 respectively) whereas the higher CI values were of the order of 0.5, for fruits of Apocynaceae, Magnoliacee, Musaceae and Thymelaeaceae, flowers of Solanaceae and fungus of Marasmiaceae. As concerns RI, that corrects for autapomorphies, lower values (RI = 0.00) were found in fruits, flowers and fungus in 28 (out of 43) families (Fig. 2), and higher values were found for fruits of Musaceae (RI =0.67).

Even so, some interesting patterns could be evaluated through the ancestral state character reconstruction depicted in Figure 2: (1) the putative symplesiomorphy between all species of *tripunctata* lineage for the use of Arecaceae fruits, which seems to have been nevertheless lost in *D. guaru* and *D. pallidipennis*; (2) the several synapomorphies between *D. griseolineata* and *D. maculifrons* for the use of Fabaceae, Magnoliaceae and Moraceae fruits; and (3) the diverse homoplasies shared between members of the *cardini* group and *D. cuaso, D. paraguayensis* and/or *D. mediopunctata*. Besides, analysis of Figure 2 also reveals that *D. cardinoides*, *D. griseolineata, D. mediopunctata* and *D. paraguayensis* are the species showing the more generalist patterns in resources use.

(Place to Figure 2)

Discussion

Our phylogenetic comparative data showed some contrasting results according to the method and/or phylogenetic scale. This is a common outcome, since phylogenetic covariances may be greatly altered according to the examined trait, to the definition and measurement of niche adopted, to the method and even to the phylogenetic level assessed (Wiens and Graham 2005; Losos 2008; Pearman et al 2008; Cooper et al 2010; Wiens et al 2010; Peterson 2011). In this sense, Partial Mantel tests performed with the entire *Drosophila* subgenus evidenced phylogenetic signal for both abiotic and biotic niche overlap measures, although such a pattern could not be detected when less inclusive strategies were conducted or, in general, when ENM devoid tests were employed. Thus, no significant association was found when individual environmental variables were analyzed with Bloomberg's *K* statistic,

and when three of the five employed variables were analyzed with Pagel's ʎ even in the most inclusive strategy.

Revell et al (2008) recommended against the use of phylogenetic signal patterns in the inference of evolutionary processes, since they demonstrated through different sets of simulations that it may have the same behavior in distinct evolutionary processes. However, other authors argue that this approach could still be used, provided that it is backed up by complementary strategies showing that closely related species are more similar than based on their phylogenetic relationships (Losos 2008; Cooper et al 2010; Crisp and Cook 2012). Losos (2008), for example, states that the presence of phylogenetic signal is fundamental but insufficient to demonstrate PNC; otherwise, their absence would be enough to discard it. In this sense, in addition to the phylogenetic signal tests performed in ENM-based and ENM-devoid perspectives, for which the sole drift model provides a reasonable null hypothesis to test the hypothesis of adaptive radiation, we also performed niche identity tests, which are better covered by the niche retention model (see Cooper et al 2010 and Wiens et al 2010). Regarding the phylogenetic signal tests, absence or weak phylogenetic dependencies could either mean that the trait varies randomly across the phylogeny (absence of NC), or that it shows evolutionary stasis (strong NC) (Wiens et al 2010). Conversely, rejection of niche equivalency patterns in the niche identity tests means that ENMs generated for two species exhibit statistically significant ecological differences (absence of NC).

Generally speaking, our data seem to support the presence of phylogenetic signal at least in the most inclusive levels (as detected by the Partial Mantel tests) and niche equivalency in several levels (as detected by the identity tests). In this case, although the presence of phylogenetic signals in face of a general pattern of niche equivalency could be interpreted in two ways [related species may display similar patterns either due to common evolutionary constraints (PNC) or because they share similar environments, being subject to common selection pressures in face of dispersal limitations (phylogenetic structured adaptation) (Freckleton and Jetz 2009; Cooper et al 2010; Kellermann et al 2012a, b)], the significant results obtained in the partial Mantel tests after controlling for geographic overlap attests that the detected phylogenetic signal can be in fact attributed to PNC instead of spatial autocorrelation patterns. Thus, the phylogenetic signal detected for the *Drosophila* subgenus through the Partial Mantel test for all abiotic (ENM-based) and biotic niche

overlap measures and through Pagel's ʎ (ENM-devoid) for two of the included variables (Bio 6 and Bio 18), added to several cases of significant niche identity pairwise tests, suggest that PNC rather than ND was the general pattern for the evaluated species.

Nevertheless, despite the detection of phylogenetic dependencies when the entire *Drosophila* subgenus is considered, this result was not maintained once limited clades within it are taken into account. For example, when the *tripunctata* or the *virilis-repleta* lineages were individually considered, no phylogenetic signal was detected. This phylogenetic scale problem may be an outcome of the power of Mantel tests in relation to tree size (Harmon and Glor 2010; Hardy and Pavoine 2012) or may reflect the fact that phylogenetic associations have arisen early in the phylogeny, so that both lineages that have been evolved separately for more than 30 Ma (Robe et al 2010a), could have coursed different evolutionary trajectories, with between-clade convergence being a rare phenomenon (see Losos et al 2008).

These results, together with the conclusions of Morales-Hojas and Vieira (2012) highlight that Throckmorton's hypothesis for the evolution of the *Drosophila* subgenus should be seen with caution, since this lineage does neither present general signs of increased speciation rates (as found by Morales-Hojas and Vieira 2012), nor shows an absence of phylogenetic signal (as found here) expected for adaptive radiations (Blomberg et al 2003). However, it is important to keep in mind that a radiation might yet have occurred through diversification in one or more traits not here considered. Furthermore, taking a closer look to the depicted scenario, we can see that there appear to be cases applying also to a scenario of ND.

The *tripunctata* lineage or radiation

Although sample size has been pointed as a potential problem in Mantel tests (Harmon and Glor 2010; Hardy and Pavoine 2012), the number of evaluated species for the *tripunctata* lineage was not straightforwardly reduced in regard to that of the *Drosophila* subgenus (n = 19 and 25, respectively). Thus, statistical power seems not to be the best explanation for the absence of any significant phylogenetic signal at this scale. As the identity tests showed that this lineage also followed the general pattern of its subgenus, with a low number of closely related species rejecting the null hypothesis of niche equivalency, niche retentionmight be the best explanation for the absence of phylogenetic signal (Wiens et al 2010). Niche retention might also explain

the rejection of niche conservatism in Robe et al (2010b), which tested PNC in the *tripunctata* lineage through an evaluation of the correlation between pairwise phylogenetic and climatic Euclidean distances in a sole ENM devoid perspective. These authors also did not find any significant phylogenetic signal for this lineage, although they alluded to ND to explain this general pattern.

The ASCR analysis regarding the use of different resources (Fig. 2) showed that several species of this group are generalist, adapted to the use of flower, fruits and fungus as breeding sites (e.g. *D. cardinoides*, *D. cuaso* and *D. paraguayensis*). We also found a concentration in the distribution of autapomorphies for particular species (e. g. *D. cardinoides*, *D. mediopunctata* and *D. griseolineata*), but this is probably an artifact of deficient sampling [despite the use of the largest dataset as yet registered for breeding sites in this group by Gottschalk (2008)]. An interesting pattern was, nevertheless, obtained in the ASCR for Arecaceae fruits exploitation, which appears to be an ancestral property for the *tripunctata* lineage species, yet maintained in most of the group.

Following the general pattern of the lineage, among the five evaluated sister pairs, only one rejected the null hypothesis of niche equivalency. Most of the species pairs that did not reject niche similarity have wide range overlap measures ($R > 0.75$) and different divergence ages [*D. cuaso/D. paraguayensis* at about 1.78 MYA, *D. maculifrons/D. griseolineata* at about 4.6 MYA, *D. neocardini/D. polymorpha* at about 7.66 MYA and *D. bandeirantorum*/*D. pallidipennis* at about 17.2 MYA (Robe et al 2010a)], suggesting a decrease in the evolutionary rates of the evaluated traits in more recent past, compatible with a niche filling model (Cooper et al 2010). In this case, the speciation mode may be (1) geographic, with allopatric divergence in face of NC and further expansion leading to a secondary contact (Fig. 3), or (2) ecological, with sympatric divergence in non-environmental niche properties. Since the recorded resource overlap patterns suggested that across evolutionary history competition has not led to a partition between the studied species, the identity of these putative non-environmental properties still remains to be addressed. The pattern detected for the fifth sister-pair (*D. mediopunctata/D. unipunctata*) is nevertheless completely different since niche conservatism is rejected despite a significant range overlap ($R = 0.69$). So, it is possible to hypothesize that this speciation was ecological, occurring through abiotic niche divergence either in sympatry, as a form of reducing competition, or in allopatry with secondary contact. Despite the uncertainties related to this scenario, it

is possible to see that *Drosophila* species could greatly contribute to the controversy regarding the impact geographical versus ecological forces have in the evolution of different species (Wiens 2004; Rundle and Nosil 2005; Wang et al 2013).

(Place to Figure 3)

The *virilis-repleta* lineage

The *virilis-repleta* lineage seems to follow a pattern that is different from that found for the subgenus *Drosophila* and *tripunctata* lineage. This lineage did not present a significant phylogenetic signal in Partial Mantel tests, a result that could be related to the small number of terminal nodes (Harmon and Glor 2010), although ND was corroborated by the identity tests. However, the low number of species included here makes the general conclusion about this group questionable, once other authors have found a NC pattern for some of its groups (Robe et al 2013; Kellerman et al 2012a; Oliveira et al 2012). In the sister species perspective, only two comparisons regarding the *virilis-repleta* lineage could be accordingly included, and these present contrasting patterns: (1) *Drosophila gasici* and *D. mesophragmatica* diverged around 3 MYA (Robe et al 2010a) and seem to be sympatric in most of their distribution (R= 0.86; Fig. S1). Abiotic identity tests performed with this pair of species were not able to reject the null hypothesis of niche equivalency (Fig. S2), suggesting geographical speciation, with NC in allopatry and secondary contact; or ecological speciation, with ND in non-environmental niche properties in sympatry. (2) *D. gaucha* and *D. pavani* diverged only about 1.25 MYA (Robe et al 2010a) and are able to hybridize under laboratory conditions (Godoy-Herrera et al 2005). Since range overlap between these species is small (R=0.43; Fig. S1-G) and they also rejected niche equivalency, it is possible to suggest allopatric speciation with niche divergence as the putative divergence pattern (Fig. 3). The allopatric speciation hypothesis suggested in both cases agrees with the reasoning provided by Mota et al (2008) that distinct Andean refuge areas formed during the Pleistocene (Salgado-Laboriau 1994), affecting the speciation of the *D. mesophragmatica* group.

Conclusion

So, despite the contrasting results presented by the different set of analyzes, only the integrative approach using different phylogenetic signal tests, in ENM and

ENM-devoid perspectives, added to the wide set of identity tests lead to a refined picture of niche dynamics patterns within the *Drosophila* subgenus. Contrasting patterns were evidenced for each of the included lineages, reflecting that the way effects of dispersal abilities and magnitude of fitness differences between habitats interact to determine whether niche properties expands or contracts (Kirkpatrick and Barton 1997, reviewed in Pearman et al 2007) also appear to be phylogenetically autocorrelated. Finally, we present subsides to suggest PNC at the environmental and trophical resource scales in the Neotropical *Drosophila* subgenus and its *tripunctata* lineage. Alghough PNC was not supported for the *virilis-repleta* lineage, increase in sampling size is needed in order to clarify the evolutionary scenary for this group.

Acknowledgements

We are grateful to Dr. Daniel AS Graichen, Dr. Elgion LS Loreto, Dr. Fernando F Franco and Francisco CC Barreto for their contributions. We also are grateful to the Brazilian Funding Agency Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support (Universal-CNPq 14/2013, process number 471174/2013-0).

References

Bächli G (2012) TaxoDros: The Database on Taxonomy of Drosophilidae, v. 1.03, Database 2011/12. Available from http://www.taxodros.uzh.ch/

Bächli G (2014) TaxoDros: The Database on Taxonomy of Drosophilidae, v. 1.04, Database 2014/1. Available from http://www.taxodros.uzh.ch/

Bächli G, Vilela CR, Ratcov V (2000) Morphological differences among *Drosophila paraguayensis* Duda 1927 and its close relatives (Diptera, Drosophilidae). Mitt Schweiz Entomol Ges 73: 67–92

Blauth ML, Gottschalk MS (2007) A novel record of Drosophilidae in the Cerrado biome of the Mato Grosso, west-central Brazil. Drosoph Inf Serv 90: 90-96

Blomberg SP, Garland TJr, Ives AR (2003) Testing for phylogenetic signal in comparative data: behavior traits are more labile. Evolution 57(4): 717–745

Chaves NB, Tidon R (2008) Biogeographical aspects of drosophilids (Diptera, Drosophilidae) of the Brazilian savanna. Rev Bras Entomol 52(3): 340-348

Cooper N, Jetz W, Freckleton RP (2010) Phylogenetic comparative approaches for studying niche conservatism. J Evol Biol 23(12): 2529-2539

Coyne JA, Orr HA (2004) Speciation. Sinauer Associates, Sunderland

De Ré FC (2011) Filogeografia de populações Brasileiras de *Drosophila maculifrons* e *Drosophila griseolineata*. Dissertation, Universidade Federal de Santa Maria

De Toni DC, Hofmann PRP, Valente VLS (2001) First record of *Zaprionus indianus* (Diptera, Drosophilidae) in the state of Santa Catarina, Brazil. Biotemas 14(1): 71-85

Durando CM, Baker RH, Etges WJ et al (2000) Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. Mol Phylogenet Evol 16: 296-307

Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst 40: 677-697

Freckleton RP, Harvey PH, Pagel M (2002) Phylogenetic analysis and comparative data: a test and review of evidence. Am Nat 160(6): 712-726

Freckleton RP, Jetz W (2009) Space versus phylogeny: disentangling phylogenetic and spatial signals in comparative data. Proc R Soc B 276: 21-30

Frota-Pessoa O (1952) Flower-feeding Drosophilidae. Drosoph Inf Serv 26: 101-102

Glennon KL, Rissler LJ, Church SA (2012) Ecogeographic isolation: a reproductive barrier between species and between cytotypes in *Houstonia* (Rubiaceae). Evol Ecol 26: 909–926

Godoy-Herrera R, Burnet B, Connolly K (2005) Hybrid disadvantage in the larval foraging behaviour of the two neotropical species of *Drosophila pavani* and *Drosophila gaucha*. Genetica 124: 33–40

Godsoe W (2010) Regional variation exaggerates ecological divergence in niche models. Syst Biol 59(3): 298–306

Gotelli NJ, Graves GR (1996) Null Models in Ecology. Smithsonian Institution Press, Washington

Gotelli NJ, Entsminger GL (2001) EcoSim: Null models software for ecology. Version 7.0. Acquired Intelligence Inc. and Kesey-Bear. Available from http://homepages.together.net/~gentsmin/ecosim.htm

Gottschalk MS (2008) Utilização de recursos tróficos por espécies Neotropicais de Drosophilidae (Diptera). D. Phil Thesis, Universidade Federal do Rio Grande do Sul

Gottschalk MS, Hofmann PRP, Valente VLS (2008) Diptera, Drosophilidae: historical occurrence in Brazil. Check List 4(4): 485–518

Gottschalk MS, Bizzo L, Döge JS et al (2009) Drosophilidae (Diptera) associated to fungi: differential use of resources in anthropic and Atlantic Rain Forest areas. Iheringia Ser Zool 99(4): 442-448

Graham CH, Ron SR, Santos JC et al (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. Evolution 58(8): 1781–1793

Grimaldi DA (1990) A phylogenetic, revised classification of the genera in the Drosophilidae (Diptera). Bull Am Mus Nat Hist 197: 1–13

Hardy OJ, Pavoine S (2012) Assessing phylogenetic signal with measurement error: A comparison of Mantel tests, Blomberg et al.'s K, and phylogenetic distograms. Evolution 66(8): 2614–2621

Harmon LJ, Glor GE (2010) Poor statistical performance of the Mantel test in phylogenetic comparative analyses. Evolution 64(7): 2173–2178

Hatadani LM, McInerney JO, de Medeiros HF et al (2009) Molecular phylogeny of the *Drosophila tripunctata* and closely related species groups (Diptera: Drosophilidae). Mol Phylogenet Evol 51: 595-600

Heed WB (1956) Apuntes sobre la ecologia y la dispersion de los Drosophilidae (Diptera) de El Salvador. Comunicaciones del Instituto Tropical de Investigaciones Cientificas El Salvador San Salvador 5: 59-74

Hernandez PA, Graham CH, Master LL et al (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29: 773-785

Hijmans RJ, Cameron SE, Parra JL (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25: 1965-1978

Kamilar JM, Cooper N (2013) Phylogenetic signal in primate behavior, ecology and life history. Phil Trans R Soc 368

Kamino LHY, Stehmann JR, Amaral S et al (2012) Challenges and perspectives for species distribution modelling in the neotropics. Biol Lett 8: 324–326

Kellerman V, van Heerwaarden B, Sgrò CM et al (2009) Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. Science 325: 1244-1246

Kellerman V, Overgaard J, Hoffmann AA et al (2012a) Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically. PNAS 109(40): 16228-16233

Kellermann V, LoeschckeV, Hoffmann AA et al (2012b) Phylogenetic constraints in key functional traits behind species' climate niches: patterns of desiccation and cold resistance across 95 *Drosophila* species. Evolution 66: 3377–3389

Kozak KH, Wiens JJ (2006) Does niche conservatism promotes speciation: a case study in north American salamanders. Evolution 60(12): 2604-2621

Kozak KH, Wiens JJ (2010) Niche conservatism drives elevational diversity patterns in appalachian salamanders. Am Nat 176(1): 40-54

Litsios G, Pellisier L, Forest F et al (2012) Trophic specialization influences the rate of environmental niche evolution in damselfishes (Pomacentridae). Proc R Soc B 279: 3662-3669

Lobo JM, Tognelli MF (2011) Exploring the effects of quantity and location of pseudo-absences and sampling biases on the performance of distribution models with limited point occurrence data. J Nat Conserv 19: 1–7

Losos JB (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecol Lett 11: 995–1007

Maddison WP, Maddison DR (2011) Mesquite: a modular system for evolutionary analysis. Version 2.75. Available from http://mesquiteproject.org

Markow TA, O'Grady P (2008) Reproductive ecology of *Drosophila*. Funct Ecol 22: 747-759

Martins MB, Santos RCO (2007) Sítios de criação de *Drosophila* na Reserva Mocambo, Belém. Pará. In: Mocambo: Diversidade e Dinâmica biológica da Área de Pesquisa Ecológica do Guampa (APEG). Belém, Museu P. Emilio Goeldi/Embrapa Amazonia oriental, pp 318-331

Mata TA, Tidon R, Cortes LG et al (2010) Invasive and flexible: niche shift in the drosophilid *Zaprionus indianus* (Insecta, Diptera). Biol Invasions 12: 1231-1241

McInerny GJ, Etienne RS (2012a) Ditch the niche – is the niche a useful concept in ecology or species distribution modelling? J Biogeogr 39: 2096-2102

McInerny GJ, Etienne RS (2012b) Stitch the niche – a practical philosophy and visual schematic for the niche concept. J Biogeogr 39: 2103-2111

McInerny GJ, Etienne RS (2012c) Pitch the niche – taking responsibility for the concepts we use in ecology and species distribution modeling. J Biogeogr 39: 2112-2118

Medina-Muñoz MC, Godoy-Herrera R (2004) Dispersal and prepupation behavior of Chilean sympatric *Drosophila* species that breed in the same site in nature. Behav Ecol 15(2): 316-322

Morales-Hojas R, Vieira J (2012) Phylogenetic patterns of geographical and ecological diversification in the subgenus *Drosophila*. PLoS ONE 7(11): e49552

Mota NR, Robe LJ, Valente VLS et al (2008) Phylogeny of the *Drosophila mesophragmatica* group (Diptera, Drosophilidae): an example of Andean evolution. Zool Sci 25(5): 526-532

Münkemüller T, Lavergne S, Bzeznik B et al (2012) How to measure and test phylogenetic signal. Methods Ecol Evol 3: 743-756

O'Grady PM, Markow TA (2009) Phylogenetic taxonomy in *Drosophila*: problems and prospects. Fly 3(1): 10-14

Olalla-Tárraga MA, McInnes L, Bini LM et al (2011) Climatic niche conservatism and the evolutionary dynamics in species range boundaries: global congruence across mammals and amphibians. J Biogeogr 38: 2237-2247

Oliveira DCSG, Almeida FC, O'Grady PM (2012) Monophyly, divergence times, and evolution of host plant use inferred from a revised phylogeny of the *Drosophila repleta* species group. Mol Phylogenet Evol 64: 533–544

Pagel M (1999) Inferring the historical patterns of biological evolution. Nature 401: 877–884

Pearman PB, Guisan A, Broennimann O et al (2008) Niche dynamics in space and time. Trends Ecol Evol (Amst.) 23(3): 149- 158

Peterson AT, Soberón J, Sánchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. Science 285: 1265-1267

Peterson AT (2011) Ecological niche conservatism: a time-structured review of evidence. J Biogeogr 38: 817–827

Philippe H, Brinkmann H, Lavrov VD et al (2011) Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol 9(3): 1-10

Phillips JS, Dudík M, Schapire RE (2004) A maximum entropy approach to species distribution modeling**.** In: *Proceedings of the Twenty-First International Conference on Machine Learning.* AMC Press, New York, pp 655-662

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Modell 190: 231-259

Phillips SJ, Dudik M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31: 161-175

Pipkin SB, Rodriguez RL, León J (1966) Plant host specificity among flower-feeding neotropical *Drosophila* (Diptera: Drosophilidae). Am Natur 100: 135-156

Pyron A, Burbrink FT (2009) Lineage diversification in a widespread species: roles for niche divergence and conservatism in the common kingsnake, *Lampropeltis getula*. Mol Ecol 18: 3443-3457

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org

Raxworthy CJ, Ingram CM, Rabibisoa N et al (2007) Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. Syst Biol 56(6): 907-923

Revell LJ, Harmon LJ, Collar DC (2008) Phylogenetic signal, evolutionary process and rate. Syst Biol 57(4): 591-601

Revell LJ (2012) Phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3: 217-223

Robe LJ, Valente VLS, Budnik M et al (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

Robe LJ, Loreto ELS, Valente VLS (2010a) Radiation of the ,,*Drosophila*“ subgenus (Drosophilidae, Diptera) in the Neotropics. J Zoolog Syst Evol Res 48(4): 310-321

Robe LJ, Valente VLS, Loreto ELS (2010b) Phylogenetic relationships and macro-evolutionary patterns within the *Drosophila tripunctata* “radiation” (Diptera: Drosophilidae). Genetica 138: 725-735

Robe LJ, de Ré FC, Ludwig A et al (2013) The *Drosophila flavopilosa* species group (Diptera, Drosophilidae): an array of exciting questions. Fly 7(2): 59-69

Rocchini D, Hortal J, Lengyel S (2011) Accouting for uncertainty when mapping species distribution: the need for map of ignorance. Prog Phys Geogr 35(2): 211-226

Roque F, Tidon R (2008) Eight new records of drosophilids (Insecta; Diptera) in the Brazilian savanna. Drosoph Inf Serv 91: 94- 98

Rosenberg MS, Anderson CD (2011) PASSaGE: Pattern Analysis, Spatial Statistics and Geographic Exegesis. Version 2. Methods Ecol Evol 2(3): 229-232

Rundle HD, Nosil P (2005) Ecological speciation. Ecol Lett 8: 336-352

Saavedra CCR, Callegari-Jaques SM, Napp M et al (1995a) A descriptive and analytical study of four neotropical drosophilid communities. J Zoolog Syst Evol Res 33: 62-74

Saavedera CCR, Valente VLS, Napp M (1995b) An ecological/genetic approach to the study of enzymatic polymorphisms in *Drosophila maculifrons*. Rev Bras Genet 18(2): 147-164

Salgado-Laboriau ML (1994) O Período Quaternário. In: História Ecológica da Terra. Editora Edgard Blücher, São Paulo, pp 255–280

Santos RCO, Vilela CR (2005) Breeding sites of Neotropical Drosophilidae (Diptera). IV. Living and fallen flowers of *Sessea brasiliensis* and *Cestrum* spp. (Solanaceae). Rev Bras Entomol 49(4): 544-551

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extension of the Mantel Test of matrix correspondence. Syst Zool 35(4): 627-632

Stockwell RB, Peterson AT (2002) Effects of sample size on accuracy of species distribution models**.** Ecol Model 148(1): 1-13

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

Tatarenkov A, Ayala FJ (2001) Phylogenetic relationships among species groups of the *virilis-repleta* radiation of *Drosophila.* Mol Phylogenet Evol 21(2): 327-331

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

Tosi D, Martins M, Vilela CR et al (1990) On a new cave-dwelling species of bat-guano-breeding *Drosophila* closely related to *D. repleta* Wollaston (Diptera, Drosophilidae). Rev Bras Genet 13(1): 19-31

Val FC, Marques MD, Vilela CR (1981) Drosophilidae of Neotropical region. In: Ashburner M, Carson HL, Thompson JN (eds.) The genetics and biology of *Drosophila*. Academic Press, Orlando, pp 123-168

Valadão H, Hay JDV, Tidon R (2010) Temporal dynamics and resource availability for drosophilid fruit flies (Insecta, Diptera) in a gallery forest in the Brazilian savanna. Int J Ecol 1: 1-7

Wang IJ, Glor RE (2013) Quantifying the roles of ecology and geography in spatial genetic divergence. Ecol Lett 16: 175–182

Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62(11): 2868–2883

Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33: 607-611

Wiens JJ (2004) Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. Evolution 58(1): 193–197

Wiens JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. Annu Rev Ecol Evol Syst 36: 519–539

Wiens JJ, Ackerly DD, Allen AP et al (2010) Niche conservatism as an emerging principle in ecology and conservation biology. Ecol Lett 13: 1310–1324

Wisz MS, Hijmans RJ, Peterson AT et al (2008) Effects of sample size on the performance of species distribution models. Divers Distrib 14: 763-773.

Yoder JB, Clancey E, Roches SD et al (2010) Ecological opportunity and the origin of adaptive radiations. J Evol Biol 23: 1581-1596

Yotoko KSC, Medeiros HF, Solferini VN et al (2003) A molecular study of the systematics of the *Drosophila tripunctata* group and the *tripunctata* radiation. Mol Phylogenet Evol 28: 614–619

Tables e figures

Table 1: Taxonomic distribution of the species selected for the analysis, with their number of unique occurrence points (N).

Lineage	Group	Species	N
tripunctata	*guarani*	*D. guaru* Dobzhansky and Pavan 1943	22
		D. ornatifrons Duda 1927	40
	guarumuru	*D. griseolineata* Duda 1927	67*
		D. maculifrons Duda 1927	56*
	tripunctata	*D. bandeirantorum* Dobzhansky and Pavan 1943	60
		D. cuaso Bächli, Vilela and Ratcov 2000	13
		D. mediopicta Frota-Pessoa 1954	32
		D. mediopunctata Dobzhansky and Pavan 1943	60
		D. mediostriata Duda 1925	62
		D. nappae Vilela, Valente and Basso-da-Silva 2004	22
		D. paraguayensis Duda 1927	34
		D. paramediostriata Townsend and Wheeler 1955	20
		D. roehrae Pipkin and Heed 1964	22
		D. unipunctata Patterson and Mainland 1943	24
	cardini	*D. cardini* Sturtevant 1916	204
		D. cardinoides Dobzhansky and Pavan 1943	115
		D. neocardini Streisinger 1946	39
		D. polymorpha Dobzhansky and Pavan 1943	115
	pallidipennis	*D. pallidipennis* Dobzhansky and Pavan 1943	69
virilis-repleta	*mesophragmatica*	*D. gasici* Brncic 1957	21
		D. gaucha Jaeger and Salzano 1953	22
		D. mesophragmatica Duda 1927	22
		D. pavani Brncic 1957	45
	annulimana	*D. annulimana* Duda 1927	26
	canalinea	*D. canalinea* Patterson and Mainland 1944	28

* data extracted from De Ré (2011)

Table 2: Niche and range overlap indices for the evaluated pairs of sister species. I = Hellinger's I, D = Schoener's D, R = range overlap, C= Czekanowski's index. The significance of each measure is given between brackets, where bold values are significant after a Bonferroni correction.

Sister species pair	I	D	R	C
D. griseolineata/*D. maculifrons*	0.95 (0.06)	0.77 (0.05)	0.97	0.25
D. neocardini/*D.polymorpha*	0.97 (0.89)	0.82 (0.65)	0.91	0.54
D. mediopunctata/*D. unipunctata*	0.75 (**0.00**)	0.48 (**0.00**)	0.69	-
D. cuaso/*D. paraguayensis*	0.93 (0.50)	0.71 (0.37)	0.92	0.15
D. bandeirantorum/*D. pallidipennis*	0.94 (0.03)	0.73 (**0.01**)	0.90	0.4
D gasici/D. mesophragmatica	0.93 (0.02)	0.70 (**0.01**)	0.86	-
D. gaucha/*D. pavani*	0.74 (**0.00**)	0.43 (**0.00**)	0.43	-

Table 3: Partial Mantel test performed in order to infer the presence of a phylogenetic signal for the environmental (evaluated through D and I) or the biotic niche overlap measures (evaluated through C), controlling for geographical distance (evaluated through G_1 or G_2) in three phylogenetic scales: the entire *Drosophila* subgenus, the *tripunctata* radiation and the *virilis-repleta* radiation.

	Matrices	N	Correlation	p value
Drosophila subgenus	PxIxG_1	24	0.33	**<0.001**
	PxIxG_2	24	0.33	**<0.001**
	PxDxG_1	24	0.44	**<0.001**
	PxDxG_2	24	0.30	**<0.01**
	PxCxG_1	12	0.37	**<0.01**
	PxCxG_2	20	0.44	**<0.001**
	AxIxG_1	24	0.31	**<0.001**
	AxIxG_2	24	0.32	**<0.001**
	AxDxG_1	24	0.42	**<0.001**
	AxDxG_2	24	0.42	**<0.001**
	AxCxG_1	20	0.38	**<0.01**
	AxCxG_2	20	0.39	**<0.01**
tripunctata lineage	PxIxG_1	18	-0.01	0.55
	PxIxG_2	18	-0.05	0.71
	PxDxG_1	18	0.04	0.30
	PxDxG_2	18	-0.01	0.51
	PxCxG_1	17	0.05	0.31
	PxCxG_2	17	0.05	0.30
	AxIxG_1	18	0.01	0.47
	AxIxG_2	18	0.01	0.47
	AxDxG_1	18	0.03	0.35
	AxDxG_2	18	0.03	0.35
	AxCxG_1	17	0.07	0.23
	AxCxG_2	17	0.06	0.24
virilis-repleta lineage*	PxIxG_1	6	0.33	0.15
	PxIxG_2	6	0.28	0.22
	PxDxG_1	6	0.41	0.11
	PxDxG_2	6	0.36	0.15
	PxCxG_1	3	-	-
	PxCxG_2	3	-	-
	AxIxG_1	6	0.31	0.19
	AxIxG_2	6	0.25	0.25
	AxDxG_1	6	0.39	0.12
	AxDxG_2	6	0.33	0.17
	AxCxG_1	3	-	-
	AxCxG_2	3	-	-

P = Phylogenetic patristic distance, A = divergence time, I = Hellinger's I, D = Schoener's D, G_1 = mean geographical coordinate, G_2 = mean of pairwise distance, N = number of species. Significant correlations after a Bonferroni correction are depicted in bold.

Table S1: Phylogenetic signal measured through Bloomberg's *K* and Pagel's ʎ statistics using the mean values presented by the set of registers compiled for each species in regard to the selected environmental variables.

		Bloomberg's *K*		Pagel's ʎ			
		K	P	ʎ	logL	logL0	p
Drosophila subgenus	Bio6	0.326	0.068	0.67121	-112.183	-116.531	**0.003**
	Bio14	0.175	0.482	0.36548	-101.747	-103.118	0.098
	Bio15	0.096	0.924	0.11122	-92.079	-92.1849	0.645
	Bio18	0.211	0.340	0.61281	-143.722	-148.518	**0.002**
	Bio19	0.160	0.594	0.00007	-130.383	-130.383	1.000
tripunctata lineage	Bio6	0.283	0.874	0.00007	-76.883	-76.883	1.000
	Bio14	0.467	0.427	0.00007	-72.854	-72.854	1.000
	Bio15	0.581	0.152	0.00007	-58.599	-58.599	1.000
	Bio18	0.509	0.325	0.20328	-92.789	-92.993	1.000
	Bio19	0.365	0.683	0.00007	-98.491	-98.490	1.000
virilis-repleta lineage	Bio6	0.710	0.035	0.86343	-30.379	-30.928	0.295
	Bio14	0.083	0.868	0.00007	-25.680	-25.680	1.000
	Bio15	0.063	0.957	0.00007	-25.547	-25.547	1.000
	Bio18	0.179	0.546	0.00006	-39.547	-39.547	1.000
	Bio19	0.356	0.266	0.79809	-31.457	-31.359	1.000

Environmental variables: Bio6: Minimum temperature of coldest month, Bio 14: Precipitation of driest month, Bio 15: Precipitation seasonality (coefficient of variation), Bio18: Precipitation of warmest quarter, Bio19: Precipitation of coldest quarter. Significant correlations after a Bonferroni correction are depicted in bold.

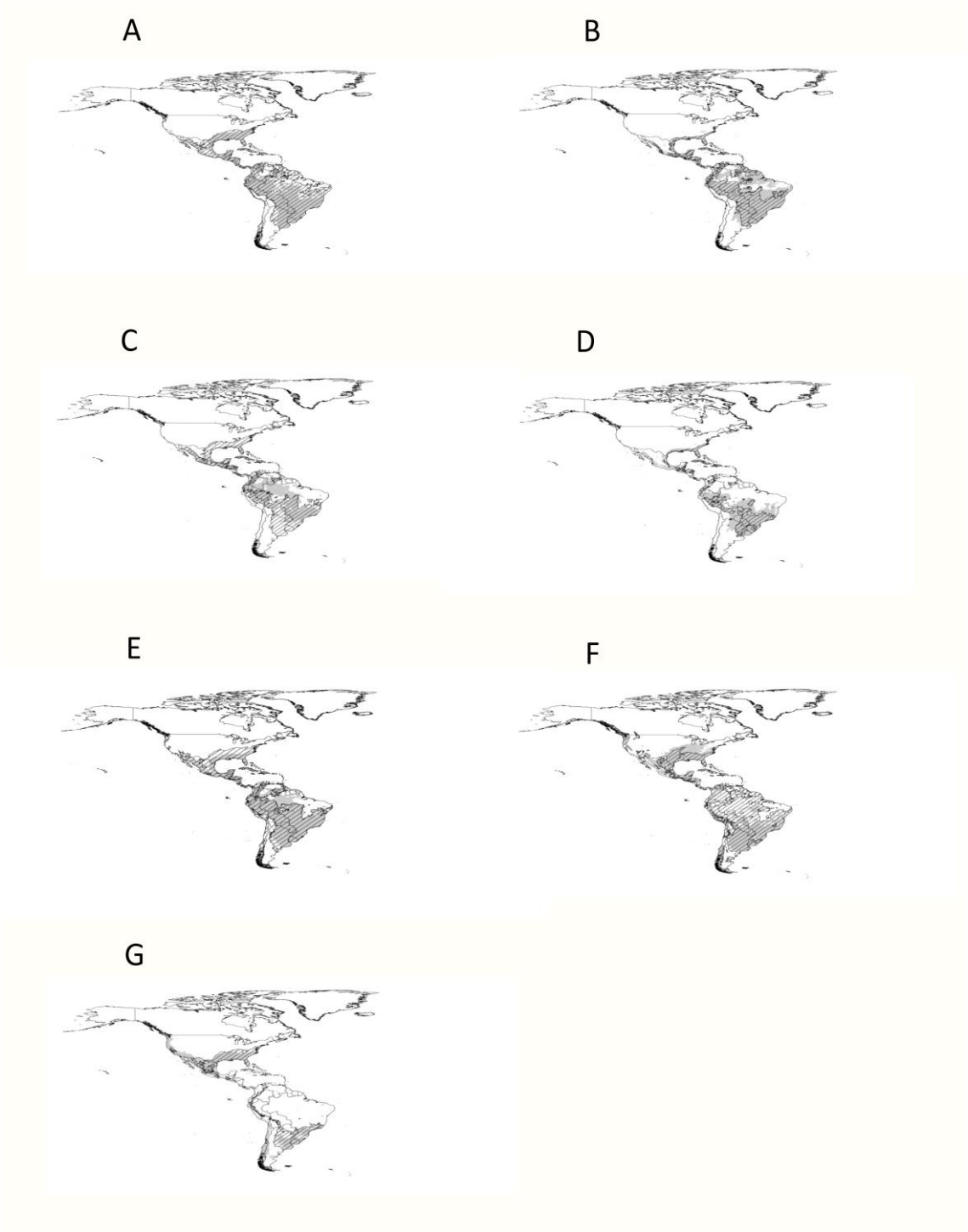

Figure 1: Overlapped environmental suitability models of the seven pairs of sister species. (A) *D. maculifrons*/*D. griseolineata*, (B) *D. neocardini*/*D. polymorpha*, (C) *D. mediopunctata*/*D. unipunctata,* (D) *D. cuaso*/*D. paraguayensis*, (E) *D. bandeirantorum*/*D. pallidipennis*, (F) *D. gasici*/*D. mesophragmatica*, (G) *D. gaucha*/*D. pavani*. In each case, the modeled potential distribution of the first mentioned species is represented by diagonal lines, whereas the predicted presence areas of the second species are presented in gray. In this sense, gray areas with diagonal lines are regions of potential range overlap.

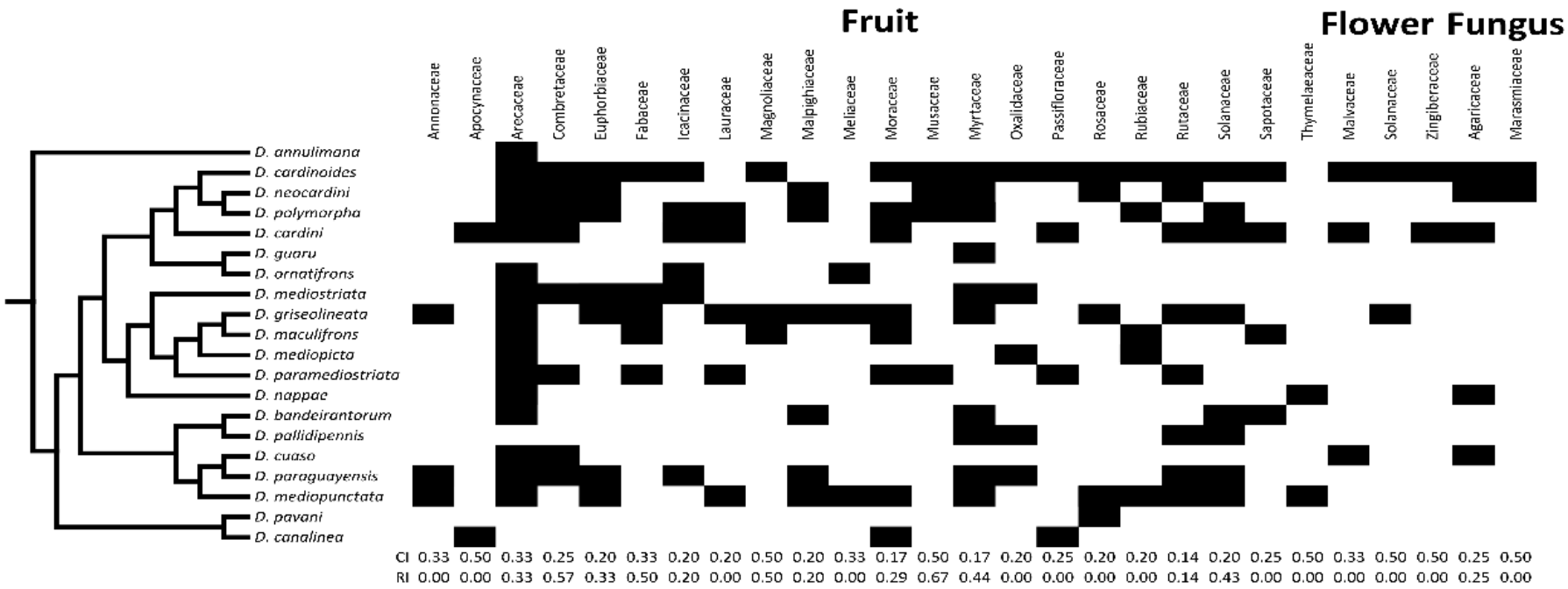

Figure 2: Breeding sites recorded to the *Drosophila* subgenus species. Black squares represent recorded emergence from the resource. Consistence (CI) and Retention (RI) indexes for each resource are shown below the presence/absence matrix. The tree shown in left was modified from Robe et al (2010a), with emphasis in the *tripunctata* radiation. Autapomorphies were omitted.

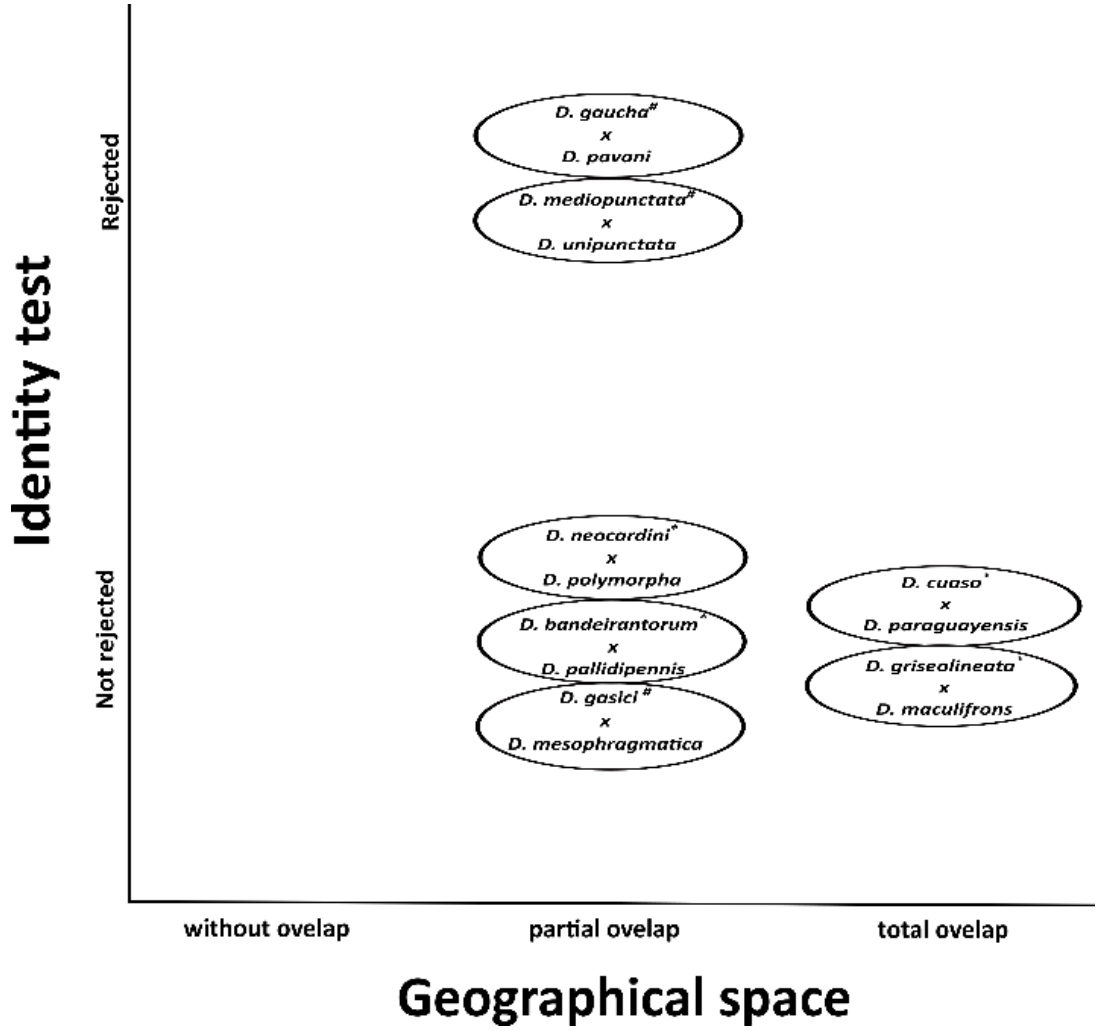

Figure 3: Niche equivalency and environmental overlap between sister species (adapted from Graham et al 2004). When range overlap value between sister species was higher than 0.96, they were considered entirely sympatric. * and # refer to biotic overlap measures: "*" depicts significant resource overlap and **#** refers to absence of information.

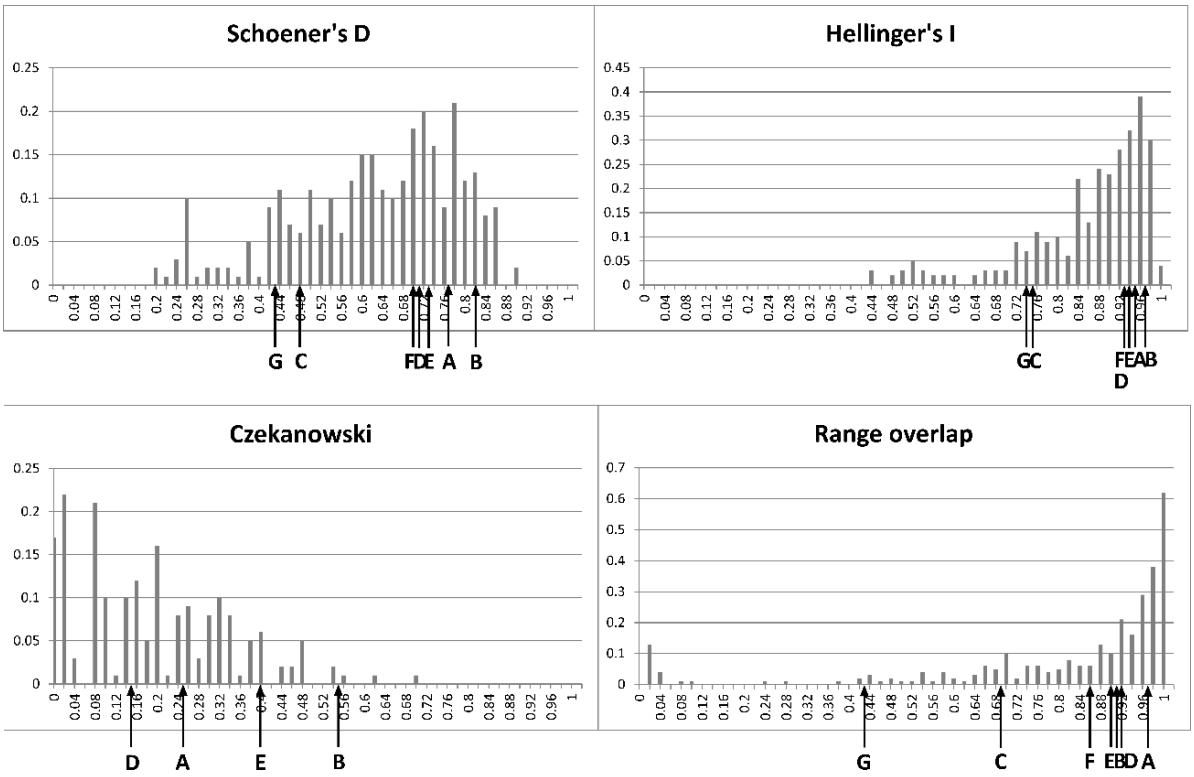

Figure S1: Relative frequencies distribution of pairwise comparisons regarding abiotic (measured through Schoener's D and Hellinger's I statistics), biotic (measured through Czekanowski's index) and range overlap values. Vertical bars represent the relative frequency of each overlap value among the set of pairwise comparisons. Arrows depict the values obtained for each of the evaluated sister species pairs, (A) *D. maculifrons*/*D. griseolineata*, (B) *D. neocardini*/*D. polymorpha*, (C) *D. mediopunctata*/*D. unipunctata,* (D) *D. cuaso*/*D. paraguayensis*, (E) *D. bandeirantorum*/*D. pallidipennis*, (F) *D. gasici*/*D. mesophragmatica*, (G) *D. gaucha*/*D. pavani*.

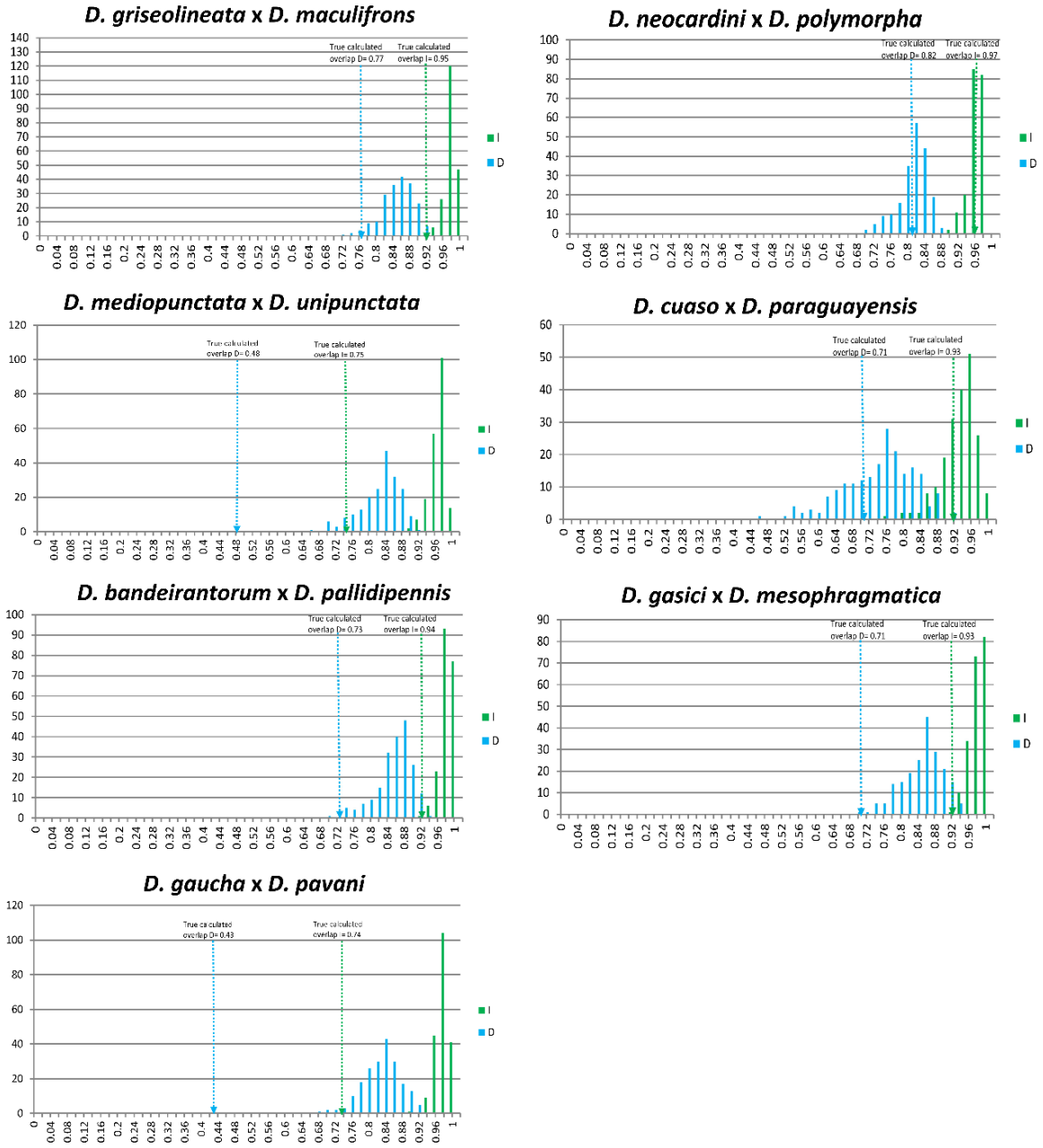

Figure S2: Niche equivalency hypothesis test for each pair of sister species. Vertical bars represent the null hypothesis distribution of Schoener's D (blue) and Hellinger's I (green) statistics, obtained after pooling together the registers of two species and then randomizing them 200 times in order to produce two new samples with the same number of observations as the empirical data which had their ENMs evaluated for D and I overlap patterns. Arrows show empirical D (blue) and I (green) value calculated for each sister species pair. Hypothesis of niche equivalency was rejected if empirical values were significantly lower than expected by chance. Significance values are detached.

3.2. Capítulo 2

Título:

Comparative ecological niche modeling and evolutionary ecology of Neotropical mycophagous Drosophilidae (Diptera) species

Artigo publicado no periódico:

Studies on Neotropical Fauna and Environment

ISSN:

0165-0521

DOI:

10.1080/01650521.2014.925370

Studies on Neotropical Fauna and Environment, 2014
http://dx.doi.org/10.1080/01650521.2014.925370

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Comparative ecological niche modeling and evolutionary ecology of Neotropical mycophagous Drosophilidae (Diptera) species

Lizandra Jaqueline Robe[a,b], Stela Machado[b], Andreza Ribeiro Bolzan[b], João Pedro Junges dos Santos[b], Felipe Berti Valer[c], Ashbel Perrut dos Santos[d], Monica Laner Blauth[c] & Marco Silva Gottschalk[c]*

[a]*Universidade Federal do Rio Grande - FURG, Instituto de Ciências Biológicas, Campus Carreiros, Rio Grande, RS, Brazil;* [b]*Universidade Federal de Santa Maria - UFSM, Programa de Pós Graduação em Biodiversidade Animal, Santa Maria, RS, Brazil;* [c]*Universidade Federal de Pelotas - UFPel, Instituto de Biologia, Departamento de Ecologia, Zoologia e Genética, Campus Capão do Leão, Capão do Leão, RS, Brazil;* [d]*Universidade do Estado de Mato Grosso, Campus Universitário de Tangará da Serra, Tangará da Serra, MT, Brazil*

(*Received 6 August 2013; accepted 14 May 2014*)

This study provides new records on the distribution of 22 species of mycophagous Drosophilidae and reports the southernmost and/or northernmost register points for most of them. Climatic modeling analysis revealed that the sampled species are heterogeneous as concerns the patterns of potential geographical ranges. A regular pattern emerging from the evaluated models is that the minimum temperature of the coldest month is one of the most determinant variables of habitat suitability for most species, although the particular response patterns may be further subdivided. In particular, some putative geographically restricted species appear to be highly susceptible to minor increases in the minimal temperature ranges, indicating that they are potential targets for the temperature rises expected to affect the Neotropics over the next decades and centuries. Moreover, as our analyses detected niche conservatism signals among the studied species, potential range shifts in face of future climatic oscillations may also be questioned.

Keywords: Hirtodrosophila; Mycodrosophila; Paraliodrosophila; Zygothrica; species distribution models (SDMs)

Introduction

Drosophilidae encompasses 76 described genera, with more than 4200 described species (Bächli 2013). Flies from this family are widely spread throughout the world, being found from sea level to high mountains, and from the tropics to tundras (Throckmorton 1975). Powell (1997) argued that originally drosophilid larvae probably fed on fermenting leaf litter and rotting vegetation, with a shift to dependence on microorganisms associated with fermentation and decay of different resources allowing further diversification. So there is a diverse array of Drosophilidae species reported in fruits, fungus, sap, flowers, pollen and rotten leaves (Carson 1971; Markow & O'Grady 2008), with these resources being used as feeding and/or as courtship/breeding sites. Despite this general association of drosophilids with yeast microorganisms, some species detach for their close affiliation to macroscopic fungi fructification bodies, and these will be referred as mycophagous species throughout this paper.

Mycophagous habits appear to have evolved independently several times within Drosophilidae (Throckmorton 1975; Powell 1997), which agrees with the notion that niche divergence may have dominated the evolution of Drosophilidae species (Robe et al. 2010). Nevertheless, mycophagy may represent a derived synapomorphy for the *Zygothrica* genus group (Grimaldi 1990), that encompasses *Hirtodrosophila* Duda, 1923, *Mycodrosophila* Oldenberg, 1914, *Paraliodrosophila* Duda, 1925 and *Zygothrica* Wiedemann, 1830, including 160, 127 and 124 species, respectively (Bächli 2013). However, even within these genera, mycophagous habits may be quite heterogeneous, with *Hirtodrosophila* and *Mycodrosophila* usually found in association with fungi in different stages, whereas *Zygothrica* mycophagy seems to be more labile, with fewer than 10% of species feeding fungi as larvae (Courtney et al. 1990). Mycophagy was also suggested for the species that encompass the poorly known genus *Paraliodrosophila*, which shares several characteristics with the *Mycodrosophila* and *Hirtodrosophila* genera (Vilela & Bächli 2007). Despite the high diversity presented by these four genera, few species have been registered for Brazil (Gottschalk et al. 2008), and these are generally represented by only a few records (Bächli 2013). This is probably an artifact of biased sampling effort, with

*Corresponding author. Email: gotts007@yahoo.com

collections mainly performed with banana-baited traps, which only rarely attract species of feed habits other than frugivory.

Concerning the particular preferences in host usages, there are claims that most mycophagous drosophilid species are polyphagous or generalist in virtue of the ephemeral and unpredictable nature of the colonized food resources and the chemical and nutritional homogeneity of the explored fungal fructification bodies (Jaenike et al. 1978; Courtney et al. 1990). Compatible with this idea, Lacy (1984) showed that host chemistry seems to have little effect on host preferences, and mycophagous drosophilids do not avoid fungi considered toxic or undigestible for other species. Nevertheless, according to Wertheim et al. (2000), despite the fact that co-existence between mycophagous species is mainly mediated by spatial aggregation and not by resource partitioning, several of the studied species have narrow diets. Regardless of the predominant niche breadth pattern, it is generally accepted that the distribution of mycophagous drosophilids may be severely limited by abiotic factors related to the physiological restrictions of the explored resources or of the flies themselves. In this regard, several authors stressed the great importance abiotic factors have for both the explored fungi (Shorrocks & Charlesworth 1980; Ohenoja & Metsanheimo 1982; Courtney et al. 1990) and the Drosophilidae species distributions (Spieth 1987; Kellermann, Loeschcke et al. 2012; Kellermann, Overgaard et al. 2012).

In recent years, the development of species distribution models (SDMs) or ecological niche models (ENMs) have allowed new insights into ecology, biogeography, evolution, conservation biology and climate change (Guisan & Thuiller 2005). SDMs are models that relate species distribution data (occurrence or abundance) with digital layers of environmental and/or spatial characteristics with the objective of improving knowledge and/or predicting the species' distribution across a landscape in a model-based interpolation to unsampled sites (Elith & Leathwick 2009). Despite the potential of this approach, the limited number of geographical records available for several species may represent a caveat to the applicability of this tool, due to its dependency on the degree to which the environmental dimensions examined actually define the species' distributional limits (Pearson et al. 2007). Even so, the capacity of SDMs to predict at least part of the possible distribution of a species could be a first step in planning sample strategies, providing support for studies that aim to understand patterns of distribution and species diversity. Moreover, for species for which more robust models are attained, SDMs can provide an important way of enhancing the knowledge involving the evolutionary and ecological patterns and processes subjacent to their biology.

This study adds to the current knowledge of the *Hirtodrosophila*, *Mycodrosophila*, *Paraliodrosophila* and *Zygothrica* Neotropical diversity, by providing new records on the distribution of 22 species. Additionally, we also applied SDMs to predict the potential geographic distribution of the studied species by combining the new occurrence records with those as yet available. The results so obtained were used to gain some ecological insights concerning, for example, the factors that contributed mostly to the modeled distributions and the response curves presented by each species for the most important variables. Finally, niche overlap measures were obtained for each pairwise SDM comparison, in order to evaluate the patterns of niche conservation presented by the studied species.

Materials and methods

Sample collections

Sample collections were carried out at 15 points distributed across southern and southeastern Brazil, within the Atlantic Forest and Pampa biomes; in addition, one collection was performed in the Brazilian Cerrado biome and four collections were performed in the Brazilian Amazonian biome (Table 1). Adult flies flying over or resting on fructification bodies of fungi were captured in the field with an entomological aspirator or an entomological net. Sometimes, fructification bodies were collected and stored at the laboratory, where emerged flies were swept daily. These field collections were performed along active search transects, with each point being inspected for approximately 2–4 h in each of the sampling days (Table 1).

The collected specimens were preserved in 70% or 100% ethanol and later identified based on general morphology and male genitalia slides, prepared according to Bächli et al. (2005). Identifications were performed through comparisons with the descriptions, keys and/or illustrations presented by Frota-Pessoa (1945), Burla (1956), Wheeler & Takada (1971), Grimaldi (1987, 1990) and Vilela & Bächli (2004, 2007).

Species distribution modeling

Geographic distributions for each of the sampled species were obtained from the Taxodros Database (Bächli 2013), which represents the sampling efforts of hundreds of drosophilid researchers across almost a century, and these were reviewed and added to the

Table 1. List of sampling sites with their respective geographical coordinates and collection dates.

No.	Sample sites: City	Sample sites: Location	Collection date	Geographical coordinates
1	Florianópolis (SC)	Morro da Lagoa da Conceição[a]	20 November 2009	27°35′27″ S, 48°28′33″ W
2	Ivoti (RS)	Cascata São Miguel[a]	20 February 2010	29°35′00″ S, 51°07′06″ W
3	Pejuçara (RS)	Balneário do Picolé[a]	27 February 2010, 4 April 2010	28°23′12″ S, 53°39′55″ W
4	Santa Maria (RS)	Bosque da Universidade Federal de Santa Maria (UFSM)[c]	23–24 March 2009	29°43′19″ S, 53°42′47″ W
5		Jardim Botânico (UFSM)[c]	23–24 March 2009	29°43′02″ S, 53°43′34″ W
6		Morro do Elefante[c]	27 April 2009	29°40′73″ S, 53°43′55″ W
7		Cascata do Mezzomo[c]	13 February 2009	29°38′35″ S, 53°36′01″ W
8	Santiago (RS)	Tênis Clube[b]	1 February 2010	29°11′09″ S, 54°53′50″ W
9		Parque Zamperete[b]	1 February 2010	29°12′51″ S, 54°51′27″ W
10	Viamão (RS)	Parque Saint'Hilaire[b]	19 February 2010	30°05′17″ S, 51°06′07″ W
11	Tuneiras do Oeste (PR)	Reserva Biológica das Perobas[a]	27 April 2011	23°53′10″ S, 52°49′01″ W
12	Diamante do Norte (PR)	Estação Ecológica do Caiuá[a]	30 April 2011	22°36′28″ S, 52°52′26″ W
13	Teodoro Sampaio (SP)	Parque Estadual do Morro do Diabo[a]	4 May 2011	22°37′29″ S, 52°10′31″ W
14	Pelotas (RS)	Horto Botânico Irmão Teodoro Luís[b]	22 February 2011, 25 March 2011, 29 April 2011, 27 May 2011	31°48′58″ S, 52°25′55″ W
15	Rio Grande (RS)	Estação Ecológica do Taim[b]	23 August 2011	32°32′25″ S, 52°32′34″ W
16	Melgaço (PA)	Floresta Nacional de Caxiuanã[d]	10 November 2011	01°47′32″ S, 51°26′02″ W
17	Colorado do Oeste (RO)	Sítio do Ildo[d]	12 January 2012	13°06′21″ S, 60°34′38″ W
18		Sítio do Maurício[d]	5 January 2012	13°00′38″ S; 60°35′25″ W
19		Sítio do João[d]	23 December 2011	13°06′40″ S, 60°35′45″ W
20	Tangará da Serra (MT)	Mata do Kart Cross[e]	25 January 2011	14°37′08″ S, 57°29′09″ W

Notes: [a]Sites in the Atlantic Forest; [b]Sites in the Pampa biome; [c]Sites at the interface between the Atlantic Forest and the Pampa biome; [d]Sites in the Amazonia biome; [e]Sites in the Cerrado biome.

occurrence records provided by Gottschalk et al. (2009) and by our collections (Table 2). The number of presence sites per species ranges from three (for *H. subgilva* and *Z. zygopoeyi*) to 61 (for *Z. prodispar*), with an average of 14 (Table 3).

These presence-only datasets were then used to construct the potential distribution maps for each of the encountered species, using the ecological niche model reconstruction algorithm DOMAIN (Carpenter et al. 1993) that generates maps of multivariate ecological similarity or distance to the sites at which the taxon is known to occur, as implemented in MODECO (http://gis.ucmerced.edu/ModEco), and the maximum entropy machine learning method performed by MAXENT (Phillips et al. 2006), that chooses the most unconstrained model among those presenting probability distributions that satisfy the constraint that the average estimated value of each environmental variable corresponds to its empirical average (Phillips et al. 2004, 2006). In both cases, model projections were established based on nine of the 19 continuous climatic variables plus altitude as available at WorldClim 1.3 database (Hijmans et al. 2005) at a 2.5-min resolution (5 km^2). Variable selection was performed through evaluation of the pairwise Pearson correlation coefficients among layers, as implemented in ENMTools (Warren et al. 2008, 2010), with the retained variables always showing small correlations among each other ($r < |0.75|$). The nine selected variables were: altitude (Alt), mean diurnal range (Bio-2), maximum temperature of the warmest month (Bio-5), minimum temperature of the coldest month (Bio-6), precipitation seasonality (Bio-15), precipitation of the wettest quarter (Bio-16), precipitation of the driest quarter (Bio-17), precipitation of the warmest quarter (Bio-18) and precipitation of the coldest quarter (Bio-19).

DOMAIN and MAXENT models were employed here because both algorithms use presence-only datasets. Absence inferences can be especially misleading with Drosophilidae species since they can introduce bias not only when the habitat is suitable but unoccupied or when the species is not at equilibrium, but also when the species is in fact present but remained undetected or was not correctly identified. In fact, Anderson et al. (2003) recommended the use of observed absences only when a site has been surveyed extensively, or when a species can be confidently detected by a single survey, which is not the case of our species. Moreover, considering the small number of registers presented for some species, DOMAIN and MAXENT were chosen because both are relatively independent of sample sizes, being able to build reliable models even with small species occurrence

Table 2. List of collected *Hirtodrosophila*, *Mycodrosophila*, *Paraliodrosophila* and *Zygothrica* species, with their respective taxonomic position (according to Bächli 2013) and sampling site(s) as registered by this study (referred according to the numbers given in Table 1).

Genera	Species group	Species	Sampling site(s)
Hirtodrosophila	*glabrifrons* group	*H. levigata* (Burla 1956)	1, 2, 4, 6, 7, 10, 11, 14
	hirticornis group	*H. gilva* (Burla 1956)	12
		H. mendeli (Mourão, Gallo & Bicudo 1965)	7, 8, 11, 12, 14
		H. morgani (Mourão, Gallo & Bicudo 1965)	12, 14, 19, 20
		H. subgilva (Burla 1956)	1, 7
	Not assigned to group	*H. subflavohalterata* (Burla 1956)	16, 17
Mycodrosophila		*M. elegans* (Wheeler & Takada 1963)	12, 13, 16, 19
		M. projectans (Sturtevant 1916)	3, 4, 6, 10, 12, 13, 14, 15, 16
Paraliodrosophila		*P. antennata* (Wheeler 1957)	3, 5, 8, 12, 13, 16, 17, 18, 19
Zygothrica	*atriangula* group	*Z. atriangula* Duda 1927	16
		Z. parapoeyi (Burla 1956)	2
		Z. poeyi (Sturtevant 1921)	1, 2, 6, 10, 11, 14, 17
		Z. virgatinigra (Burla 1956)	16
	bilineata group	*Z. bilineata* (Williston 1896)	14
	dispar group	*Z. dispar* (Wiedmann 1830)	2, 10, 11, 12, 14
		Z. prodispar Duda 1925	2, 11, 12, 14, 20
	hypandriata group	*Z. hypandriata* (Burla 1956)	1, 6, 7, 10
	orbitalis group	*Z. orbitalis* (Sturtevant 1916)	1, 2, 7, 12
	vittimaculosa group	*Z. vittinubila* (Burla 1956)	11
		Z. zygopoeyi (Burla 1956)	11
	Not assigned to group	*Z. parvipoeyi* (Burla 1956)	7
		Z. ptilialis (Burla 1956)	6, 7, 10, 11, 14

datasets (Hernandez et al. 2006; Wisz et al. 2008). In this sense, Pearson et al. (2007) stated that sample sizes may be successfully reduced below 10 points if MAXENT is used. So, in general, both these two algorithms appear to fulfill our different requirements.

For DOMAIN, a sensitivity measure was used as validation strategy based on a Jackknife k-fold cross-validation procedure, with k equal to the number of occurrence localities and one sampling point withheld from the calibration set and encompassing the test set at each partition, as recommended by Pearson et al. (2007) for low sample sizes. In this case, reclassification of the SDMs to binary maps of predicted presence and absence was accomplished using a similarity threshold as stringent as 0.95. The models generated by MAXENT, on the other hand, were tested through the reconstruction of 25 replicates of bootstrap, in each of which 25% of the randomly chosen registers were withheld as tests. Area under the receiving operating curve (AUC) [which evaluates both omission (false negatives) and commission errors (false positives) by plotting the sensitivity values (true positive fractions) against 1-specificity (false positive fraction) for all available thresholds (Fielding & Bell 1997)], was used as a diagnostic test performed in face of the testing datasets. An AUC of 0.5 indicates that the model performance is no better than a random prediction and an AUC of 1.00 yields theoretically perfect discrimination ability to the model.

Each of the nine WordClim climatic variables used on modeling was also evaluated for its contribution to the derived DOMAIN and MAXENT SDMs. For DOMAIN, measures of the most limiting factors (MLF) were obtained through mapping the variable for which the percentile score of each cell was lowest or highest. For MAXENT, the relative contribution was measured through the increase in the SDM regularized gain added by the corresponding variable, or through the decrease subtracted from it if the change was negative. Response curves showing how each environmental variable affects the MAXENT predictions were also evaluated.

Finally, pairwise measures of predicted niche overlap were obtained for the MAXENT models through the use of two identity tests implemented in ENMTools: Schoener's D and Hellinger's I statistics. Both these indices range from 0 (niche models have no overlap) to 1 (niche models identical), and do not require thresholds for distinguishing presence and absence (Warren et al. 2008). Statistical significance of the niche overlap measures was evaluated in ENMTools using the identity test (Warren et al. 2008, 2010), which tests if SDMs generated for two species are more different than expected if they were drawn from the same underlying distribution.

Table 3. Prediction success of SDMs related to sensitivity for DOMAIN models and to AUC value for MAXENT models.

	Number of points*	DOMAIN model Sensitivity	MAXENT model AUC value
H. gilva	4	0.50	0.95
H. levigata	11	0.82	0.95
H. mendeli	8	1.00	0.91
H. morgani	8	0.75	0.78
H. subflavohalterata	4	0.50	0.87
H. subgilva	3	0.00	–
M. elegans	10	0.78	0.78
M. projectans	30	0.83	0.92
*P. antennata***	14	0.93	0.89
Z. atriangula	7	0.57	0.83
Z. bilineata	28	0.90	0.93
Z. dispar	37	0.92	0.93
Z. hypandriata	11	1.00	0.94
Z. orbitalis	19	0.89	0.99
Z. parapoeyi	4	0.75	0.96
Z. parvipoeyi	4	0.75	0.96
Z. poeyi	22	0.91	0.95
Z. prodispar	61	0.84	0.87
Z. ptilialis	9	1.00	0.93
Z. virgatinigra	4	0.50	0.86
Z. vittinubila	6	0.67	0.93
Z. zygopoeyi	3	0.00	–

Notes: *As obtained after adding our registers and those presented by Gottschalk et al. (2009) to Bächli (2013) dataset. ** Cited as *Hirtodrosophila thoracis* (Williston 1896) by Gottschalk et al. (2009). The species shaded in gray had their SDMs withdrawn due to their poor prediction performance.

Pairwise range overlap measures (R) were also obtained in ENMTools based on binary MAXENT models, with presence/absence patterns contrasted through the use of the mean minimum training presence threshold. Correlations between D, I and R were estimated through a Mantel test, with 5000 permutations, as performed in Past 2.16 (Hammer et al. 2001).

Results

A total of 22 species belonging to the genera *Hirtodrosophila*, *Mycodrosophila*, *Paraliodrosophila* and *Zygothrica* were identified among the collected individuals (Table 2). The currently known distribution of each of these species was expanded by this study, and our collection points encompass the southernmost and northernmost records available for one species (*H. morgani*), the southernmost record for other 16 species (*H. levigata*, *H. mendeli*, *H. subgilva*, *M. elegans*, *M. projectans*, *P. antennata*, *Z. bilineata*, *Z. dispar*, *Z. hypandriata*, *Z. parapoeyi*, *Z. parvipoeyi*, *Z. poeyi*, *Z. prodispar*, *Z. ptilialis*, *Z. vittinubila* and *Z. zygopoeyi*), and the northernmost record for other two species (*H. subflavohalterata* and *Z. atriangula*).

Concerning the SDMs, only 13 of the 22 species models passed through our validation criteria of presenting more than six register points associated with a DOMAIN and a MAXENT SDM prediction performance greater than 0.75 (as measured through sensitivity and AUC value, respectively) (Table 3). Sensitivity values greater than 0.75 reflect omission rates (false negatives) lower than 0.25, whereas AUC values greater than 0.75 should be interpreted as at least a 75% probability that the model can correctly distinguish between a presence and a background record if the point is selected randomly from the set of presences and absences (Phillips et al. 2006; Pearson 2007), so that both criteria assigned a good performance to the sustained models. The species that presented such a moderate/high predictive performance were *H. levigata*, *H. mendeli*, *H. morgani*, *M. elegans*, *M. projectans*, *P. antennata*, *Z. bilineata*, *Z. dispar*, *Z. hypandriata*, *Z. orbitalis*, *Z. poeyi*, *Z. prodispar* and *Z. ptilialis*, and their distribution models are evaluated in Figure 1. Interestingly, most species which presented poorer performances (*H. gilva*, *H. subflavohalterata*, *H. subgilva*, *Z. atriangula*, *Z. virgatinigra*, *Z. vittinubila* and *Z. zygopoeyi*) also had no more than six records (Table 3). It is also important to emphasize that for the 13 maintained average models, all the observed records fell within areas of high habitat suitability, predicted as presence for at least one model (similarity threshold of 0.95 for DOMAIN and minimum training presence for MAXENT), although MAXENT generally presented higher omission rates as compared to DOMAIN. In fact, most of MAXENT's presence prediction was generally contained within DOMAIN's suitability areas, although the former extended further southwards and westwards in several cases (exceptions provided by *M. elegans* and *Z. orbitalis*). Nevertheless, we need to be aware that this difference could be an artifact of unscaled threshold choice.

In spite of this, the SDMs recovered by both algorithms for the three species with larger sample sizes (*M. projectans*, *Z. dispar* and *Z. prodispar*) were in general agreement with each other, suggesting that DOMAIN and MAXENT may be capturing different niche properties in the case of species with lower sample sizes. Using a conservative approach, it is possible to trace the potential distribution of each species in face of its actual occurrences by intersecting the habitat suitability models generated by both algorithms. This strategy shows that the studied species are widely heterogeneous regarding the patterns of geographical ranges (Figure 1), with *Z. prodispar*, *Z.*

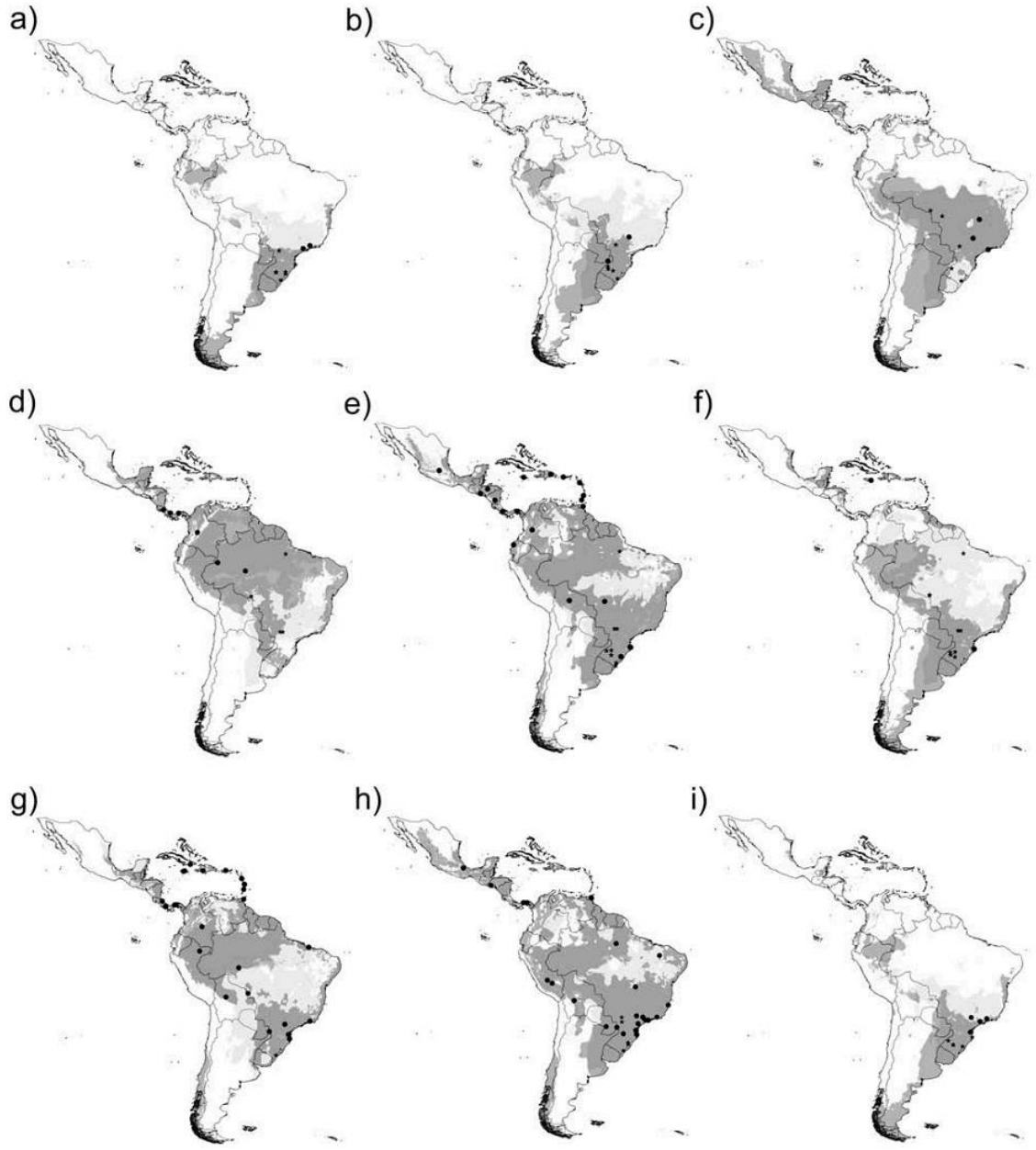

Figure 1. Environmental suitability distribution models for the species (a) *Hirtodrosophila levigata*; (b) *H. mendeli*; (c) *H. morgani*; (d) *Mycodrosophila elegans*; (e) *M. projectans*; (f) *Paraliodrosophila antennata*; (g) *Zygothrica bilineata*; (h) *Z. dispar*; (i) *Z. hypandriata*; (j) *Z. orbitalis*; (k) *Z. poeyi*; (l) *Z. prodispar*; and (m) *Z. ptilialis* as projected only by the point-to-point similarity based method derived by DOMAIN (light gray), only by the machine-learning maximum entropy method performed in MAXENT (intermediate gray), or by both algorithms (dark gray). Occurrence points obtained by this study are shown as black stars, whereas those obtained from the searchable web portal of Bächli (2013) are shown by black circles. Presence (scales of gray) and absence (white) patterns were extracted from the continuous probability distributions (%), using the MTP (minimum training presence) and a similarity threshold of 0.95 for MAXENT and DOMAIN, respectively.

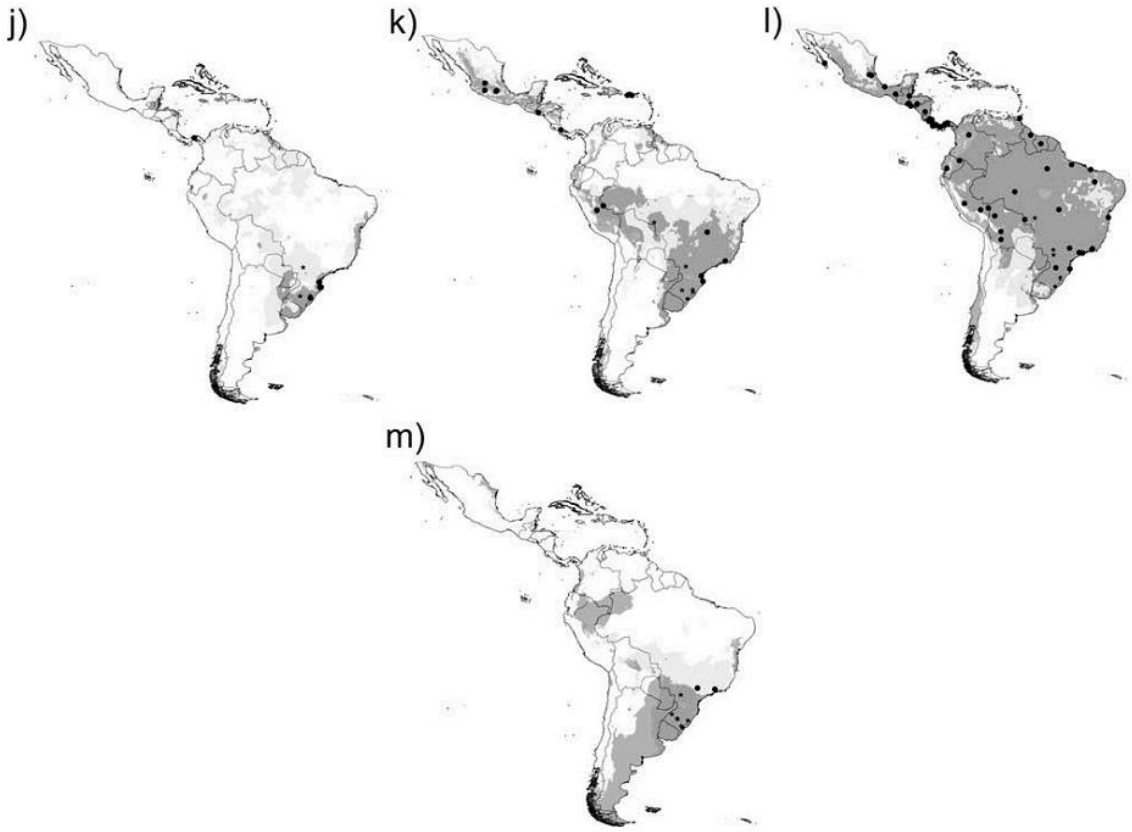

Figure 1. (Continued).

dispar and *M. projectans* outstanding as widely distributed, presenting high environmental suitability values across almost the entire Neotropical region. Conversely, *H. levigata*, *H. mendeli*, *Z. hypandriata*, *Z. orbitalis* and *Z. ptilialis* appear to be endemic to smaller and frequently disjunct regions (especially for MAXENT). Nevertheless, several of these disjunct patches appear to be unoccupied by the modeled species.

According to MAXENT analyses, the environmental variables that contribute mostly to the models of the mycophagous species (Figure 2) are precipitation seasonality (Bio-15, mean = 18%), precipitation of the warmest quarter (Bio-18, mean = 16%) and minimum temperature of the coldest month (Bio-6, mean = 12%). Interestingly, DOMAIN's analysis of the most limiting factors for each species positioned the minimal temperature of the coldest month as the most important variable (mean = 30%), followed by the maximum temperature of the warmest month (Bio-5, mean = 16%). In fact, Bio-6 is DOMAIN's most limiting factor for six out of the 13 species (*H. levigata*, *H. mendeli*, *H. morgani*, *M. elegans*, *Z. hypandriata* and *Z. ptilialis*) whereas three of the remaining species present Bio-5 as the most limiting factor (*M. projectans*, *P. antennata* and *Z. prodispar*) (Figure 2), showing that the distribution of these mycophagous species may be widely affected by temperature extremes.

Concerning the response curves generated by MAXENT based on the modeled patterns of environmental suitability related to each climatic variable, some general trends can be established. For Bio-15, for example, most modeled species appear to present a negative response to an increase in precipitation seasonality values (Figure 3), although the opposite pattern can be found for *Z. prodispar*. Complementary to this, the response curves for Bio-18 reveals that patterns of environmental suitability are positively correlated with the precipitation of the warmest quarter for most species, with the exception of *Z. prodispar* and

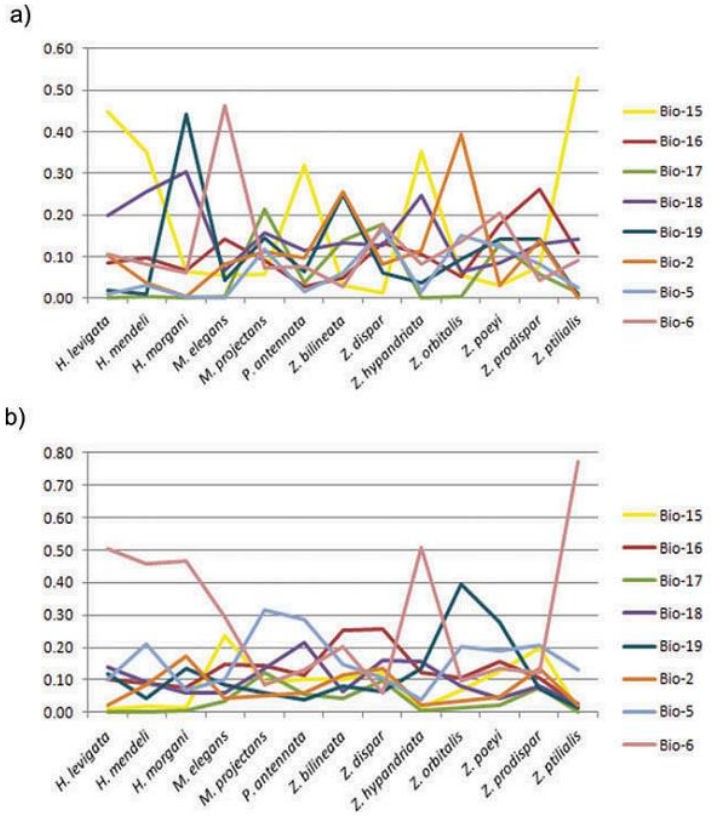

Figure 2. (Color online) Contribution of each of the nine studied environmental variables to (a) MAXENT SDMs and (b) DOMAIN most limiting factor for each mycophagous species.

maybe *P. antennata*, *Z. dispar* and *Z. poeyi*, for which the standard deviations prevent visualization of more consistent patterns (Figure 4). A very interesting scenario emerges from the analysis of response curves related to Bio-6 (Figure 5): (1) a group of species composed by *H. levigata*, *H. mendeli*, *H. morgani*, *Z. hypandriata* and *Z. ptilialis* present higher predicted environmental suitability values in areas with lower minimum temperatures (of the order of –20°C to 5°C); (2) a unique species, *M. elegans*, shows higher predicted environmental suitability values in areas with higher minimum temperatures (of the order of 20–25°C); (3) a set of species encompassing *M. projectans*, *P. antennata*, *Z. bilineata*, *Z. dispar*, *Z. orbitalis*, *Z. poeyi* and *Z. prodispar* show higher predicted environmental suitability values in areas with intermediate minimum temperatures (of the order of 5–15°C).

Patterns of niche and range overlap between species as inferred from comparisons between SDMs confirmed some of the above depicted environmental suitability similarities among species. All species presented some level of predicted sympatry or range overlap (Table 4), and niche overlap values also varied (Table 5): higher similarities were found in the comparisons involving *H. levigata*, *H. mendeli*, *Z. hypandriata* and *Z. ptilialis* (D = 0.85 to 0.96 and I = 0.98 to 1.00); lower similarities were found in the comparisons between *Z. orbitalis* and any of *H. morganil M. elegans*/*Z. prodispar* (D = 0.34–0.36 and 0.64 and I = 0.63–0.65). However, the identity tests performed to determine whether SDMs generated from two species are more different than expected if they were drawn from the same underlying distribution revealed that only 9–10% of the pairwise comparisons rejected the null hypothesis of niche identity after a

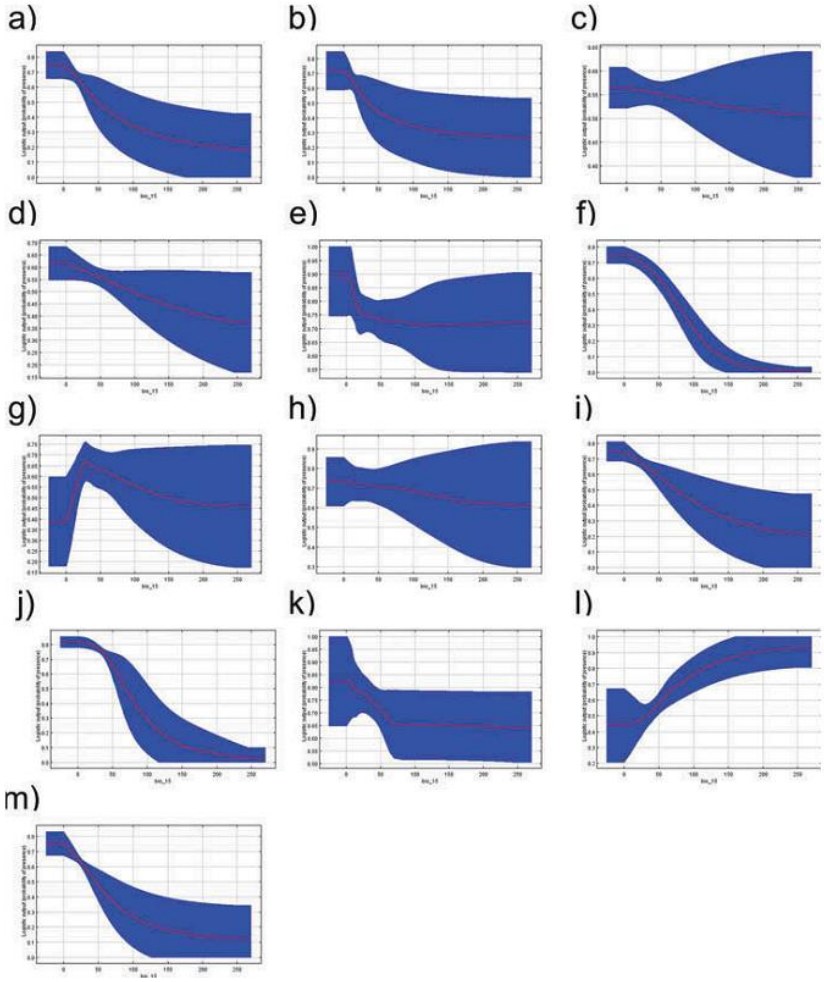

Figure 3. (Color online) Response curves of (a) *Hirtodrosophila levigata*; (b) *H. mendeli*; (c) *H. morgani*; (d) *Mycodrosophila elegans*; (e) *M. projectans*; (f) *Paraliodrosophila antennata*; (g) *Zygothrica bilineata*; (h) *Z. dispar*; (i) *Z. hypandriata*; (j) *Z. orbitalis*; (k) *Z. poeyi*; (l) *Z. prodispar*; and (m) *Z. ptilialis*, showing how precipitation seasonality (Bio-15) affects MAXENT prediction in each case. The curves show how the logistic prediction changes as Bio-15 is varied, keeping all other environmental variables at their average sample value. The curves show the mean response of the 25 replicate Maxent runs (red line) and the mean ± one standard deviation (blue shading).

Bonferroni correction. Most of these significantly different comparisons referred to both *Z. prodispar* and *Z. orbitalis* in relation to the other mycophagous species (Table 5). Mantel tests revealed that D and I niche overlap indices are significantly correlated ($R_{IxD} = 0.97$, $p = 2 \times 10^{-4}$), although the evaluations performed in order to measure the relationship between niche overlap patterns and space revealed that both indices are not spatially autocorrelated ($R_{R\ x\ D} = 0.2$, $p > 0.1$; $R_{R\ x\ I} = 0.16$, $p > 0.1$).

Discussion

Ecological properties of the mycophagous species

The correlative approach of niche modeling here implemented for 13 of the sampled mycophagous species revealed a high heterogeneity concerning the modeled potential distribution ranges, so biome putative diversity seems to vary across the Neotropical region. At one extreme, *M. projectans*, *Z. dispar* and *Z. prodispar* appear to be geographically dispersed across most of the Neotropics, despite the fact that

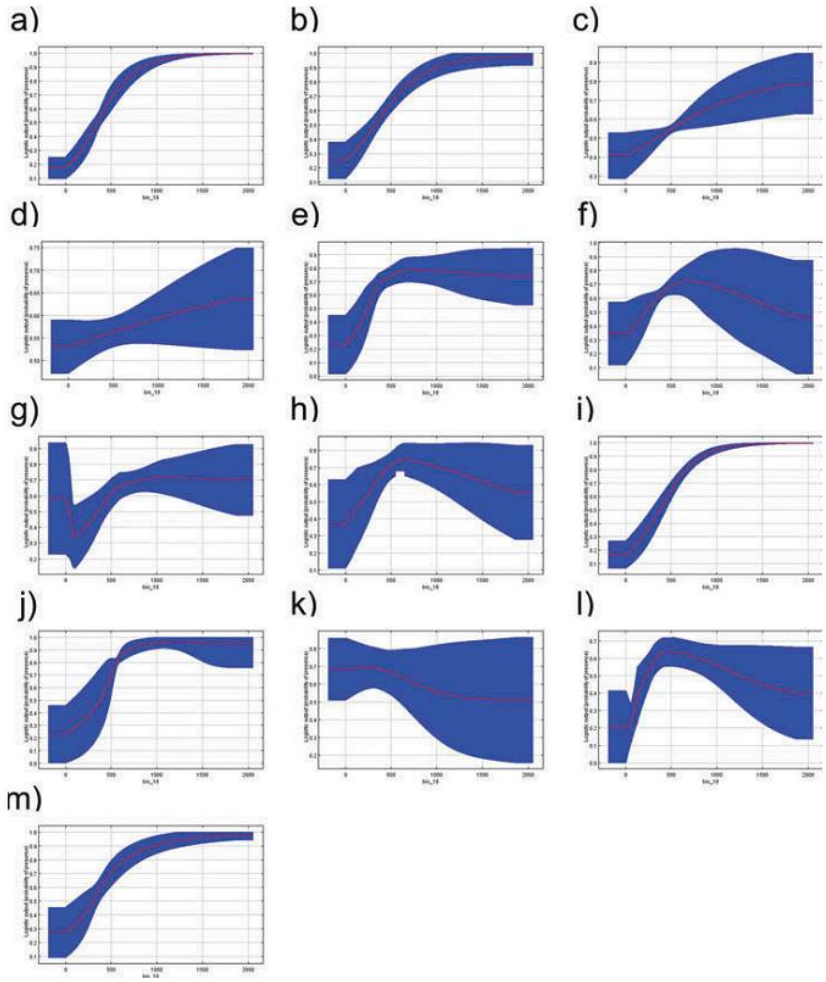

Figure 4. (Color online) Response curves of (a) *Hirtodrosophila levigata*; (b) *H. mendeli*; (c) *H. morgani*; (d) *Mycodrosophila elegans*; (e) *M. projectans*; (f) *Paraliodrosophila antennata*; (g) *Zygothrica bilineata*; (h) *Z. dispar*; (i) *Z. hypandriata*; (j) *Z. orbitalis*; (k) *Z. poeyi*; (l) *Z. prodispar*; and (m) *Z. ptilialis*, showing how precipitation of the warmest quarter (Bio-18) affects MAXENT prediction in each case. The curves show how the logistic prediction changes as Bio-18 is varied, keeping all other environmental variables at their average sample value. The curves show the mean response of the 25 replicate Maxent runs (red line) and the mean ± one standard deviation (blue shading).

M. projectans had been as yet only sporadically registered for Brazil (Gottschalk et al. 2008; Bächli 2013). At the other extreme, *H. levigata*, *H. mendeli*, *Z. hypandriata*, *Z. orbitalis* and *Z. ptilialis* are deemed as geographically more endemic species, being predicted to occur only in southern and southeastern Brazil, although MAXENT predicts a significant probability of distribution also for northwestern region of the Amazonian Forest. These disjunct areas seemingly unoccupied by the modeled species may be resultant of subsampling, biological interactions (as competition), dispersal limitations, or may even represent the distribution of a potentially unknown species (Raxworthy et al. 2003, 2007), putatively closely related to the modeled one, but yet undifferentiated at the ecological level.

In addition to the importance of distribution projection *per se*, modeling results can also be used to find the relative importance of each variable to the predicted distribution of the species, providing insights into the environmental critical conditions affecting each taxon. Furthermore, an analysis of predicted distribution in face of variation in these variables can lead to suitability ranges which

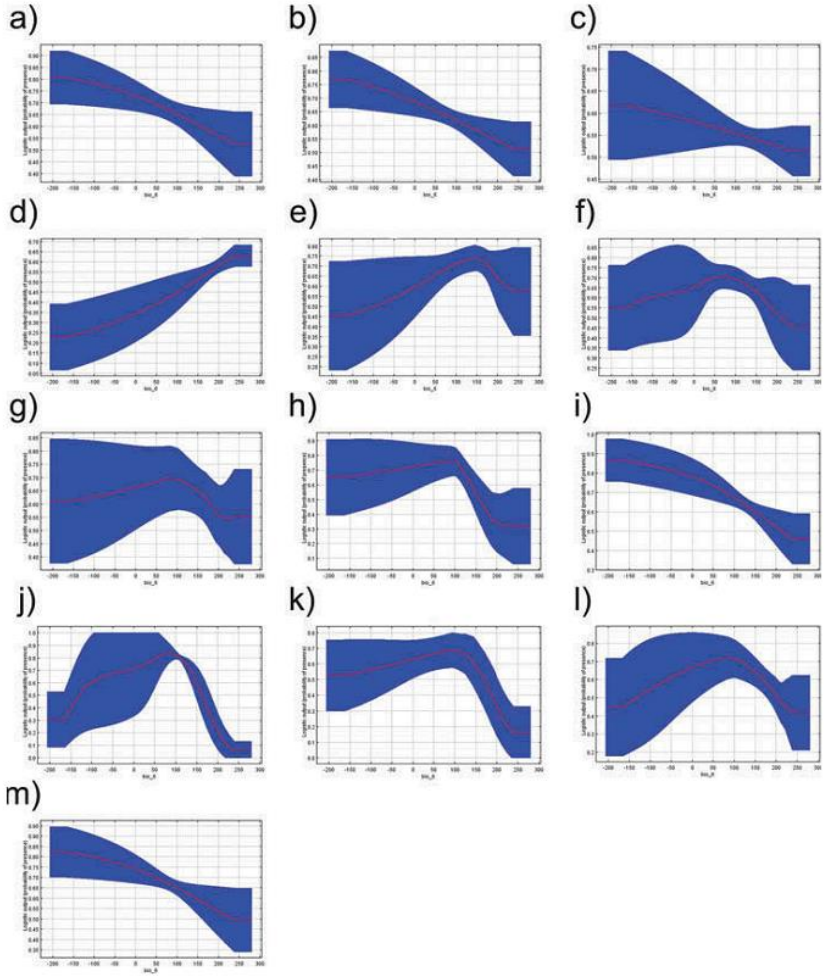

Figure 5. (Color online) Response curves of (a) *Hirtodrosophila levigata*; (b) *H. mendeli*; (c) *H. morgani*; (d) *Mycodrosophila elegans*; (e) *M. projectans*; (f) *Paraliodrosophila antennata*; (g) *Zygothrica bilineata*; (h) *Z. dispar*; (i) *Z. hypandriata*; (j) *Z. orbitalis*; (k) *Z. poeyi*; (l) *Z. prodispar*; and (m) *Z. ptilialis*, showing how the minimal temperature of the coldest month (Bio-6) affects MAXENT prediction in each case. The curves show how the logistic prediction changes as Bio-6 is varied, keeping all other environmental variables at their average sample value. The curves show the mean response of the 25 replicate Maxent runs (light gray line) and the mean ± one standard deviation.

putatively define habitat preferences along climatic gradients. In this respect, although for MAXENT variables related to precipitation (Bio-15 and Bio-18) are the most important factors in defining the SDMs, the minimal temperature of the coldest month (Bio-6) appeared as the most limiting factor affecting DOMAIN's modeled distribution of the mycophagous Drosophilidae species. These results are in agreement with the previously reported association between desiccation and cold resistance with climate range characteristics, as annual precipitation and minimum temperature of the coldest month, respectively (Kellerman, Loeschcke et al. 2012), so that these climatic niche traits appear to be closely linked to the geographic distribution of different Drosophilidae species. Moreover, Kellerman, Overgaard et al. (2012) showed that although it is generally assumed that heat resistance is related to the maximum temperature of the warmest months, the values of that trait are strongly negatively associated with precipitation, suggesting that factors related to water imbalance are more important in driving not only desiccation patterns but also the upper thermal limits. It was also previously suggested

Table 4. Pairwise range overlap R measures.

	H. levigata	*H. mendeli*	*H. morgani*	*M. elegans*	*M. projectans*	*P. antennata*	*Z. bilineata*	*Z. dispar*	*Z. hypandriata*	*Z. orbitalis*	*Z. poeyi*	*Z. prodispar*	*Z. ptilialis*
H. levigata		0.926	1.000	0.834	0.808	0.825	0.691	0.766	0.977	1.000	0.614	0.391	0.918
H. mendeli			1.000	0.795	0.823	0.903	0.694	0.744	0.936	0.991	0.649	0.425	0.991
H. morgani				0.933	0.999	1.000	0.976	1.000	1.000	1.000	1.000	0.937	1.000
M. elegans					0.998	0.939	0.988	0.913	0.833	1.000	0.870	0.938	0.785
M. projectans						0.844	0.734	0.754	0.819	0.948	0.604	0.685	0.624
P. antennata							0.803	0.681	0.842	0.999	0.522	0.452	0.786
Z. bilineata								0.474	0.701	0.710	0.485	0.904	0.539
Z. dispar									0.754	0.901	0.760	0.710	0.578
Z. hypandriata										1.000	0.601	0.382	0.945
Z. orbitalis											0.864	0.757	0.976
Z. poeyi												0.724	0.459
Z. prodispar													0.295
Z. ptilialis													

Table 5. Pairwise niche overlap D (above the diagonal) and I (below the diagonal) measures.

	H. levigata	*H. mendeli*	*H. morgani*	*M. elegans*	*M. projectans*	*P. antennata*	*Z. bilineata*	*Z. dispar*	*Z. hypandriata*	*Z. orbitalis*	*Z. poeyi*	*Z. prodispar*	*Z. ptilialis*
H. levigata		0.860	0.730	0.661	0.657	0.804	0.584	0.636	0.956	0.496	0.584	0.571*[2]	0.893
H. mendeli	0.982		0.794	0.711*[1]	0.637	0.811	0.564	0.618	0.855	0.431	0.572	0.602*[2]	0.884
H. morgani	0.945	0.968		0.783*[1]	0.615	0.696	0.554	0.629	0.702	0.347*[1]	0.607	0.716	0.698
M. elegans	0.910	0.936	0.963		0.646	0.708	0.662	0.585	0.645	0.356*[2]	0.533*[1]	0.769	0.640
M. projectans	0.885	0.875	0.871	0.900		0.712	0.765	0.807	0.653	0.527	0.687	0.690*[1]	0.611
P. antennata	0.945	0.956	0.904	0.927	0.910		0.655	0.649	0.815	0.519	0.568	0.604*[1]	0.801
Z. bilineata	0.842	0.830	0.827*[1]	0.899	0.950	0.885		0.666	0.584	0.463*[1]	0.570	0.708	0.545
Z. dispar	0.870	0.859	0.877	0.868	0.968	0.861	0.906		0.624	0.489*[1]	0.774	0.684*[2]	0.584
Z. hypandriata	0.998	0.981	0.933	0.902	0.884	0.953	0.841	0.862		0.512	0.571	0.552*[2]	0.894
Z. orbitalis	0.777	0.716	0.636*[1]	0.653*[2]	0.810	0.806	0.762	0.770*[1]	0.792		0.429*[2]	0.337*[2]	0.474
Z. poeyi	0.843	0.833	0.862	0.796*[1]	0.894	0.793	0.806*[1]	0.942	0.831	0.708*[2]		0.635*[1]	0.535
Z. prodispar	0.844*[1]	0.861*[2]	0.921	0.950	0.924	0.848*[1]	0.922	0.922*[2]	0.830*[1]	0.634*[2]	0.859*[2]		0.522*[2]
Z. ptilialis	0.990	0.990	0.934	0.903	0.857	0.955	0.815	0.835	0.991	0.754	0.807	0.814*[2]	

Note: *These measures denote comparisons in which overlap between species distribution models generated from the actual data seem to be more different than those obtained through a null distribution constructed through 100 randomizations of occurrence identities to produce new samples with the same number of observations as the empirical data [*[1]$0.01 \le p < 0.05$; *[2]$p < 0.01$].

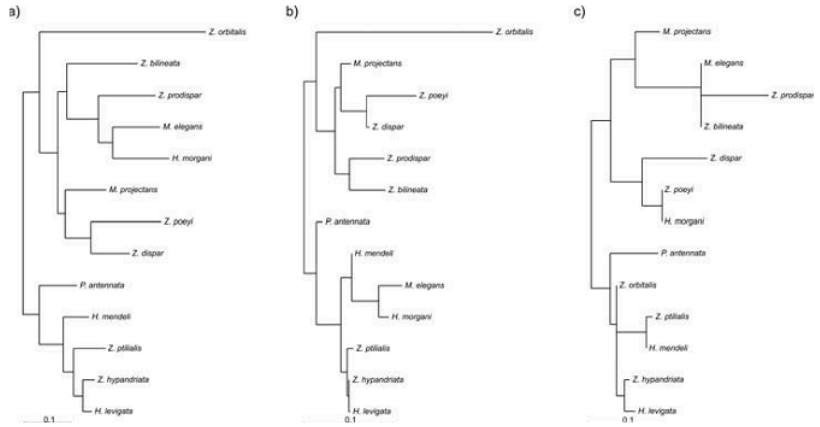

Figure 6. Neighbor-joining phylograms depicting similarity patterns between the 13 mycophagous modeled species as concerns the niche overlap indices D (a) and I (b), and the range overlap statistic R (c).

that mycophagous species may be less desiccation-resistant than the other *Drosophila* species (Spieth 1987), and that their resource abundance appears to be extremely sensitive to climatic conditions, especially rainfall (Shorrocks & Charlesworth 1980).

In accordance with this pattern, the tolerance ranges related to the precipitation variables appear to be strikingly conserved among the studied species. With the exception of *Z. prodispar*, mycophagous species tend to present higher environmental suitability values in areas with lower precipitation seasonality and higher precipitation of the warmest quarter. In contrast, the pattern emerging from the Bio-6 response curves is by far more variable, and three sets of response curves can be envisioned: *H. levigata*, *H. mendeli*, *H. morgani*, *Z. hypandriata* and *Z. ptilialis* appear to be better adapted to lower minimum temperatures; *M. elegans* presents higher suitability values in areas with higher minimum temperatures; *M. projectans*, *P. antennata*, *Z. bilineata*, *Z. dispar*, *Z. orbitalis*, *Z. poeyi* and *Z. prodispar* appear to be better adapted to intermediate minimum temperatures. Consistent with this, Kellermann, Loeschcke et al. (2012) showed that phylogenetic associations related to the evolution of desiccation resistance (intimately related to precipitation variables) in *Drosophila* acted at higher taxonomic levels than those related to the evolution of cold resistance (strongly correlated to the average minimum temperature). Interestingly, *H. levigata*, *H. mendeli*, *Z. hypandriata* and *Z. ptilialis* not only present coincident response curves regarding the minimum temperature resistance, but are also among the potentially more restricted species, presenting the most similar environmental niche properties in general, as measured by D and I indices. Although this result could be easily attributed to spatial autocorrelation, niche overlap and range overlap measures were not significantly correlated, at least in regard to the 13 modeled species considered as a whole.

Concerning the niche overlap patterns, although no taxonomic structure could be recovered [once the four studied genera (*Hirtodrosophila*, *Mycodrosophila*, *Paraliodrosophila* and *Zygothrica*) and the two species groups with at least two species sampled (the *hirticornis* group of *Hirtodrosophila*, and the *dispar* group of *Zygothrica*) are highly intermingled in the reconstructed ecological phylograms (Figure 6)], some levels of niche conservatism are suggested by our data. Besides the similarity presented by some species regarding response curves and pairwise niche overlaps, most comparisons between the SDMs did not reject the null hypothesis of niche identity, suggesting niche conservatism among most of the evaluated mycophagous species. In fact, only *Z. orbitalis* and *Z. prodispar* presented a reiterated signal of niche divergence in relation to other mycophagous species, probably reflecting the fact that these species present the narrowest and widest potential distribution ranges, respectively. The envisioned niche identity patterns in spite of other potential distribution range differences is probably related to the fact that most species present higher and/or lower environmental suitability values coincident to the same regions, so that the ultimate signal is provided by shared and/or unsuited geographical areas, once the similarity D and I measures are obtained by comparing the estimates of habitat suitability for each grid cell after normalization so that they sum to 1 within the studied area (Warren et al.

2010). Even so, taken as a whole, these results suggest that despite previous evidence of niche divergence (Throckmorton 1975; Robe et al. 2010) niche conservatism may also have played an important role in the evolutionary history of some Drosophilidae species. Nevertheless, it remains to be seen if this pattern is sustained after unbiased continued sampling, and after accomplishment of additional niche conservatism tests, especially phylogenetic signal tests in the face of a required phylogenetic hypothesis for this group of species. Even the monophyly of each genus and/or group is still largely questionable, as suggested by the *Hirtodrosophila* polyphyly presented by DaLage et al. (2007) and van der Linde et al. (2010).

There are some caveats regarding the potential distribution projections that need to be further considered: if the modeled distribution is to be used to target field surveys, it needs to be made clear that some regions outlined in the maps as sites of potential distribution may not be inhabited at all by the referred species, due to biotic interactions, dispersal constraints or historical contingencies (Pearson 2007). On the other hand, if the observed occurrence records do not reflect accurately the environmental space occupied by the species, so that some relevant environmental gradients are not adequately sampled, a niche model does not necessarily predict the full extent of the actual/potential distribution of a species (Pearson 2007). Nevertheless, the fact that this study attained the southernmost and/or northernmost register for each of the sampled species, with the exceptions of *H. gilva*, *Z. orbitalis* and *Z. virgatinigra*, supports the idea that it helped to approach the most limiting factors associated with their climatic niche patterns.

SDMs and related issues

As mentioned above, the modeling results should be used critically, especially in face of the putative errors and biases reflected by the occurrence records. Elith et al. (2006) showed that predictive performance does not vary with the number of occurrence points available for modeling. Several posterior studies have also shown that different methods are able to successfully predict species distribution despite low sample sizes (Hernandez et al. 2006; Pearson et al. 2007; Wisz et al. 2008). Even so, we found that a threshold of at least eight records is necessary to obtain consistent results, so that nine species had their SDMs suppressed due to their poor predictive performance probably related to small sample sizes. Although the poor performance of these nine SDMs was mainly restricted to DOMAIN predictive success, MAXENT results for these species were also omitted because AUC values can be overestimated if the evaluation data contains absence points from a very large area (Wisz et al. 2008). Thus, in general, our results suggest that naturally data-depauperate drosophilid species (those with less than eight occurrence records) provide error-prone, variable and inconsistent distribution models that are widely affected by sampling design, especially when it is biased at its source and is not a random representation of the actual species distribution. Nevertheless, there is also the possibility that the threshold artificial selection decreased DOMAIN's prediction success for the smaller sample size categories, because as this methodology measures the distance between each point and its most similar occurrences, as the sample size decreases, each distance tends to increase, decreasing the probability of suitability (Hernandez et al. 2006).

Among the remaining 13 SDMs, we preferred to adopt a MAXENT versus DOMAIN comparative strategy, although MAXENT results could have been favored due to its general superior performance (Elith et al. 2006), which is also evident with reduced sample sizes (Hernandez et al. 2006; Pearson et al. 2007; Wisz et al. 2008). MAXENT's use of a regularization parameter is also taken as a predictor that this algorithm reduces the chances of overfitting (Phillips & Dudik 2008), which could explain its higher omission rates when the entire set of presences is considered. In fact, overfitting may be an issue for DOMAIN models, for which all the observed records fell within areas of high occurrence probability, although in general this algorithm also presented higher predicted distribution ranges. As the main aim of the provided SDMs is to guide future field searches enhancing the knowledge of drosophilid diversity in poorly known regions, areas predicted to be inhabited by a given species or species set by both modeling algorithms certainly provide a better search strategy.

Conclusion

This study adds to the knowledge about the distribution ranges of mycophagous species of Drosophilidae, by expanding the recognized distribution, northwards and/or southwards, of 22 species of *Hirtodrosophila, Mycodrosophila, Paraliodrosophila* and *Zygothrica*. It also contributes to the understanding of ecological and evolutionary processes that can affect the distribution of 13 species of these genera, by modeling their environmental suitability patterns. It was detected that among the environmental variables with most influence on the potential distribution of species, susceptibility to increasing precipitation seasonalities and decreasing precipitation conditions (especially in

warmer months, due to the drying power of the air) appear to be widely conserved among the studied mycophagous species. Suitability ranges for the minimal temperature of the coldest month are more variable and associated with different responses, with some species putatively sensitive to the temperature rises expected across the Neotropics over the next decades and centuries (Intergovernmental Panel on Climate Change 2007). The suggestion of niche conservatism may also hinder the perspective of potential range shifts, highlighting the risk that some mycophagous Drosophilidae species may become threatened in the next years.

Acknowledgments

We are grateful to the authorities responsible for Brazilian Ecological Stations, Reserves or Parks and to the private properties owners that allowed us to perform Drosophilidae searches in their area, and that are accordingly listed in Table 1 of this manuscript. All these collections were authorized by the Brazilian Ministério do Meio Ambiente (MMA), in the form of the Sistema de Autorização e Informação em Biodiversidade (SISBIO), and we thank them for their commitment to ethical guidelines and conservation efforts concerning Brazilian scientific activities. This study was supported by Universal-CNPq [473372/2011-8].

References

Anderson RP, Lew D, Peterson AT. 2003. Evaluating predictive models of species' distribution: criteria for selecting optimal models. Ecol Model. 162(3):211–232.

Bächli G. 2013. TaxoDros: The database on taxonomy of Drosophilidae [Internet]. version 1.04 (2013/06). Available from: http://www.taxodros.uzh.ch/

Bächli G, Vilela CR, Escher SA, Saura A. 2005. The Drosophilidae (Diptera) of Fennoscandia and Denmark. Fauna Entomologica Scandinavica (Book 39). Leiden (The Netherlands): Brill Academic Publishers.

Burla H. 1956. Die Drosophilidengattung *Zygothrica* und ihre Beziehung zur *Drosophila*-Untergattung *Hirtodrosophila*. Mitt Mus Naturkunde Berl, Zoolog Reihe. 32(2):189–321.

Carpenter G, Gillson AN, Winter J. 1993. DOMAIN: a flexible modeling procedure for mapping potential distributions of plants and animals. Biodivers Conserv. 2(6):667–680.

Carson HL. 1971. The ecology of *Drosophila* breeding sites. Harold L. Lyon Arboretum Lecture Number Two. The University of Hawaii, Honolulu; p. 1–27.

Courtney SP, Kibota TT, Singleton TA. 1990. Ecology of Mushroom-feeding Drosophilidae. Adv Ecol Res. 20:225–274.

DaLage JL, Dergoat GJ, Maczkowiak F, Silvain JF, Cariou, ML, Lachaise D. 2007. A phylogeny of Drosophilidae using the Amyrel gene: questioning the *Drosophila melanogaster* species group boundaries. J Zoolog Syst Evol Res. 45(1):47–63.

Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 29(2): 129–151.

Elith J, Leathwick JR. 2009. Species distribution models: ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst. 40:677–697.

Fielding AH, Bell JF. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv. 24(1):38–49.

Frota-Pessoa O. 1945. Sobre o subgênero *Hirtodrosophila*, com descrição de uma nova espécie (Diptera, Drosophilidae, Drosophila). Rev Bras Entomol. 5(4):469–483.

Gottschalk MS, Bizzo L, Döge JS, Profes MS, Hofmann PRP, Valente VLS. 2009. Drosophilidae (Diptera) associated to fungi: differential use of resources in anthropic and Atlantic Rain Forest areas. Iheringia Ser Zool. 99(4):442–448.

Gottschalk MS, Hofmann PRP, Valente VLS. 2008. Diptera, Drosophilidae: historical occurrence in Brazil. Check List. 4(4):485–518.

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

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

Guisan A, Thuiller W. 2005. Predicting species distribution: offering more than simple habitat models. Ecol Lett. 8(9):993–1009.

Hammer O, Harper DAT, Ryan PD. 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electronica. 4(1): 9.

Hernandez PA, Graham CH, Master LL, Albert DL. 2006. The effect of sample size and species characteristics on performance of different species distribution modelling methods. Ecography. 29(5): 773–785.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 25(15):1965–1978.

Intergovernmental Panel on Climate Change. 2007. Climate change 2007: synthesis report. Geneva (Switzerland): Intergovernmental Panel on Climate Change.

Jaenike J. 1978. Resource predictability and niche breadth in the *Drosophila quinaria* species group. Evolution. 32(3):676–678.

Kellermann V, Loeschcke V, Hoffmann AA, Kristensen TN, Flojgaard C, David JR, Svenning JC, Overgaard J. 2012. Phylogenetic constraints in key functional traits behind species' climate niches: patterns of desiccation and cold resistance across 95 *Drosophila* species. Evolution. 66(11):3377–3389.

Kellermann V, Overgaard J, Hoffmann AA, Flojgaard C, Svenning JC, Loeschcke V. 2012. Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically. Proc Natl Acad Sci USA. 109(40):16228–16233.

Lacy RC. 1984. Predictability, toxicity, and trophic-niche breadth in fungus-feeding Drosophilidae (Diptera). Ecol Entomol. 9(1):43–54.

Markow TA, O'Grady P. 2008. Reproductive ecology of *Drosophila*. Funct Ecol. 22(5):747–759.

Ohenoja E, Metsanheimo K. 1982. Phenology and fruiting body production of macrofungi in subarctic Finnish Lapland. In: Laursen GA, Ammirati JF, editors. Arctic and alpine mycology. Seattle (WA): University of Washington Press; p. 390–409.

Pearson RG. 2007. Species' distribution modeling for conservation educators and practitioners [Internet]. Synthesis. American Museum of Natural History. Available from: http://ncep.amnh.org

Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J Biogeogr. 34(1):102–117.

Phillips JS, Dudík M, Schapire RE. 2004. A maximum entropy approach to species distribution modeling. In: Proceedings of the Twenty-First International Conference on Machine Learning. New York: AMC Press; p. 655–662.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. Ecol Model. 190(3–4):231–259.

Phillips SJ, Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. 31(2):161–175.

Powell JR. 1997. Progress and prospects in evolutionary biology: the *Drosophila* model. New York: Oxford University Press; p. 512.

Raxworthy CJ, Ingram CM, Rabibisoa N, Pearson RG. 2007. Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using Day Geckos (*Phelsuma*) of Madagascar. Syst Biol. 56(6):907–923.

Raxworthy CJ, Martínez-Meyer E, Horning N, Nussbaum RA, Schneider GE, Ortega-Huerta MA, Peterson AT. 2003. Predicting distributions of known and unknown reptile species in Madagascar. Nature. 426(6968):837–841.

Robe LJ, Valente VLS, Loreto ELS. 2010. Phylogenetic relationships and macro-evolutionary patterns within the *Drosophila tripunctata* radiation (Diptera: Drosophilidae). Genetica. 138(7):725–735.

Shorrocks B, Charlesworth P. 1980. The distribution and abundances of the British fungal-feeding *Drosophila*. Ecol Entomol. 5(1):61–78.

Spieth HT. 1987. The *Drosophila* fauna of a native California forest (Diptera: Drosophilidae). Pan-Pac Entomol. 63:247–255.

Throckmorton LH. 1975. The phylogeny, ecology and geography of *Drosophila*. In: King RC, editor. Handbook of genetics. New York: Plenum; p. 421–469.

van der Linde K, Houle D, Spicer GS, Steppan SJ. 2010. A supermatrix-based molecular phylogeny of the family Drosophilidae. Genet Res (Camb). 92(1):25–38.

Vilela CR, Bächli G. 2004. On the identities of nine Neotropical species of *Hirtodrosophila* (Diptera, Drosophilidae). Mitt Schweiz Entomol Ges. 77:161–195.

Vilela CR, Bächli G. 2007. Revision of the Neotropical genus *Paraliodrosophila* (Diptera, Drosophilidae). Mitt Schweiz Entomol Ges. 80:291–317.

Warren DL, Glor RE, Turelli M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution. 62(11):2868–2883.

Warren DL, Glor RE, Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography. 33(3):607–611.

Wertheim B, Sevenster JG, Eijs IEM, van Alphen JJM. 2000. Species diversity in a mycophagous insect community: the case of spatial aggregation vs. resource partitioning. J Anim Ecol. 69(2):335–351.

Wheeler MR, Takada H. 1971. Male genitalia of some representative genera of American Drosophilidae. Univ Texas Publs. 7103:225–240.

Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. Diversity Distrib. 14(5):763–773.

3.3. Capítulo 3

Título:

Neotropical mycophylic drosophilids (Diptera: Drosophilidae): DNA barcoding as a way of overcoming the taxonomic impediment

Manuscrito submetido ao periódico:

Zoological Journal of the Linnean Society

ISSN:

0269-7653 (versão impressa)
1573-8477 (versão online)

Neotropical mycophylic drosophilids (Diptera: Drosophilidae): DNA barcoding as a way of overcoming the taxonomic impediment

Stela Machado[1] stelamchd@gmail.com

João Pedro Junges dos Santos[1] jungesj@gmail.com

Pedro Mesquita Fonseca[1] pedro.graal@gmail.com

Andreza Ribeiro Bolzan[1] andreza.bolzan@gmail.com

Jean David[2,3] Jean.David@legs.cnrs-gif.fr

Elgion Lucio da Silva Loreto[1] elgion@base.ufsm.br

Marco Silva Gottschalk[4] gotts007@yahoo.com

Lizandra Jaqueline Robe[1,5] lizbiogen@gmail.com

[1] Universidade Federal de Santa Maria (UFSM), Programa de Pós Graduação em Biodiversidade Animal, Avenida Roraima 1000, CEP 97105-900 - Santa Maria, RS - Brazil, Phone: +55 55 3220-8912

[2] Laboratoire Evolution, Génomes, Spéciation (LEGS), CNRS, 91198 Gif sur Yvette France et Université Paris-Sud 11, 91405 Orsay F

[3] Département Systématique et Evolution, Muséum National d'Histoire Naturelle, UMR 5202 (OSEB) 45 rue Buffon, 75014 Paris, F

[4] Universidade Federal de Pelotas (UFPel), Instituto de Biologia, Departamento de Ecologia, Zoologia e Genética, Campus Capão do Leão, Mail Box 354, CEP 96010-900 - Capão do Leão, RS – Brazil, Phone: +55 53 3275-7340

[5] Universidade Federal do Rio Grande (FURG), Instituto de Ciências Biológicas – ICB, Av. Itália, km 08 Campus Carreiros, CEP 96203-900 - Rio Grande, RS – Brasil, Phone: +55 53 3293-5128

Corresponding Author:

Lizandra Jaqueline Robe

Short running title: DNA Barcode of mycophylic drosophilids

Abstract

Within Drosophilidae, the *Zygothrica* genus group corresponds to a single speciose mycophylic clade, which remains insufficiently investigated because: most species cannot be reared in laboratory; most species are collected in small number or as occasional samplings; the taxonomic impediment and related difficulties involving the correct identification of these specimens. Thus, this group encompasses an excellent opportunity to testing and implementing the barcode molecular method for analyzing species. Extensive prospections in various parts of Brazil provided 222 individuals which could be distributed into three categories: 22 taxa corresponding to previously described species; 24 corresponding very likely to undescribed species and 10 to dubious cases corresponding mainly to single or few females. Intraspecific distances ranged from 0% to 5.3%, whereas interspecific congeneric and intergeneric distances ranged from 2.4% to 21.7% and from 7.4% to 20%, respectively. Despite the absence of a barcoding gap within two genera, all species were revealed as reciprocally monophyletic, and 62.5% of them presented diagnostic-characters. Coalescent-based methodologies were also congruent with these results, never lumping species initially considered different. Several putative new species were cryptic in regard to other, highlighting the potential of COI barcodes as a tool complementary to traditional taxonomy in unraveling cryptic diversity.

Keywords: Cryptic species - Drosophilinae - *Hirtodrosophila* - Integrative approach - mycophagous drosophilid - *Mycodrosophila* - monophyly - numt's detection - *Paraliodrosophila* - *Zygothrica*.

Introduction

Neotropical region harbors a great diversity of species (Lewinsohn & Prado, 2005; Giam *et al*., 2012), encompassing several major biodiversity hotspots of the world (Myers *et al*., 2000). This region covers some of the recognized Megadiverse countries (Mittermeier *et al*., 1997; Mittermeier, 1988), among which we can highlight Brazil, which hosts approximately 10-13% of the world's biodiversity (Lewinsohn & Prado, 2005). However, as occurs in other Megadiverse countries, most of Brazilian diversity remains unknown, with estimates that for each described species, ten remain undescribed (Lewinsohn & Prado, 2005; Landim & Hingst-Zaher, 2010). One of the major issues for cataloging this diversity is the taxonomic impediment, related to the reduced number of specialized taxonomists as compared to the high number of species to be described (Lewinsohn & Prado, 2005; Kvist, 2013). This is especially troublesome since most studies in biology, from inventories to molecular analysis, require reliable species identification, which is a hard and longstanding task for some groups of species, principally in small and diverse groups like insects. DNA barcoding methodology has emerged as a way to help non-specialists to overcome these difficulties (Hebert & Gregory, 2005).

DNA barcoding proposes the use of a standardized DNA region to (1) identify and assign unknown individuals to described species, and (2) enhance the discovery of new species (Hebert *et al*., 2003a). For animals, the gene region proposed as the standard barcode is a fragment encompassing 658 bp of the 5' section of the mitochondrial cytochrome c oxidase subunit I (COI) (Hebert *et al.*, 2003a; Hebert, Ratnasingham & deWaard, 2003b). DNA barcoding intends to work within the Linnaean system, preserving the principles by which species are named and classified, but establishing a taxonomic framework based on DNA sequence data (Hebert & Gregory, 2005; Goldstein & DeSalle, 2010), and has been tested in diverse animal groups (Clare *et al*., 2011; Tavares *et al*., 2011; Robe, Machado & Bartholomei-Santos, 2012; Pramual & Adler, 2014; Zahiri *et al*., 2014).

Since its original proposal, DNA barcoding has largely enhanced cryptic species discovery and potential new species assignment (for some examples see Hebert *et al*., 2004; Clare *et al*., 2011; Crawford *et al*., 2012; Mutanen et al., 2013; Gill *et al*., 2014). Despite these scientific advances and the fact that an integrative approach is needed to confirm each putative new species (Hebert & Gregory, 2005),

some traditional taxonomists are refractory to DNA barcoding, since they see it as a threat to traditional taxonomy and conventional species description (Ebach & Holdrege, 2005; Will, Mishler & Wheeler, 2005). Other shortcomings of this methodology have also been pointed (Moritz & Cicero, 2004; DeSalle, Egan & Siddall, 2005; Hurst & Jiggins, 2005; Rubinoff, 2006; Galtier *et al*., 2009; Jinbo, Kato & Ito, 2011; Leite, 2012; Srivathsan & Meier, 2012), and these are associated to: (1) potential biological and molecular idiosyncrasies of mitochondrial DNA evolution, that may occur due to mitochondrial introgression, endosymbiont infection in Arthropoda with mitochondrial hitchhiking, or nuclear copies of mitochondrial DNA fragments (numts), for example; (2) the consequences of non-uniformity between speciation processes, leading to incomplete lineate sorting coupled with recent speciation, or overlap between intra and interspecific distances due to variation in the evolutionary rates; and (3) methodological issues related to the effects of distance, character and tree-based analysis in species boundary recognition. Thus, it is important that the effectiveness of the DNA barcoding approach is individually tested in each taxonomic group before its thorough usage in order to identify occurrence and effects of each of these issues.

There are several biases related to the historical implementation of DNA barcode, with differential usage among taxonomic groups (Kvist, 2013) and geographical regions (Porter *et al*., 2014). Arthropoda, for example, is a speciose group, which received a great deal of attention, but still has only 12% of the formally described species represented in Barcoding reference databases (Kvist, 2013). Concerning geographic sampling, Porter *et al*., (2014) found that some orders of insect are represented in iBOL mainly by sequences derived from a single country, with 80% of Diptera sequences coming from Canada. The effectiveness of DNA barcoding in species identification is highly dependent on database coverage (Kvist, 2013; Porter *et al*., 2014), and in groups with high rates of intraspecific diversity like insects, this issue becomes even more critical (Meier *et al*., 2006; Yassin *et al*., 2010; Zahiri *et al*., 2014).

Drosophilidae is a diverse group with more than 4,200 described species distributed in 77 genera (Bächli, 2014). With the exception of Polar Regions, this diversity is distributed all around the planet, although greater diversity levels are found in woods and forests, principally in the tropical region (Throckmorton, 1975). Throughout this area, the family is also highly diversified in terms of resource use,

with species adapted to the use of fruits, fungus, sap, flowers, pollen, among others (Throckmorton, 1975; Val, Marques & Vilela, 1981; Tosi *et al*., 1990; Markow & O'Grady, 2008; O'Grady & Markow, 2009). However, this diversity is certainly even greater, and studies performed in the Neotropical region often find unrecognized species (Medeiros & Klaczko, 2004; Gottschalk *et al.*, 2007; Gottschalk, Hofmann & Valente, 2008; Gottschalk *et al*., 2009; Valadão, Hay & Tidon, 2010; Garcia *et al*., 2012; Poppe *et al*., 2014). Another important issue for Neotropical drosophilids is the traditional use of traps with fermented fruits in collections, since these attract mainly frugivorous species of the genus *Drosophila* (Gottschalk *et al*., 2008; Robe *et al*., 2014). Thus, the diversity of Drosophilidae species associated with other resources is barely assessed, being mainly recorded in pinpoint description works (Gottschalk *et al*., 2008).

The *Zygothrica* genus group encompasses mycophylic species of *Hirtodrosophila*, *Mycodrosophila*, *Paraliodrosophila* and *Zygothrica* (Grimaldi, 1990a), and is a good example of this subsampling scenario (Gottschalk *et al*., 2009; Robe *et al*., 2014). Although Grimaldi (1990a) suggested that its association with macroscopic fungi for oviposition, feeding and/or courtship may be a synapomorphy, some evidences suggest the *Zygothrica* genus group may not be monophyletic at all, given the putative para/polyphyly of *Hirtodrosophila* (DaLage *et al*., 2007; van der Linde *et al*., 2010; Yassin, 2013). Despite Courtney, Kibota & Singleton (1990) assigned to these mycophylyc species a great potential to ecologic studies, since then, few studies were performed (Rouquette & Davis, 2003; Yorozuya, 2006; Tuno *et al*., 2007; Yorozuya, 2009; Stump *et al*., 2011; Kasuya *et al*., 2013; Robe *et al*., 2014). In fact, little is known about the patterns and processes related to their evolution and ecology, especially as Neotropical species are concerned.

The aim of the present study is to explore COI sequence divergence within and between species of Neotropical mycophylic drosophilids that are included in the *Zygothrica* genus group and contrast the obtained patterns with those obtained for other drosophilid species. We ask whether COI sequences of a set of collected species present the DNA barcoding desired properties necessary for effectiveness of distance-based, phylogenetic-based, character and coalescence-based methods, in a perspective that once this potential is confirmed, prompt identification could promote a change in the incomprehension scenario in which these species are merged. Moreover, while confirming potential utility of COI barcodes, this study *per*

se represents a step forward in the work toward the major goal of understanding Neotropical mycophylic drosophilid's biodiversity, since it uncovers high levels of cryptic and unknown diversity while questioning their generic boundaries.

Material and methods

Sampling

Collections were performed in 20 sites distributed in Brazilian Amazonian, Atlantic Forest and Pampa biomes (Fig. 1, Fig. S1, Tables S1 and S2). Each site was coursed looking for macroscopic fungal fruiting bodies, and when adult drosophilids were encountered resting or flying over them, they were aspirated with a modified entomological aspirator (Machado *et al*., 2014). In some cases, the fungi were carried and maintained in laboratory until adults emerged, under 25± 1°C, with constant light and humidity. The captured flies were stored in absolute ethanol and identified through their external morphology and male genitalia patterns through the use of identification keys and aedeagus comparisons with illustrations available in the literature (Frota-Pessoa, 1945; Burla, 1956; Wheeler & Takada, 1963; Grimaldi, 1987, 1990a, 1990b; Vilela & Bächli, 1990, 2004, 2005, 2007). Genitalia slides were prepared following Wheeler & Kambysellis (1966), with the modifications described by Kaneshiro (1969). Females were provisionally allocated to the name of the most similar male collected in the same point and later defined based on sequence identity/positioning in relation to male's barcodes, since diagnostic characters are absent for this gender in most species of the group. Anyway, the terminalia and spermathecae of several females were removed and treated with the same technic as the genitalia of males, and permanent slides were prepared. All male and female genitalias slides were deposited in the Drosophilidae Collection of the Museu de Ciências Naturais Carlos Ritter of Universidade Federal de Pelotas (MCNCR-UFPel) as vouchers material.

(Place to Figure 1)

DNA manipulation

Total DNA was extracted from each individual specimen using the NucleoSpin Tissue XS kit (MACHEREY-NAGEL). An amplicon spanning the 5' region of the mitochondrial cytochrome c oxidase subunit I (COI) was amplified with the use of the

following combinations of primers: TYJ1460 (Simon *et al*., 1994) and C1N2329M (5'-ACT GTA AAT ATA TGA TGA GCT CAT ACA -3'), modified from Simon *et al*., (1994), HCO1490 and LCO2198 (Folmer *et al*., 1994), TYJ1460 and LCO2198, and HCO1419 and C1N2329M. The PCR reactions contained around 10-100 ng of total DNA, 100-250 µM of each dNTP, 1.5-2.5 mM of $MgCl_2$, 0.15-0.3 µM of each primer, 1x buffer and 0.5 units of Platinum Taq DNA Polymerase (Invitrogen). The amplifications started with an initial denaturation at 94 ºC (3 min), followed by 11 cycles at 94 ºC (45 s), 55 ºC-0.5 ºC/cycle (45 s), 72 ºC (50 s) and 24 cycles at 94 ºC (35 s), 50ºC (35 s), 72 ºC (50 s) and finished with a final extension at 72 ºC (10 min). In some cases, individual optimization of PCR cycling pattern was necessary. The PCR products were directly sequenced using the same primers from the PCR reaction. Sequencing reactions were carried on a MegaBACE 500 automatic sequencer with the use of the DYEnamic ET® Sequencing Kit (Amersham) or by Macrogen (http://www.macrogen.com/eng/).

Data analysis

Sequence manipulation

The generated electropherograms were assembled with the use of the Staden Package Gap 4 program (Staden, 1996) and each obtained contig was individually checked for adequate sequence quality and accuracy in each polymorphic site visualized in the intraspecific alignment. Homologous sequences were aligned with the Clustal W algorithm, as implemented in Mega 6 (Tamura *et al*., 2013) and checked for the presence of stop-codons and frameshifts in order to avoid the presence of nuclear mitochondrial pseudogenes (Numts) in our data matrices. Additionally, the ratio between the number of nonsynonymous substitutions per nonsynonymous sites (dN) and the number of synonymous substitutions per synonymous sites (dS) was assessed in Mega 6 using Nei-Gojobori method (Nei & Gojobori, 1986). A codon-based Z test was also conducted in order to check pairwise probabilities of rejecting the null hypothesis of strict-neutrality (dN = dS) in favor of the alternative hypothesis of purifying selection (dN < dS). Sequences that did not reject neutrality in this test were further evaluated in regard to their divergence and dN/dS ratios in order to further control for the presence of recent Numts, that did not yet attain frameshifts or nonsense mutations (for more details,

see Robe *et al*., 2012). The sequences that passed this Numt detection protocol were deposited in Bold system (Table S3), and those that presented warning signals were deposited in GenBank (Accession numbers KP317996 and KP317997).

These newly obtained mycophylic drosophilid sequences were further added to the unique two determined COI sequences available in GenBank for the focused genera (from *H. trilineata* - EU126511 and *M. dimidiata* - EU493682) and to one sequence from *Stegana adentata* (Steganinae), used as an outgroup - HQ842774. This alignment encompassed the first database used in the analyses (hereafter referred as database 1 or mycophylic database). A second dataset (hereafter referred as database 2 or Drosophilinae database) was constructed with these sequences added to all COI sequences available for Drosophilinae in GenBank, in order to test the stability of mycophylic DNA barcode results in a broader scale. Available Drosophilinae COI sequences were tracked through the use of the keywords “Drosophilidae AND COI” in the GenBank Nucleotide search tool. Nevertheless, from the 2,372 recovered results, only 1,975 nucleotide sequences presented at least 500bp of overlap with database 1, and these were included in database 2 (Table S4).

Tree-based analysis

Neighbor-joining trees (NJ) were generated for database 1 and 2 using K2P distance model (Kimura, 1980), as implemented in Mega 6. Based on these preliminary trees, species status of males was revised through additional evaluation of male genitalia slides and in some cases of photographs taken from the entire specimens. Females that had only provisional names previously attributed (see above) were further renamed based on their nesting patterns relative to male clusters.

Distance-based analysis

Pairwise distances were calculated for mycophylic drosophilids using K2P distance model as implemented in Mega 6. This software was also used to obtain minimum, maximum and mean distance values within each taxonomic category (intraspecific, interspecific congeneric and interspecific intergeneric), with standard error estimated through 1,000 bootstrap replicates. Based on these measures, taxonomic resolution ratio (Costa *et al*., 2007) and barcoding gap were calculated. To

visually infer the presence/absence of barcoding gap, a boxplot graphic was generated in Spider package (Brown *et al*., 2012) based on the distribution of intra and interspecific distances.

In order to establish which threshold is more adequate to our mycophylic dataset, different strategies were used: (i) the function "localMinima" was used to obtain the best threshold as a function of the transition between intra and inter-specific distances based on the density of genetic distances, without prior knowledge of species identity, as performed in Spider package; (ii) the assumption-free approach developed by Lefébure *et al*. (2006) was implemented in order to determine the threshold which simultaneously minimizes the percentage of intra and interspecific misdiagnosis; (iii) the Automatic Barcode Gap Discovery tool (ABGD) (Puillandre *et al*., 2012) was used to detect the barcoding gap beyond the confidence limits of pairwise intraspecific distances which partitions the dataset into the maximum number of groups (hypothetical candidate species) such that the distance between sequences from different groups is larger than the threshold. ABGD was also used to test the efficiency of each of the above thresholds, and of the fixed 1% and 3% threshold criteria of BOLD and Hebert *et al*. (2003a), respectively.

Character-based analysis

As distance and tree-based analyses have been frequently criticized as arbitrary criterions for distinguishing species (DeSalle *et al*., 2005), we also performed a character-based analysis, which has been suggested as a robust alternative methodology (DeSalle *et al*., 2005). In fact, it was argued that this method provides better results with missing and small datasets, and does not incur in false-positive identifications (Kelly *et al*., 2007). For this task, we used the "nucDiag" function implemented in Spider package to generate a list of diagnostic nucleotides for both species and genera, in datasets 1 and 2.

Coalescence and Bayesian-based analysis

One additional limitation of most tree and distance-based analyses is that species are previously assigned based on morphological and/or populational criteria and then barcoding gap, thresholds and/or reciprocal monophyly are used to define the species boundaries (Fujisawa & Barraclough, 2013). To overcome this hurdle, Bayesian methods have been proposed, like the Generalized Mixed Yule Coalescent

(GMYC) method (Pons *et al*., 2006; Fontaneto *et al*., 2007; Fujisawa & Barraclough, 2013). This method is based on the detection of transitions between inter and intra-specific evolutionary processes, where distinct genetic clusters that diversified into independent lineages will be assigned to different species, rejecting the null hypothesis that all individuals belong to the same species (for details see Fujisawa & Barraclough, 2013). GMYC requires an ultrametric tree in order to perform the analysis, so we used Beast 1.7.4 (Drummond *et al*., 2012) to generate a cronophylogenetic tree under a model of uncorrelated lognormal substitution rates and an "ucld.mean" prior set to 1.0. The run was performed using Yule pure birth model tree prior and GTR+I+G nucleotide substitution model, as chosen by jModelTest 2 (Darriba *et al*., 2012), with 50,000,000 iterations sampling each 1,000, and discarding 50% of the sample as burn-in. The runs were visually inspected using TRACER 1.6 and summarized in TreeAnnotator 1.7.1. The single threshold analysis of GMYC was performed using Splits package (https://github.com/tfujisawa/splits_tmp) in R.

Results

COI sequences generated for mycophylic drosophilids range from 537 to 717pb. Altogether, our dataset includes 222 specimens encompassing 56 species, 21 belonging to *Hirtodrosophila* (six of which were previously described, and six of which remained undetermined due to the exclusive presence of females), nine to *Mycodrosophila* (three already described), two to *Paraliodrosophila* (both described) and 24 to *Zygothrica* (11 described) (Table 1, Table S2). As expected for mitochondrial genes, nucleotide composition was highly biased, showing a high general AT content (66.5%). This bias was similar for all specimens, and between first (52%) and second codon positions (58.4%), although third codon positions presented an even higher AT-content (88.7%). Moreover, no strong evidence of Numts was found throughout the mycophylic database, since neither frameshifts nor indels were detected within the sequences. Besides, all pairwise comparisons that did not reject neutrality in favor of the alternative hypothesis of purifying selection occurred in intraspecific comparisons with reduced levels of divergence (0% to 1.2%, with mean of 0.18%). Among these, dN/dS values were higher than 0.33 only in one comparison involving two sequences of *H. subflavohalterata* affinis 5 (dN/dS = 1.86,

distance of 1.2%), so that one or both of these two sequences warn to the putative presence of recent Numt sequences.

(Place to Table 1)

Tree-based analysis

The first Neighbor-joining tree generated showed some minor conflicts between molecular and morphological identification. Nevertheless, in most cases, these occurred with females, a result that was already expected since in several cases female's identity could only be provisionally inferred based on external morphology, with the aid of the morphologically "similar" males collected at the same point. Additionally, in some minor cases, morphological identification of isolated males was discordant with recovered species clades. These cases always occurred among members of a cryptic species pair/group, as *H. mendeli* affinis and *H. subflavohalterata affinis* H002, *H. subflavohalterata* affinis 5 and *Pa. antennata*, and species of the *Z. dispar* group. In all these cases, genitalia slides and external morphology photographs were further checked in order to assess the real identity, which always confirmed the main molecular phylogenetic pattern.

High levels of cryptic diversity were revealed both, at the morphological and molecular levels. At the morphological level, 23 unknown and potentially new species were detected, most of which were cryptic concerning external morphology but presented some differences in male genitalia patterns (Table 1). At contrast, some levels of putatively new cryptic diversity were only revealed at the molecular level, and this applied for *H. levigata* (*H. levigata* affinis), *H. subflavohalterata* 4 (*H. subflavohalterata* 5), *M. neoprojectans* affinis (*M. neoprojectans* affins 1 and 2), *M.* M001 (*M.* affinis M001), *M. projectans* (*M. projectans* affinis 1 and 2), *Zy. laevifrons* (*Zy. laevifrons* 1 and 2) and *Zy. prodispar* (*Zy. prodispar* affinis 1). Nevertheless, in three of the last cases (*H. levigata* affinis, *H. subflavohalterata* affinis 4 and *M. projectans* affinis 1) cryptic diversity was revealed only at the molecular level possibly due to the absence of associated males, which precluded further aedeagus inspection. Joining these results, as 10 of these 34 potentially new sampled morphotypes/MOTU's (Molecular Operational Taxonomic Units) (Floyd *et al*., 2002) were only represented by females, we estimated in 24 the minimum number of putative new species recorded in this study.

After, this initial identity review, the NJ tree was reconstructed, when all mycophylic species were revealed as reciprocally monophyletic with bootstrap values higher than 98%. This result was recovered for both, the mycophylic (Fig. 2) and the Drosophilinae dataset (data no shown). This last result also shows that all the putatively new or even undetermined species are not strictly related to any other Drosophilinae taxon yet sampled for the 5' region of the COI mitochondrial gene. As concerns the interspecific topology, some intrageneric well-supported clades have been recovered [like *H. subflavohalterata* affinis 1, 3, 4, and *H.* H010 (bootstrap of 89%), and *H. subflavohalterata* affinis H002, 2, 5 and *H. mendeli* affinis (bootstrap of 100%)], although all the studied genera appeared intermingled at a basal polytomy (Fig. 2). This polytomy is probably an effect of the high saturation of COI sequences in divergence values higher than 5% (Fig. S2), a property that allowed adequate species delimitation but disallowed further phylogenetic inferences.

(Place to Figure 2)

Distance-based analysis

Intraspecific distances ranged from 0% in all genera, to 5.1%, 3.3%, 1.5%, and 5.3% in the *Hirtodrosophila*, *Mycodrosophila*, *Paraliodrosophila* and *Zygothrica* genera, respectively (Table S5). As the minimum interspecific distances found for *Hirtodrosophila*, *Mycodrosophila*, *Paraliodrosophila* and *Zygothrica* were generally low (2.4%, 4.6%, 13.6% and 3.7%, respectively – Table S5), we could not find a barcoding gap for *Hirtodrosophila* and *Zygothrica*. This overlap between intra and interspecific distances in the *Zygothrica* genus group can be better observed in Figure 3, although most of this effect is due to individual cases of high intraspecific distances, as those found in *H. subgilva* (5.1%) and *Zy. ptilialis* (5.3%), or low interspecific values, as those found between *H. subflavohalterata* affinis 2/*H. subflavohalterata* affinis H002 (minimum of 2.4%), *H. mendeli* affinis with *H. subflavohalterata* affinis 2 or with *H. subflavohalterata* affinis H002 (minimum of 2.7% or 2.6%, respectively) and *Zy. vittimaculosa/Zy. vittimaculosa* affinis (minimum of 3.7%). In the two cases of high intraspecific distances, we preferred to adopt a conservative strategy and consider putatively independent lineages as members of a single species, since they were cryptic even at the male genitalia level and formed a single clade. At the four cases of low interespecific divergences, we maintained the species status for each lineage since they present at least some subtle differentiation

at the genitalia level (case of *H. subflavohaterata* affinis 2 x *H. subflavohaterata* affinis H002 and *Zy. vittimaculosa* x *Zy. vittimaculosa* affinis) or are not cryptic at all (case of *H. subflavohaterata* affinis 2 or H002 and *H. mendeli* affinis). Except for these mentioned isolated cases, the other pairs of cryptic species showed high interspecific distances (minimum of 4.6%) which could be mostly attributed to synonymous substitutions, since dN/dS between cryptic species were always lower than 0.072 (Table 1).The thresholds optimized according to our individual datasets varied according to the approach used in the estimation: whereas low threshold values were suggested by ABGD (1.15-1.5%) and Spider "localMinima" function (1.92%), higher values were suggested through the implementation of the assumption-free approach developed by Lefébure *et al*. (2006) (5.25%). When these and other standard thresholds were tested for their performance, higher success rates were obtained with ABGD (96.85%) (Table S6).

(Place to Figure 3)

Character-based analysis

Diagnostic characters could not be found for any of the four genera of mycophylic drosophilids, although one to five diagnostic characters were detected for 62.5% of the studied species, and these were distributed in 61 of the 717 evaluated sites (Fig. 4). In the case of cryptic species pairs and putatively new species, diagnostic characters were found in 41.1% and 42.9% of the cases, respectively. Interestingly, most of the diagnostic characters are represented by desoxyguanilate nucleotides.

(Place to Figure 4)

Coalescence and Bayesian-based analysis

Using this approach, the null hypothesis was rejected (likelihood null model = 1582.475, maximum likelihood of GMYC model = 1627.831, likelihood ratio test = 90.71, LR test = 0, with a threshold time of -0.01577 specifying the location of nodes defining species). The supported GMYC model suggests the presence of 48 clusters (confidence interval of 42 - 49), when species represented by only one terminal node are excluded, and 77 (confidence interval of 62 - 84) when all terminal nodes are considered (Table S7). In general, GMYC clusters were similar to those presented by the NJ tree, to the exception that it additionally subdivided some species, like *H.*

mendeli, *H. morgani*, *Zy. orbitalis* (4 clusters) and *Zy. ptilialis* (5 clusters) (Fig. 2, Table S7). Nevertheless, GMYC did never cluster species considered different in the phylogeny-based, tree-based and character-based methodologies. Comparing the clusters suggested by GMYC and ABGD, it was possible to see that ABGD thresholds ranging from 0.14% - 0.3% better resemble GMYC partitioning.

Discussion

Neotropical mycophylic drosophilids' diversity

Although the *Zygothrica* genus group's diversity is centered on circumtropical regions (Wheeler & Takada, 1963; Vilela & Bächli, 2004; Grimaldi, 2010), it is not limited to this area. *Hirtodrosophila* (160 species – Bächli, 2014) and *Mycodrosophila* (127 species – Bächli, 2014) are considered cosmopolitan genera (Wheeler & Takada, 1963; Vilela & Bächli, 2004; Grimaldi, 2010), whereas *Zygothrica* (124 species – Bächli, 2014) is found in Neotropical and Indopacific regions (Grimaldi, 1987) and the small genus *Paraliodrosophila* (five species – Bächli, 2014) has their distribution restricted to Neotropics (Vilela & Bächli, 2007). This genus group encompasses more than 400 described species (Bächli, 2014), but it is certainly even more diverse, since there are yet several species waiting description (Grimaldi, 2010).

In fact, in this work, we demonstrate that Neotropical mycophylic drosophilids are much more diverse than previously recorded. Therefore, besides 22 already described species, we sampled at least 24 putative new species, which should be further described and added to the list of 75 Brazilian species included in the *Zygothrica* genus group (Vilela & Bächli, 2007; Gottschalk *et al*., 2008; Robe *et al*., 2014). Nevertheless, cryptic or unregistered diversity levels are probably yet underestimated, since other ten independent MOTUs could not be adequately evaluated at the morphological level since they were represented only by females, although general external characteristics already suggested them as putative new species.

Among the 24 putative new species detected here, 16 were first recognized as undescribed by the co-author specialized taxonomists (JPJS and MSG) (Table 2). Despite the fact that in several cases these highlighted species were externally cryptic to some other determined sampled species, DNA barcoding corroborated

their independent status. For the remaining eight putative new species, even the aedeagus structure was cryptic to other sampled species, and evolutionary independency was first suggested by the phylogenetic approach, and afterwards corroborated by different DNA barcoding methodological strategies (Table 2). So, although it is frequently recognized that the phylogenetic species concept tends to overestimate the number of species (Thorp *et al*., 2010), the general congruency among different methodologies adds support to the high putative new species levels (see below). In fact, seven of the eight entirely cryptic putative new species presented diagnostic characters and the minimum interspecific distance presented in each of these eight cases were of the order of 5-14%.

(Place to Table 2)

High levels of cryptic/unknown diversity were already expected since this group has been widely neglected in previous ecological and evolutionary studies related to Neotropical drosophilids. In fact, Brazilian species of the genera *Hirtodrosophila*, *Mycodrosophila, Paraliodrosophila* and *Zygothrica* have been the target group only for descriptions (Frota-Pessoa, 1945; Burla, 1956; Wheeler & Takada, 1963; Grimaldi, 1987, 1990b; Vilela & Bächli, 2004) and some other punctual works (Grimaldi, 1987; Gottschalk *et al*., 2008; Roque & Tidon, 2008; Robe *et al*., 2014). Together with Robe *et al*. (2014), the present work comprises one of the most extensive sampling for mycophylic drosophilids in the Neotropical region, with four collection sites located in the Amazonian biome (north of Brazil), nine and seven sites in the Atlantic Forest and Pampa biome, respectively (south of Brazil) (Fig. 1, Fig. S1, Table S1). As concerns the diversity distribution among genera, Grimaldi (1987) suggested *Zygothrica* as the most speciose lineage among Neotropical Drosophilidae, but *Hirtodrosophila* also detached for its high levels of previously unrecognized/cryptic diversity (9-15 putative new species).

A frequent concern in DNA barcoding studies refers to the co-amplification of Numts (Song *et al*., 2008; Leite, 2012), which could lead to an overestimation of diversity levels (Song *et al*., 2008). Nevertheless, in this study, indirect search for such artifacts revealed only a single warning case, which was not even related to new species assignment. Since Numts tend to lose function after incorporation to nuclear genome, they generally evolve under neutrality and progressively degenerate, presenting high frequencies of nonsynonymous substitutions, indels causing frameshifts and substitutions leading to stop codons (Funk & Omland, 2003).

This makes them often easily identified (Song *et al*., 2008). In our case, 99.5% of our sequences did not fit this profile, and as the single warning case still presents a limited divergence to the main sequences of the host species, this likely represents a case of recent nuclear incorporation, that does not affect the confidence and consistence of our results.

Congruence between traditional and DNA barcoding approaches

For drosophilids, species recognition is initially performed based on external morphology patterns, although in most cases this must be complemented by the analysis of male genitalia slides, principally aedeagus structure (Vilela & Bächli, 1990; Vilela, 1992; Medeiros & Klaczko, 2004). This happens because external properties like thorax or abdomen color patterns are cryptic between many species, making identification a hard task for males and even an impossible task for several female specimens. For this reason, revision of species status in the first constructed NJ tree based on accessory photographic material and male genitalia slides was essential to confirm species bordering.

Indeed, after isolated cases of status review, all species were confirmed as reciprocally monophyletic in regard to each other. This is a desired property for DNA barcoding approaches based on the use of phylogenetic trees to assign unknown specimens to known or new species, supporting species status under most species concepts (Funk & Omland, 2003). Notwithstanding, monophyly is not a universal feature in animal species, and around 19% - 23% of the metazoan species seem to be para or polyphyletic (Funk & Omland, 2003; Ross, 2014). Five reasons can explain such instances of non-monophyly (Funk & Omland, 2003): (1) imperfect taxonomy, with oversplitting or overlumping; (2) incomplete lineage sorting; (3) interspecific hybridization; (4) unrecognized paraphyly; (5) inadequate phylogenetic information. Although second and third reasons are dependent on species history, the fact that an integrative approach is being used here to settle taxonomic gaps certainly is helping to control the first reason. In fact, it was previously suggested that integrative approaches joining morphological, ecological, behavior and/or molecular aspects generally produce better results in hypothetical new species assignment/corroboration (DeSalle *et al*., 2005; Will *et al*., 2005; Yassin *et al*., 2010).

Distance analysis showed that mycophylic drosophilds present moderated levels of intraspecific diversity, with only two species (*H. subgilva* and *Zy. ptilialis*)

presenting maximum intraspecific distance values higher than 3.5% (Table 1). These two cases of high intraspecific distances and some isolated comparisons with low interspecific distances (referring mainly to *H. mendeli* affinis x *H. suflavohalterata* affinis 2 x *H. suflavohalterata* affinis H002, and *Zy. vittimaculosa* x *Zy. vittimaculosa* affinis) explain the absence of a barcoding gap both in the *Hirtodrosophila* and *Zygothrica* genera. Despite the absence of the desired property of a barcoding gap (Meyer & Paulay, 2005), the threshold of 1.15-1.5% showed low error rates within the entire *Zygothrica* genus group (3.15%, Table S6). The success rates attained with BOLD 1% (Ratnasingham & Hebert, 2007) (96.4%) was quite similar to this best threshold, although higher values showed considerable worst performances (Table S6). This highlights the importance of optimizing threshold choice for each taxonomic group (Robe *et al*., 2012), which is biologically meaningful since genetic variation depends on a variety of factors, as group-specific rates of molecular evolution, individual population history, population sizes, inbreeding, spatial organization and social organization (to a review see Amos & Harwood, 1998; Bromham, 2009).

As traditional taxonomy is based on diagnostic characters, DeSalle *et al*. (2005) suggested integrating molecular data through the use of diagnostic nucleotide analysis together with other evidences to determine species boundaries. Kelly *et al*. (2007) showed that this method is accurate, independent of the extent of the barcoding gap, and Yassin *et al*. (2010) stated that it performs better than other methods when species are paraphyletic. Using this methodology, we found diagnostic characters for 62.5% of the evaluated mycophylic drosophild species, although this success rate was highly reduced when the entire subfamily was considered. In agreement with this, Kerr *et al*. (2009) found that this approach is severely constrained by the number of species included for rule generation, despite its success with smaller datasets. These results are probably an outcome of the high nucleotide divergence in synonymous sites, which are probably very homoplastic, especially considering the highly biased A-T content of the employed gene marker. Saturation also helps to explain the absence of diagnostic characters for any of the four studied mycophylic genera, although this is probably not the complete explanation.

DNA barcode uses a typological species concept, where species are considered as discrete units (Casiraghi *et al*., 2010), and all the analyses presented above embrace at a greater or lesser extent this concept. The recently proposed

Generalized Mixed Yule Coalescent (GMYC) method (Pons *et al*., 2006; Fontaneto *et al*., 2007; Fujisawa & Barraclough, 2013) uses the premise of independent evolution under a coalescent approach, in which organisms that evolve independently due to some barrier preventing gene flow are considered as individual species (Fujisawa & Barraclough, 2013). Although conceptually different from the other applied methodologies, which also require *a priori* information about species delimitations, in general, GMYC *a posteriori* designations were concordant with those recovered with the traditional approaches (Fig. 2, Table S7), only splitting, but never lumping species previously separated. The tendency of GMYC to overestimate species number was previously reported by other authors (Esselstyn *et al*., 2012; Paz & Crawford, 2012; Miralles & Vences, 2013; Talavera, Dinca & Vila, 2013; Kekkonen & Hebert, 2014), although it is commonly treated as an artifact of geographic distance (Kekkonen & Hebert, 2014). Thus, future studies focused on the phylogeographic patterns of the subdivided species will provide better evidences to evaluate if this method is in fact overestimating species delimitations or if it is simply more sensitive than others, detecting species structure even when few individuals are available.

The set of results presented here suggest that DNA barcode may be much more promising for the mycophylic *Zygothrica* genus group than it is for the *Drosophila* genus, for which 26% of the species did not present a barcoding gap, and 23% were not monophyletic (Yassin *et al*., 2010). These results could be a by-product of the particular evolutionary history of distinct Drosophilidae lineages, with different speciation times and non-homogeneous occurrence of hybridization and/or *Wolbachia* infection with mitochondrial hitchhiking (Yassin *et al*., 2010), or may be associated to sampling shortages (Moritz & Cicero, 2004; Linares *et al*., 2009) in our study. This last contention, nevertheless, seems not to apply here, since the number of sampled species is similar across studies [56 x 68 here and in Yassin *et al*. (2010), respectively], and the number of individuals sampled per species were above our sampling range (1-18) for only five of the 17 *Drosophila* species for which the criteria of monophyly and presence of a barcoding gap were not accomplished (Yassin *et al*., 2010). Alternatively, as Yassin *et al*. (2010) only used COI sequences downloaded from GenBank, these have not gone through a careful process of morphological revision as the one used here to generate mycophylic sequences, and some of them may have been subject to identification errors. The integrative approach used here and the congruence between different methodologies is certainly helping to reinforce

the effectiveness and the importance of DNA barcoding in elucidating the yet poorly accessed Neotropical mycophylic drosophilids' diversity.

Nevertheless, if on the one hand the approach used here suggests the effectiveness of DNA barcode in the designation of drosophilid mycophylic species, genus assignment seems to be seriously hampered. Congeneric and intergeneric intespecific distances are highly overlapped (Table S5) and none genus of the *Zygothrica* genus group is reciprocally monophyletic. Although non-monophyly was previously suggested for some of the evaluated genera (DaLage *et al*., 2007; van der Linde *et al*., 2010; Yassin, 2013), saturation in COI nucleotide substitutions above intraspecific divergence levels compromises the use of our NJ tree to make further phylogenetic inferences (Nagy *et al*., 2012; Wilson, 2011). Notwithstanding, if the presented pattern is confirmed, the whole taxonomy of the group should be revised, possibly through the splitting of at least some of the evaluated genera. This should probably clear the distribution patterns and reduce the diversity levels of each genus, with the putative assignment of entirely Neotropical taxons. Regardless of this response, which should be sought through the design of specific phylogenetic studies using a set of distinct markers, the patterns of species distribution along the obtained phylogram suggest that mycophagous drosophilids are targets of great levels of morphological conservation, with several instances of symplesiomorphies, and/or are subject of frequent phenotypic convergences, since large phylogenetic/phenetic distances at the molecular level are often associated with great levels of similarity at the morphological level.

Conclusion

Despite previous evidences pointing to limitations in the use of DNA barcode within (Yassin *et al*., 2010) or outside *Drosophila* (Meier *et al*., 2006; Wiemers & Fiedler, 2007; Trewick, 2008; Nicholls *et al*., 2012), the results for mycophylic drosophilds seem to be quite different. The test of DNA barcoding in Neotropical species included in the *Zygothrica* genus group provided results that are largely congruent among different molecular approaches and also with morphological boundaries, emphasizing the importance of the DNA barcoding endeavor to enhance species assignments, especially in cases of fragmentary remains and female-specimens. Moreover, DNA barcoding also demonstrated here a high potential of

accelerating species discovery, allowing the assortment of divergent taxa that may represent new species. In some cases, the detection of genetically discrete units facilitated the discovery of morphological differences that were previously unrecognized, whereas in other cases, the encountered deep divergences indicated a lack of genetic cohesion among taxa for which robust morphological differences have not yet arisen.

Anyway, as emphasized by Hebert & Gregory (2005), DNA barcoding aids in delimiting species, and is never sufficient to describe new taxa, so that traditional taxonomy will certainly be fundamental in the corroboration and definitive acceptance of the putative new species. The task of unraveling diversity of mycophylic drosophilids, and promoting greater ease and speed in the identification process is certainly a first step toward the major goal of advancing a better comprehension of this poorly known component of Neotropical fauna and its associated evolutionary and ecological patterns. These actions are particularly urgent in face of the yet instituted climatic changes (Intergovernmental Panel on Climate Change, 2007) and biodiversity crisis (Butchart *et al*., 2010), since several Drosophilidae species and groups have been pointed as extremely sensitive to climatic conditions (Kellerman *et al*., 2012a, b; Robe *et al*., 2014).

Acknowledgments

We are grateful to the authorities responsible for Brazilian Ecological Stations, Reserves or Parks and to the private properties owners that allowed us to perform Drosophilidae searches in their area. All these collections were authorized by the Brazilian Ministério do Meio Ambiente (MMA), in the form of the Sistema de Autorização e Informação em Biodiversidade (SISBIO), and we thank them for their commitment to ethical guidelines and conservation efforts concerning Brazilian scientific activities. This study was supported by Universal-CNPq 14/2013, process numbers 471174/2013-0 and 472973/2013-4.

References

Amos W, Harwood J. 1998. Factors affecting levels of genetic diversity in natural populations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 353: 177-186.

Bächli G. 2014. TaxoDros: The database on taxonomy of Drosophilidae. Version 1.04. Available from: http://www.taxodros.uzh.ch/. Accessed: 16/10/2014.

Bromham L. 2009. Why do species vary in their rate of molecular evolution? *Biology Letters* 5: 401-404.

Brown SDJ, Collins RA, Boyer S, Lefort MC, Malumbres-Olarte J, Vink CJ, Cruickshank AH. 2012. Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. *Molecular Ecology Resources* 12(3): 562-565.

Burla H. 1956. Die Drosophilidengattung Zygothrica und ihre Beziehung zur *Drosophila*-Untergattung *Hirtodrosophila. Mitteilungen aus dem Museum für Naturkunde in Berlin* 32(2): 189–321.

Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW, Almond REA, Baillie JEM, Bomhard B, Brown C, Bruno J, Carpenter KE, Carr GM, Chanson J, Chenery AM, Csirke J, Davidson NC, Dentener F, Foster M, Galli A, Galloway JN, Genovesi P, Gregory RD, Hockings M, Kapos V, Lamarque J-F, Leverington F, Loh J, McGeoch MA, McRae L, Minasyan A, Morcillo MH, Oldfield TEE, Pauly D, Quader S, Revenga C, Sauer JR, Skolnik B, Spear D, Stanwell-Smith D, Stuart SN, Symes A, Tierney M, Tyrrel TD, Vié J-C, Watson R. 2010. Global biodiversity: indicators of recent declines. *Science* 328:1164–1168.

Casiraghi M, Labra M, Ferri E, Galimberti A, Mattia F. 2010. DNA barcoding: a six question tour to improve users' awareness about the method. *Briefings in Bioinformatics* 40(2): 440-453.

Clare EL, Lim BK, Fenton MB, Hebert PDN. 2011. Neotropical Bats: estimating species diversity with DNA Barcodes. *PLoS ONE* 6: e22648.

Costa FO, deWaard JR, Boutillier J, Ratnasingham S, Dooh RT, Hajibabaei M. 2007. Biological identification through DNA barcodes: the case of the Crustacea. *Canadian Journal of Fisheries and Aquatic Sciences* 64: 272-295.

Courtney SP, Kibota TT, Singleton TA. 1990. Ecology of mushroom-feeding Drosophilidae. *Advances in Ecological Research* 20: 225–274.

Crawford AJ, Cruz C, Griffth D, Ross H, Ibáñez R, Lips KR, Driskell AC, Bermingham E, Crump P. 2012. DNA barcoding applied to *ex situ* tropical amphibian conservation programme reveals cryptic diversity in captive populations. *Molecular Ecology Resources* 13(6): 1005-18.

DaLage JL, Dergoat GJ, Maczkowiak F, Silvan J-F, Cariou M-L, Lachaise D. 2007. A phylogeny of Drosophilidae using the Amyrel gene: questioning the *Drosophila melanogaster* species group boundaries*. Journal of Zoological Systematics and Evolutionary Research* 45(1): 47–63.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9(8): 772.

DeSalle R, Egan MG, Siddall M. 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding*. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 360: 1905-1916*.*

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.

Ebach MC, Holdrege C. 2005. DNA barcode is no substitute for taxonomy. *Nature* 434: 697.

Esselstyn JA, Evans BJ, Sedlock JL, Khan FAA, Haney LR. 2012. Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proceedings of the Royal Society of London. Series B* 279: 3678-3686.

Floyd R, Abebe E, Papert A, Blaxter M. 2002. Molecular barcodes for soil nematode identification. *Molecular Ecology* 11: 839–850.

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(5): 294-299.

Fontaneto D, Herniou EA, Boschetti C, Capiroli M, Melone G, Ricci C, Barraclough TG. 2007. Independently Evolving Species in Asexual Bdelloid Rotifers. *PLoS Biology* 5(4): e87.

Frota-Pessoa O. 1945. Sobre o subgênero *Hirtodrosophila*, com descrição de uma nova espécie (Diptera, Drosophilidae, *Drosophila*). *Revista Brasileira de Entomologia* 5(4): 469-483.

Fujisawa T, Barraclough TG. 2013. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent Approach: a revised method and evaluation on simulated data sets. *Systematic Biology* 62(5): 707-724.

Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* 34: 397-423.

Galtier N, Nabholz B, Glémin S, Hurst GDD. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Molecular Ecology 18: 4541-4550.

Garcia CF, Hochmüller CJC, Valente VLS, Schmitz HJ. 2012. Drosophilid assemblages at different urbanization levels in the city of Porto Alegre, state of Rio Grande do Sul, Southern Brazil. *Neotropical Entomology* 40(1): 32-41.

Giam X, Scheffers BR, Sodhi NS, Wilcove DS, Ceballos G, Ehrlich PR. 2012. Reservoirs of richness: least disturbed tropical forests are centres of undescribed species diversity. *Proceedings of the Royal Society of London. Series B* 279: 67–76.

Gill BA, Harrington RA, Kondratieff BC, Zamudio KR, Poff NLR, Funk C. 2014. Morphological taxonomy, DNA barcoding, and species diversity in southern Rocky Mountain headwater streams. *Freshwater Science* 33(1): 288-301.

Goldstein PZ, DeSalle R. 2010. Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. *Bioessays* 33: 135-147.

Gottschalk MS, de Toni DC, Valente VLS, Hofmann PRP. 2007. Changes in Brazilian Drosophilidae (Diptera) assemblages across an urbanization gradient. *Neotropical Entomology* 36 (6): 848-862.

Gottschalk MS, Hofmann PRP, Valente VLS. 2008. Diptera, Drosophilidae: historical occurrence in Brazil. *Check List* 4: 485-518.

Gottschalk MS, Bizzo L, Döge JS, Profes MS, Hofmann PRP, Valente VLS. 2009. Drosophilidae (Diptera) associated to fungi: differential use of resources in anthropic and Atlantic Rain Forest areas. *Iheringia Série Zoologia* 99: 442-448.

Grimaldi DA. 1987. Phylogenetics and taxonomy of *Zygothrica* (Diptera: Drosophilidae). *Bulletin of the American Museum of Natural History* 186: 103–268.

Grimaldi DA. 1990a. A phylogenetic revised classification of genera in the Drosophilidae (Diptera). *Bulletin of the American Museum of Natural History* 197: 1–139.

Grimaldi DA. 1990b. Revision of *Zygothrica* (Diptera: Drosophilidae), Part II. The First African Species, Two New Indo-Pacific Groups, and the *bilineata* and *samoaensis* Species Groups. *American Museum Novitates* 2964: 1–31.

Grimaldi DA. 2010. Drosophilidae (Small fruit flies, pomace flies, vinegar flies). In: Brown BV, Borkent A, Cumming JM, Wood DM, Woodley NE, Zumbado MA, eds. Manual of Central American Diptera. Ottawa: NRC Research Press, 1197-1206.

Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003a. Biological identification through DNA barcodes. *Proceedings of the Royal Society of London. Series B* 270: 313-321.

Hebert PDN, Ratnasingham S, deWaard JR. 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B* 7: 270(Suppl 1) S96–S99.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator. Proceedings of the National Academy of Sciences of the United States of America* 101(41): 14812-14817.

Hebert PDN, Gregory TR. 2005. The promise of DNA Barcoding for taxonomy. *Systematic Biology* 54(5): 852-859.

Hurst GDD, Jiggins FM. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society of London. Series B* 272: 1525-1534.

Intergovernmental Panel on Climate Change. 2007. Climate change 2007: synthesis report. Geneva (Switzerland): Intergovernmental Panel on Climate Change.

Jinbo U, Kato T, Ito M. 2011. Current progress in DNA barcoding and future implications for entomology. *Entomological Science* 14: 107-124.

Kaneshiro K. 1969. A study of the relationships of Hawaiian *Drosophila* species based on the external male genitalia. *The University of Texas Publication* 6918(4016): 55-69.

Kasuya N, Mitsui H, Aotsuka T, Kimura MT. 2013. Diversity and host associations of parasitoids attacking mycophagous drosophilids (Diptera: Drosophilidae) in northern and central Japan. *Entomological Science* 6: 227-234.

Kekkonen M, Hebert PD. 2014. DNA barcode-based delineation of putative species: eficiente start for taxonomic workflows. *Molecular Ecology Resources* 14: 706-715.

Kellermann V, Loeschcke V, Hoffmann AA, Kristensen TN, Fløjgaard C, David JR, Svenning J-C, Overgaard J. 2012a. Phylogenetic constraints in key functional traits behind species' climate niches: patterns of desiccation and cold resistance across 95 *Drosophila* species. *Evolution* 66(11): 3377–3389.

Kellermann V, Overgaard J, Hoffmann AA, Fløjgaard C, Svenning J-C, Loeschecke V. 2012b. Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically. *Proceedings of the National Academy of Sciences of the United States of America* 109(40): 16228–16233.

Kelly RP, Sarkar IN, Eernisse DJ, DeSalle R. 2007. DNA barcoding using chitons (genus *Mopalia*). *Molecular Ecology Notes* 7: 177–183.

Kerr KCR, Birks MS, Kalyakin MV, Red'kin YA, Koblik EA, Hebert PDN. 2009. Filling the gap - COI barcode resolution in eastern Palearctic birds. *Frontiers in Zoology* 6: 29.

Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16: 111-120.

Kvist S. 2013. Barcoding in the dark?: A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge. *Molecular Phylogenetics and Evolution* 69: 39–45.

Landim MIPF, Hingst-Zaher E. 2010. Brazil's biodiversity crisis. *ICOM News* 63: 4-15.

Lefébure T, Douady CJ, Gouy M, Gilbert J. 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: Proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution* 40(2): 435-447.

Leite LAR. 2012. Mitochondrial pseudogenes in insect DNA barcoding: differing points of view on the same issue. *Biota Neotropica* 12(3): 1-8.

Lewinsohn TM, Prado PI. 2005. How Many Species Are There in Brazil? *Conservation Biology* 19: 619–624.

Linares MC, Soto-Calderon ID, Lees DC, Anthony NM. 2009. High mitochondrial diversity in geographically widespread butterflies of Madagascar: a test of the DNA barcoding approach. *Molecular Phylogenetics and Evolution* 50: 485-495.

Machado S, dos 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.

Markow TA, O'Grady P. 2008. Reproductive ecology of *Drosophila*. *Functional Ecology* 22: 747-759.

Medeiros HF, Klaczko L. 2004. How many species of drosophila (Diptera, Drosophilidae) remain to be described in the forests of São Paulo, Brazil?: species lists of three forest remnants. *Biota Neotropica* 4(1): 01-12.

Meier R, Shiyang K, Vaidya G, Ng PKL. 2006. DNA Barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic Biology* 55(5): 715-728.

Meyer CP, Paulay G. 2005. DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology* 3(12): e422.

Miralles A, Vences M. 2013. New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in Madascincus lizards. *PLoS ONE* 8: e68242.

Mittermeier RA. 1988. Primate diversity and the tropical forest case studies from Brazil and Madagascar and the Importance of the megadiversity countries. In:

Wilson EO, Peter FM eds. *Biodiversity*. Washington: National Academies Press, 145-157.

Mittermeier RA, Robles-Gil P, Mittermeier CG. 1997. *Megadiversity: Earth's biologically wealthiest nations*. Mexico: CEMEX.

Moritz C, Cicero C. 2004. DNA Barcoding: promise and pitfalls. *PLoS Biology* 2(10): e354.

Mutanen M, Kaila K, Tabell J. 2013. Wide-ranging barcoding aids discovery of one-third increase of species richness in presumably well-investigated moths. *Scientific Reports* 3(2901): 1-7.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853-858.

Nagy ZT, Sonet G, Glaw F, Vences M. 2012. First large-scale DNA Barcoding assessment of Reptiles in the biodiversity hotspot of Madagascar, based on newly designed COI Primers. *Plos ONE* 7(3): e34506.

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* 3: 418-426.

Nicholls JA, Challis RJ, Mutun S, Stone GN. 2012. Mitochondrial barcodes are diagnostic of shared refugia but not species in hybridizing oak gallwasps. *Molecular Ecology* 21: 4051–4062.

O'Grady PM, Markow TA. 2009. Phylogenetic taxonomy in *Drosophila*: problems and prospects. *Fly* 3(1): 10-14.

Paz A, Crawford AJ. 2012. Molecular-based rapid inventories of sympatric diversity: a comparison of DNA barcode clustering methods applied to geography-based vs clade-based sampling of amphibians. *Journal of Biosciences* 35(5): 887-896.

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazel S, Kamoun S, Sumlin WD, Vogler AP. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* 55: 595-609.

Poppe JL, Schmitz H, Grimaldi D, Valente VLS. 2014. High diversity of Drosophilidae (Insecta, Diptera) in the Pampas Biome of South America, with descriptions of new *Rhinoleucophenga* species. *Zootaxa* 3779(2): 215–245.

Porter TM, Gibson JF, Shokralla S, Baird DJ, Golding GB, Hajibabaei M. 2014. Rapid and accurate taxonomic classification of insect (class Insecta) cytochrome *c* oxidase subunit 1 (COI) DNA barcode sequences using a naïve Bayesian classifier. *Molecular Ecology Resources* 14(5): 929-942.

Pramual P, Adler PH. 2014. DNA barcoding of tropical black flies (Diptera: Simuliidae) of Thailand. *Molecular Ecology Resources* 14: 262–271.

Puillandre N, Lambert A, Brouillet S, Achaz G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* 21: 1864-1877.

Ratnasingham S, Hebert PDN. 2007. BOLD: the Barcode of Life Data system (www.barcodinglife.org). *Molecular Ecology Notes* 7: 355–364.

Robe LJ, Machado S, Bartholomei-Santos ML. 2012. The DNA Barcoding and the caveats with respect to its application to some species of Palaemonidae (Crustacea, Decapoda). *Zoological Science* 29: 714–724.

Robe LJ, Machado S, Bolzan AR, dos Santos JPJ, Valer F, dos Santos AP, Blauth ML, Gottschalk MS. 2014. Comparative ecological niche modeling and evolutionary ecology of Neotropical mycophagous Drosophilidae (Diptera) species. *Studies on Neotropical Fauna and Environment* 49(2): 79-94.

Roque F, Tidon R. 2008. Eight new records of drosophilids (Insecta: Diptera) in the Brazilian savanna. *Drosophila Information Service* 91: 94- 98.

Ross HA. 2014. The incidence of species-level paraphyly in animals: a re-assessment. *Molecular Phylogenetics and Evolution* 76: 10-17.

Rouquette J, Davis AJ. 2003. *Drosophila* species (Diptera: Drosophilidae) oviposition patterns on fungi: the effect of allospecifics, substrate toughness, ovipositor structure and degree of specialization. *European Journal of Entomology* 100: 351-355.

Rubinoff D. 2006. Utility of Mitochondrial DNA Barcodes in species conservation. *Conservation Biology Series* 20: 1026-1033.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annual Review of Entomology* 87: 651-701.

Song H, Buhay JE, Whiting MF, Crandall KA. 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences of the United States of America* 105(36): 13486-13491.

Srivathsan A, Meier R. 2012. On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. *Cladistics* 28: 190–194.

Staden R. 1996. The Staden sequence analysis package. *Molecular Biotechnology* 5(3): 233-41.

Stump AD, Jablonski SE, Bouton L, Wilder JA. 2011. Distribution and mechanism of α-amanitin tolerance in mycophagous *Drosophila* (Diptera: Drosophilidae). *Environmental Entomology* 40(6): 1604-1612.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* 30: 2725-2729.

Tavares ES, Gonçalves P, Miyaki CY, Baker AJ. 2011. DNA barcode detects high genetic structure within Neotropical bird species. *PLoS ONE* 6: e28543.

Talavera G, Dinca V, Vila R. 2013. Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution* 4: 1101–1110.

Thorp JH, Dimmick WW, Covich AP. 2010. Introduction to invertebrates of Island Waters. In: Thorp JH, Dimmick WW, Covich AP, eds. *Ecology and classification of North American freshwater invertebrates*. Elsevier, 1-23.

Throckmorton LH. 1975. The phylogeny, ecology and geography of Drosophila. In: King RC, ed. *Handbook of genetics*. New York: Plenum, 421–469.

Tosi D, Martins M, Vilela CR, Pereira MAQR. 1990. On a new cave-dwelling species of bat-guano-breeding *Drosophila* closely related to *D. repleta* Wollaston (Dipetra, Drosophilidae). *Revista Brasileira Genética* 13(1): 19-31.

Trewick SA. 2008. DNA Barcoding is not enough: mismatch of taxonomy and genealogy in New Zealand grasshoppers (Orthoptera: Acrididae). *Cladistics* 24(2): 240-254.

Tuno N, Takahashi KH, Yamashita H, Osawa N, Tanaka C. 2007. Tolerance of *Drosophila* flies to Ibotenic Acid poisons in mushrooms. *Journal of Chemical Ecology* 33: 311-317.

Val FC, Marques MD, Vilela CR. 1981. Drosophilidae of Neotropical region. In: Ashburner M, Carson HL, Thompson JN, eds. *The genetics and biology of Drosophila*. Orlando: Academic Press, 123-168.

Valadão H, Hay JDV, Tidon R. 2010. Temporal dynamics and resource availability for drosophilid fruit flies (Insecta, Diptera) in a gallery forest in the Brazilian Savanna. *International Journal of Ecology* 2010: 1-8.

van der Linde K, Houle D, Spicer GS, Steppan SJ. 2010. A supermatrix-based molecular phylogeny of the family Drosophilidae. *Genetics Research* 92(1): 25–38.

Vilela CR, Bächli G. 1990. Taxonomic studies on Neotropical species of seven genera of Drosophilidae (Diptera). *Mitteilungen der Schweizerischen entomologischen Gesellschaft* 63: 1-332.

Vilela CR. 1992. On the *Drosophila tripunctata* species group (Diptera, Drosophilidae). *Revista Brasileira de Entomologia* 36: 197-221.

Vilela CR, Bächli G. 2004. On the identities of nine Neotropical species of *Hirtodrosophila* (Diptera, Drosophilidae). *Mitteilungen der Schweizerischen entomologischen Gesellschaft* 77: 161–195.

Vilela CR, Bächli G. 2005. Three new species of Neotropical Drosophilidae (Diptera). *Mitteilungen der Schweizerischen entomologischen Gesellschaft* 78: 143-163.

Vilela CR, Bächli G. 2007. Revision of the Neotropical genus *Paraliodrosophila* (Diptera, Drosophilidae). *Mitteilungen der Schweizerischen entomologischen Gesellschaft* 80: 291–317.

Wheeler MR, Kambysellis MP. 1966. Notes on the Drosophilidae (Diptera) of Samoa. *The University of Texas Publications* 6615: 533-565.

Wheeler MR, Takada H. 1963. A revision of the American species of *Mycodrosophila* (Diptera, Drosophilidae). *Annals of the Entomological Society of America* 56(3): 392-399.

Wiemers M, Fiedler K. 2007. Does the DNA barcoding gap exist? –a case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology* 4: 8.

Will KW, Mishler BD, Wheeler QD. 2005. The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology* 54: 844-851.

Wilson JJ. 2011. Assessing the value of DNA Barcodes for molecular phylogenetics: effect of increased taxon sampling in Lepidoptera. PLoS ONE 6(9): e24769.

Yassin A, Markow TA, Narechania A, O'Gardy PM, DeSalle R. 2010. The genus *Drosophila* as a model for testing tree- and character-based methods of species identification using DNA barcoding. *Molecular Phylogenetics and Evolution* 57: 509–517.

Yassin A. 2013. Phylogenetic classification of the Drosophilidae Rondani (Diptera): the role of morphology in the postgenomic era. *Systematic Entomology* 38: 349-364.

Yorozuya H. 2006. Effects of parasitoids on a mycophagous drosophilid community in northern Japan and an evaluation of the disproportionate parasitism hypothesis. *Entomological Science* 9: 13-22.

Yorozuya H. 2009. Effects of nematode parasitism on mycophagous drosophilid species in northern Japan. *Entomological Science* 12: 363-369. Zahiri R, Lafontaine JD, Schmidt BC, Dewaard JR, Zakharov EV, Hebert PD. 2014. A Transcontinental Challenge — A test of DNA Barcode performance for 1,541 species of Canadian Noctuoidea (Lepidoptera). *PLoS ONE* 9(3): e92797.

Tables and figures

Table 1. Summary information and essential statistics for each of the sampled species. Described species are presented in bold, whereas additional morphotypes/MOTUs are subdivided in: *those presenting morphological distinction at least at the genitalia level; ** those that do not present any morphological distinction in regard to at least one of its cryptic species. N= number of specimens, MI% Minimum intraspecific distance (in %), MX% Maximum intraspecific distance (in %), MD% (±) Mean intraspecific distance (in %) and standard error (in %), ID% Mean interspecific distance to cryptic species (in %), dN/dS Mean dN/dS values presented in regard to cryptic species. Species for which no difference was found in aedeagus structure between cryptic species have their respective ID% presented in bold.

	Species	Cryptic species	N	MI%	MX%	MD% (±)	ID%	dN/dS
1	***H. gilva***	with 3, 8, 16, 17	6	0	3.5	1.1 (0.2)	11.1, 14.85, 15.8, 9.1	0.02, 0.03, 0.01, 0.009
2	***H. levigata***	with 7	18	0	1.4	0.6 (0.2)	12.86	0.02
3	***H. mendeli***	with 1, 8, 16, 17	9	0	0.9	0.5 (0.2)	11.1, 12.4, 13.8, 10.2	0.02, 0.02, 0.03, 0.03
4	***H. morgani***	with 5, 9, 16, 17	7	0	3.4	1.9 (0.4)	13.9, 12.7, 14, 12.7	0.02, 0.01, 0.02, 0.01
5	***H. subgilva***	with 4, 9, 16, 17	3	0.7	5.1	3.4 (0.6)	13.9, 12.9, 13.7, 11.9	0.02, 0.02, 0.02, 0.02
6	***H. trilineata*** [#]	-	-	-	-	-	-	-
7	*H. levigata* affinis**	with 2	1	-	-	-	12.86	0.02
8	*H. mendeli* affinis*	with 1, 3, 16, 17	2	-	-	0 (0)	14.85, 12.4, 13.8, 10.2	0.03, 0.02, 0.03, 0.03
9	*H. morgani* affinis sul*	with 4, 5, 16, 17	6	0	0	0 (0)	12.7, 12.9, 7.7, 12.9	0.01, 0.02, 0, 0.008
10	*H. subflavohalterata* affinis 1*	with 11, 12, 13, 14, 15	3	0	0	0 (0)	12.6, 9.5, 9.2, 12.9, 12.3	0.01, 0.008, 0.01, 0.02, 0.01
11	*H. subflavohalterata* affinis 2*	with 10, 12, 13, 14, 15	4	0.3	1	0.6 (0.2)	12.6, 12.5, 14.6, 5.5, 2.8	0.01, 0.02, 0.01, 0.009, 0.01
12	*H. subflavohalterata* affinis 3*	with 10, 11, 13, 14, 15	2	-	-	0 (0)	9.5, 12.5, 11.4, 14.2, 15	0.008, 0.02, 0.03, 0.03, 0.02
13	*H. subflavohalterata* affinis 4**	with 10, 11, 12,	1	-	-	-	9.2, 14.6, 11.4, 14.7,	0.01, 0.01, 0.03, 0.02, 0.02

		14, 15					13.7	
14	*H. subflavohalterata* affinis 5**	with 10, 11, 12, 13, 15	9	0	3.2	1.2 (0.3)	12.9, 5.5, 14.2, 14.7, 5.7	0.02, 0.009, 0.03, 0.02, 0.009
15	*H. subflavohalterata* affinis H002*	with 10, 11, 12, 13, 14	9	0	1.3	0.4 (0.1)	12.3, 2.8, 15, 13.7, 5.7	0.01, 0.01, 0.02, 0.02, 0.009
16	*H. subgilva* affinis 1*	with 1, 3, 4, 5, 8, 9, 17	1	-	-	-	15.8, 13.8, 14, 13.7, 13.8, 7.7, 13.4	0.01, 0.03, 0.02, 0.02, 0.03, 0, 0.008
17	*H. subgilva* affinis 2*	with 1, 3, 4, 5, 8, 9, 16	3	0	0	0 (0)	9.1, 10.2, 12.7, 11.9, 10.2, 12.9, 13.4	0.009, 0.03, 0.01, 0.02, 0.03, 0.008, 0.008
18	*H.* H007*	-	1	-	-	-	-	-
19	*H.* H010*	-	1	-	-	-	-	-
20	*H.* H012*	-	1	-	-	-	-	-
21	*H.* H013*	-	1	-	-	-	-	-
22	***M. dimidiata*#**	-	-	-	-	-	-	-
23	***M. elegans***	-	3	0	1.5	1 (0.3)	-	-
24	***M. projectans***	with 29, 30	10	0.3	3.3	1.7 (0.3)	10, 10.9	0.01, 0.009
25	M. M001*	with 26	5	0	0.5	0.2 (0.1)	14	0.02
26	M. affinis M001**	with 25	1	-	-	-	**14**	**0.02**
27	*M. neoprojectans* affinis 1**	with 28	3	0.1	0.8	0.6 (0.2)	**5.5**	**0.01**
28	*M. neoprojectans* affinis 2**	with 27	4	0.2	0.8	0.4 (0.2)	**5.5**	**0.01**
29	*M. projectans* affinis 1**	with 24, 30	2	-	-	0.1 (0.1)	**10**, **12.2**	**0.01, 0.01**
30	*M. projectans* affinis 2**	with 24, 29	4	0.7	1.7	1.2 (0.3)	**10.9**, **12.2**	**0.009, 0.01**
31	***Pa. antennata***	-	10	0	1.5	0.3 (0.1)	-	-
32	***Pa. burlai***	-	2	-	-	0.7 (0.3)	-	-
33	***Zy. atriangula***	with 45	2	-	-	0.2 (0.2)	7.5	0.01

34	***Zy. dispar***	-	6	0.3	1.6	1.0 (0.2)	-	-
35	***Zy. hypandriata***	-	7	0	0.4	0.2 (0.1)	-	-
36	***Zy. orbitalis***	-	13	0	1.9	0.8 (0.2)	-	-
37	***Zy. parapoeyi***	with 51	2	-	-	0.3 (0.2)	7.3	0
38	***Zy. parvipoeyi***	-	3	0	0.3	0.2 (0.1)	-	-
39	***Zy. prodispar***	with 53	2	-	-	0.2 (0.2)	9.1	0
40	***Zy. ptilialis***	-	11	0	5.3	1.7 (0.3)	-	-
41	***Zy. virgatinigra***	-	1	-	-	-	-	-
42	***Zy. vittimaculosa***	with 55	1	-	-	-	3.7	0
43	***Zy. vittinubila***	-	1	-	-	-	-	-
44	*Zy. apopoeyi* affinis*	-	8	0	0.4	0.2 (0.1)	-	-
45	*Zy. atriangula* affinis H015*	with 33	1	-	-	-	7.5	0.01
46	*Zy. festiva* affinis*	-	1	-	-	-	-	-
47	*Zy. fuscina* affinis*	-	2	-	-	0.9 (0.4)	-	-
48	*Zy. gracilipoeyi* affinis*	-	1	-	-	-	-	-
49	*Zy. laevifrons* 1**	with 50, 52	5	0	0.1	0.1 (0.1)	**7.8**, 11.8	**0.005**, 0.02
50	*Zy. laevifrons* 2**	with 49, 52	3	0.4	0.7	0.6 (0.2)	**7.8**, 12	**0.005**, 0.02
51	*Zy. parapoeyi* affinis*	with 37	4	0	0.3	0.2 (0.1)	7.3	0
52	*Zy. poeyi* affinis*	with 49, 50	1	-	-	-	11.8, 12	0.02, 0.02
53	*Zy. prodispar* affinis 1**	with 39	4	0.4	1.3	0.9 (0.3)	**9.1**	**0**
54	*Zy. prodispar* affinis 2*	-	2	-	-	0.6	-	-

						(0.3)		
55	*Zy. vittimaculosa* affinis*	with 42	3	0	1.3	0.7 (0.2)	3.7	0
56	*Zy.* H009*	-	4	0	0	0 (0)	-	-

Table 2. Methodological congruence in regard to the suggestion of the 24 putative new species represented by males.

Putative new species	Morphological Difference	Minimum interspecific divergence higher than the best threshold	Diagnostic character	GYMC recovery*
H. mendeli affinis	√	√	√	√√
H. morgani affinis sul	√	√	√	√
H. subflavohalterata	√	√	**X**	√
H. subflavohalterata	√	√	**X**	√
H. subflavohalterata	√	√	√	√
H. subflavohalterata	**X**	√	√	√√
H. subflavohalterata	√	√	**X**	√√
H. subgilva affinis 1	√	√	√	√
H. subgilva affinis 2	√	√	**X**	√
M. M001	√	√	**X**	√
M. affinis M001	**X**	√	√	√
M. neoprojectans	**X**	√	√	√
M. neoprojectans	**X**	√	√	√
M. projectans affinis 2	**X**	√	**X**	√
Zy. apopoeyi affinis	√	√	**X**	√
Zy. festiva affinis	√	√	**X**	√
Zy. gracilipoeyi affinis	√	√	√	√
Zy. laevifrons 1	**X**	√	√	√
Zy. laevifrons 2	**X**	√	√	√
Zy. parapoeyi affinis	√	√	√	√
Zy. prodispar affinis 1	**X**	√	**X**	√
Zy. prodispar affinis 2	√	√	√	√
Zy. vittimaculosa affinis	√	√	√	√
Zy. H009	√	√	√	√

√√ indicates that additional subdivision was suggested by GMYC.

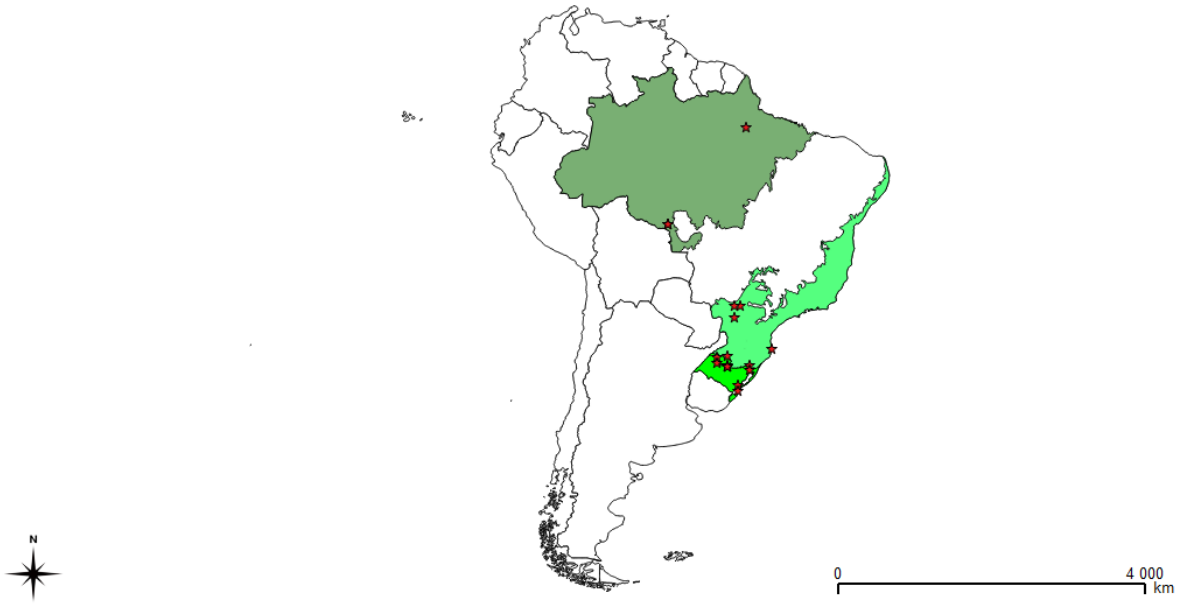

Figure 1. Map of sampling sites of the *Zygothrica* genus group mycophylic species across Brazil. All Brazilian biomes are delimited, and those in which collections were performed are highlighted. Stars represent sampling sites.

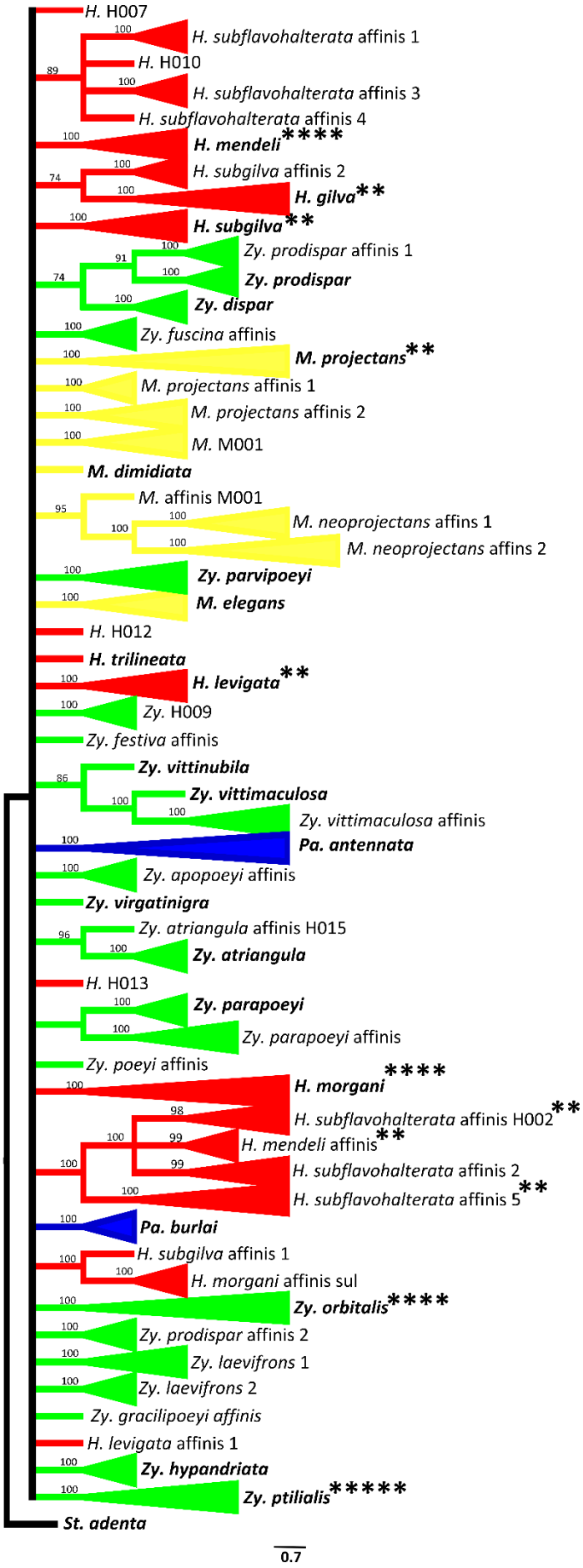

Figure 2. K2P Neighbor-joining tree generated for the *Zygothrica* genus group mycophylic species COI dataset. Branches were collapsed to represent each of the 56 independent taxonomic entities and colored according to the corresponding genus in order to favor visualization. Numbers before each collapsed clade correspond to bootstrap values and branch lengths are proportional to the scale, given in substitutions per nucleotide. The 22 described species are represented in bold whereas the other 34 collapsed clades represent undescribed or dubious cases. Asterisks represent the number of independent clusters suggested for each clade by GMYC (when the cluster suggested by GMYC totally agrees with the secondary NJ tree, the asterisk was omitted; for details see appendix – Table S7).

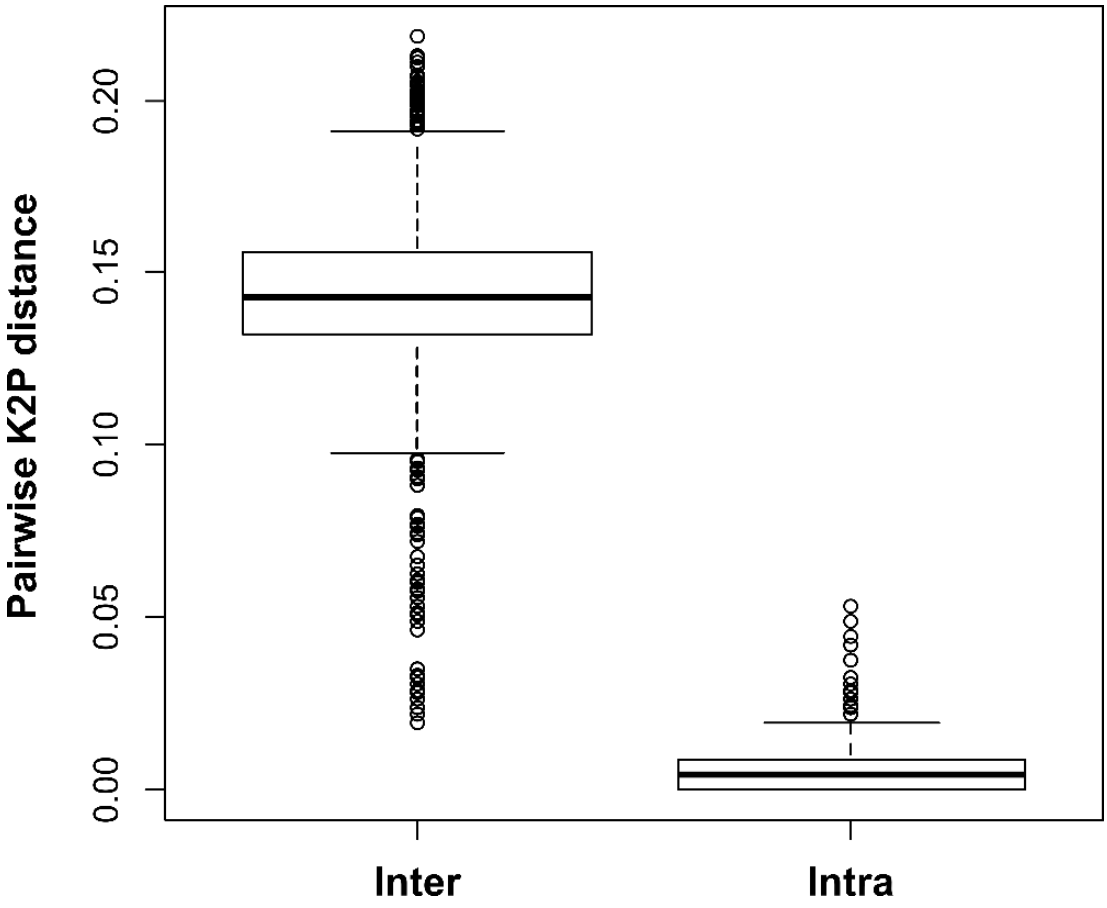

Figure 3. Comparison between intra and interspecific K2P pairwise distances obtained for the sampled *Zygothrica* genus group specimens.

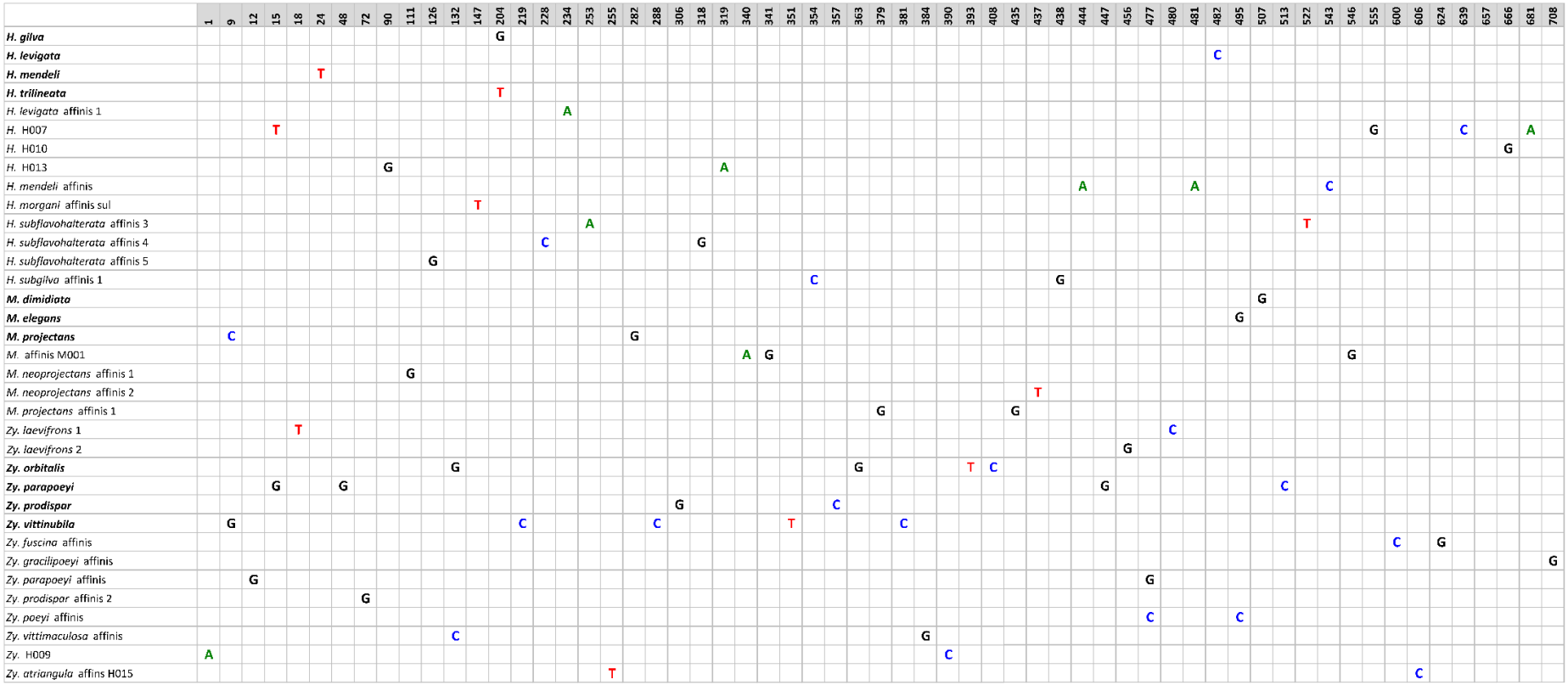

Figure 4. Diagnostic characters presented by each *Zygothrica* genus group species for which character-based analysis was successful. Sites and species that did not present diagnostic nucleotides were omitted.

Supporting information

Figure S1. Interactive map of sampling sites per species [to be opened in Quantum GIS (http://www.qgis.org/en/site/)]

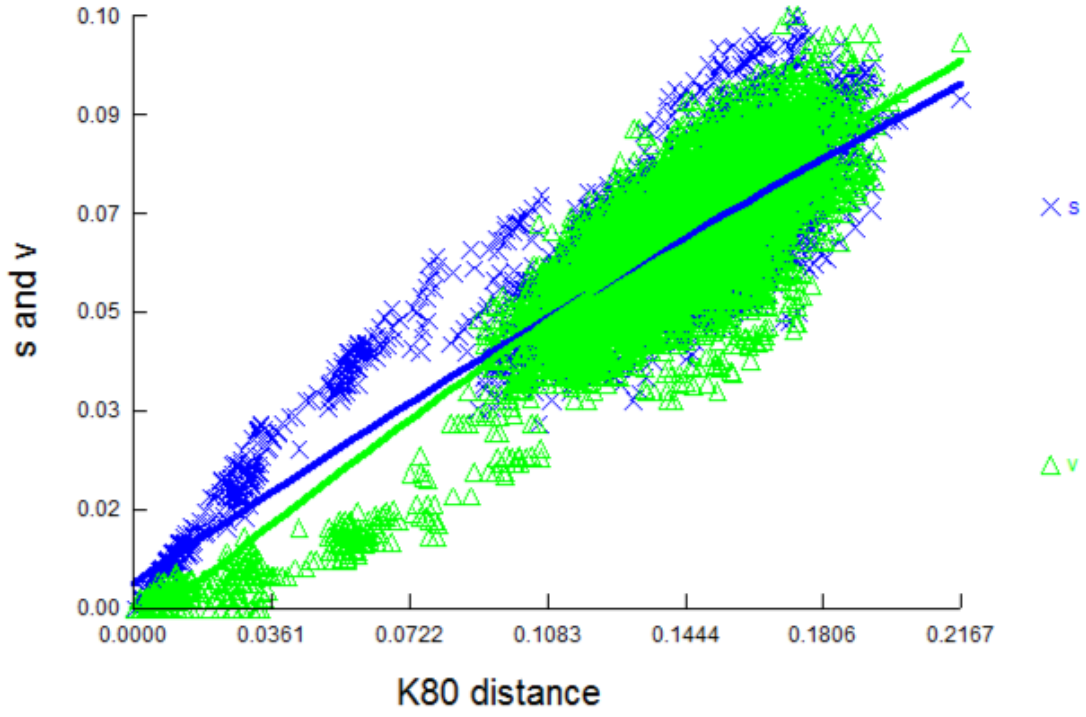

Figure S2. Saturation test of mycophylic drosophilids COI sequences using K80 (K2P) correction model.

3.4. Capítulo 4

Título: **Disentangling the Brazilian *Mycodrosophila projectans* (Diptera, Drosophilidae) species complex**

Artigo a ser submetido:

Biological Journal of the Linnean Society

ISSN:

1095-8312

Disentangling the Brazilian *Mycodrosophila projectans* (Diptera, Drosophilidae) species complex

Stela Machado[1] stelamchd@gmail.com
Elgion Lucio da Silva Loreto[1] elgionl@gmail.com
Marco Silva Gottschalk[2] gotts007@yahoo.com
Lizandra Jaqueline Robe[1,3] lizbiogen@gmail.com

[1]Universidade Federal de Santa Maria (UFSM), Programa de Pós Graduação em Biodiversidade Animal, Avenida Roraima 1000, CEP 97105-900 - Santa Maria, RS - Brazil, Phone: +55 55 3220-8912
[2]Universidade Federal de Pelotas (UFPel), Instituto de Biologia, Departamento de Ecologia, Zoologia e Genética, Campus Capão do Leão, Mail Box 354, CEP 96010-900 - Capão do Leão, RS – Brazil, Phone: +55 53 3275-7340
[3]Universidade Federal do Rio Grande (FURG), Instituto de Ciências Biológicas – ICB, Av. Itália, km 08 Campus Carreiros, CEP 96203-900 - Rio Grande, RS – Brasil, Phone: +55 53 3293-5128

Corresponding Author:
Lizandra Jaqueline Robe

Abstract

The *Zygothrica* genus group has been shown to be a diversified group, with high number of cryptic species. Here we analyze the diversity, ecology and evolution of one of the genus of this group, *Mycodrosophila*, in Brazilian biomes with special attention to the triad of cryptic species belonging to the *M. projectans* complex. We found a high diversity in both COI and COII molecular markers for most of the studied *Mycodrosophila* species, especially when this is compared with other Drosophilidae species. Ecologically, sympatry and syntopy seems to be a recurrent pattern for the putative sister cryptic species of the genus, and in the *M. projectans* complex, it was observed in fungi of different genera. Although the species of the complex showed different amplitudes of potential distribution, as assigned by ecological niche modeling, niche overlap measures and identity tests corroborate the niche conservatism in evolution of these species. This pattern is in agreement with the lack of complete morphological differentiation between them, despite the fact that their speciation was dated to the Miocene.

Keywords: phylogeography, Quaternary, Pleistocene, mycophagous, ecological niche modeling

Introduction

Drosophilidae is a speciose family encompassing more than 4.200 species and 77 genera (Bächli, 2015), which is almost world widely distributed (Throckmorton, 1975). Flies of this family are associated with diverse resources, like fruits, fungus, sap, pollen and rotten leaves (Carson, 1971; Markow & O'Grady, 2008). This diversity in resource exploration occurs due to the saprophagous nature of drosophilds, which feed mainly on bacteria and yeast involved in the fermentation of each of these resources, or on their components thereby made available (Throckmorton, 1975; Powell, 1997; Markow & O'Grady, 2008).

Nevertheless, many taxa of the family associate with macroscopic fungi found in the form of fruiting bodies, using them as resource for feeding, [eating spores and hyphae (Grimaldi, 1987)], oviposition and/or breeding sites. These taxa are commonly known as mycophagous (Courtney *et al.*, 1990) or mycophylic species (Machado *et al.*, 2014). Although this property seems to be homoplastic within Drosophilidae (Throckmorton, 1975), it may represent a synapomorphy for the *Zygothrica* genus group (Grimaldi, 1990), which encompasses five genera with different degrees of association with fungus: *Hirtodrosophila*, *Mycodrosophila*, *Paraliodrosophila*, *Paramycodrosophila* and *Zygothrica*.

Mycodrosophila is composed by about 130 obligatory mycophagous species (Courtney *et al.*, 1990), distributed all over the world (Bächli, 2015). Most part of this diversity is found in the Old World, with only 4% of the species being recorded in the Neotropical region (McEvey & Polak, 2005). Nevertheless, there are evidences that the Neotropical diversity of this genus is seriously underestimated, and a study on the effectiveness of DNA barcode on the group increased from three (Wheeler & Takada, 1963) to nine (Machado *et al.*, in preparation) the number of species encountered in Brazil. In fact, few studies have incorporated Neotropical *Mycodrosophila* (Wheeler & Takada, 1963; Grimaldi, 1987; Roque & Tidon, 2008; Gottschalk *et al.*, 2009; Robe *et al.*, 2014), and except by the revision of the American species performed by Wheeler & Takada (1963), no extensive works have been made with the genus in this region.

Mycodrosophila projectans has been considered a widely distributed mycophylic species, occurring from Mexico to Brazil (Robe *et al.*, 2014; Bächli, 2015). Nevertheless, there are evidences of significant genetic structure within this range

(Machado *et al.*, in preparation), so that this may be composed by a group of cryptic species. Moreover, cryptic diversity was also found in Brazil by the same authors within other new Brazilian *Mycodrosophila* “morphotypes”. As Machado *et al.* (in preparation) incorporated only a low number of individuals for each of the six *Mycodrosophila* putative new species sampled (n=2-5) the aim of the present study is to characterize genetically, morphologically and ecologically the Brazilian *M. projectans* species complex. In this sense, we increased the sampling for the mitochondrial COI (cytochrome oxidase c subunit I) gene, and also incorporated sequences of the mitochondrial COII (cytochrome oxidase c subunit II) and the nuclear hunchback genes in order to date the diversification of the lineages within this complex. Once male genitalia showed to be poorly effective to distinguish these cryptic species, wing morphometry was also tested as a marker useful to this purpose. Notwithstanding the diversification of the group was dated to the Miocene (approximately 16 Mya ago), none morphological marker proven to be effective in species differentiation.

Material and methods

Sampling

Collections were performed in 45 sites distributed along Amazonian, Atlantic and Pampa Brazilian Biomes (Table S1, Fig. 1a). In these places, active searches for fruiting bodies of macroscopic fungi were performed in forest fragments during the active period of the flies. Adult specimens resting or flying over the fungi were collected with the use of an entomological aspirator (Machado *et al.*, 2014) and, in some cases, the resource was carried to the laboratory and maintained until adults emergence. Flies were stored in absolute ethanol and determination to the species level was performed based on external morphology and male genitalia. The genitalia slides were prepared following Bächli *et al.* (2004) and identified according the descriptions of Wheeler & Takada (1963).

DNA and sequences manipulation

Total DNA of each specimen was extracted using the NucleoSpin Tissue XS kit (Marchery-Nagel). The different markers were amplified from a set of 103 *Mycodrosophila* specimens (Table S2) with the use of different combinations of the

following set of primers: TYJ1460 and C1N2329M (5’ ACT GTA AAT ATA TGA TGA GCT CAT ACA 3’) modified from Simon *et al.* (1994), HCO1490 and LCO2198 (Folmer *et al.*, 1994), COIMYCOF (5’ AYT TTA TTT TYG GRG CHT GR 3’) and COIMYCOS R (5’ WCC TAA TGA DCC AAA DGT TTC Y 3’) for the mitochondrial cytochrome c oxidase subunit I (COI) gene; TL2J3037 and TKN3785 (Simon *et al.*, 1994), and COII3494M (5’ GGN ARV AYD RYD CGR TTR TCD AC 3’) and COII 3400M (5’ ATY GGN CAY CAR TGR TAY TGA 3’) modified from Simon *et al.* (1994) for the mitochondrial cytochrome c oxidase subunit II (COII) gene; HB106F and HB903R (Mota *et al.*, 2008), and HBFMYCOS (5’ CAT ATA CGC AAG CAC AAG AAC C 3’) and HBRMYCOS (5’ GCT CRG CAC TGG CMG CAC 3’) for the nuclear hunchback (Hb) gene. Additionally, in order to estimate the evolutionary rate of COI, COII and hunchback (see below), the nuclear gene alpha methyldopa (AMD) was amplified for a set of mycophylic species (Table S3) with the primers AMDEX4F (Robe *et al*., 2010a) and AMDBW from Tatarenkov *et al.* (1999). Sequencing reactions were carried out on a MegaBACE 500 automatic sequencer with the use of the DYEnamic ET® Sequencing Kit (Amersham) or were performed by Macrogen (http://www.macrogen.com/eng/).

The electropherograms so obtained were assembled with the use of the Staden Package Gap 4 program (Staden, 1996), where each contig was individually checked in regard to sequence quality and editions were performed whenever necessary. Orthologous sequences were aligned with the Clustal W algorithm, as implemented in Mega 5 (Tamura *et al.*, 2011) and checked for the presence of stop-codons and frameshifts, in order to avoid the presence of Numts (nuclear copies of mitochondrial DNA fragments) in the mitochondrial datasets. Additionally, the ratio between the number of nonsynonymous substitutions per nonsynonymous sites (dN) and the number of synonymous substitutions per synonymous sites (dS) was assessed in Mega 5 using the Kimura 2-Parameters nucleotide substitution model (K2P) (Kimura, 1980). The same method was also used to perform a codon-based-Z test of purifying selection (for details see Robe *et al.*, 2012 and Machado *et al.*, in preparation).

Data analysis

Diversity analysis

As observed in the DNA barcode analyses of mycophylic drosophilids (Machado *et al.*, in preparation), *M. projectans* has at least two cryptic species that are undistinguishable even at the aedeagus level. So, we first added our new COI sequences to the mycophylic dataset previously generated by Machado *et al.* (in preparation) and constructed a Neighbor Joining tree (NJ) using the K2P model in Mega 5, in order to assess specimens identity. After that, diversity analysis [nucleotide (π) and haplotype (H and Hd) diversities, besides the number of polymorphic sites (s)] and neutrality tests [Tajima's D (Tajima, 1989), Fu's and Li's D and Fu's and Li's F (Fu, 1997)] were performed in DNAsp 5 (Librado & Rozas, 2009). The genealogical relationship between mtDNA haplotypes was estimated by median-joining (Bandelt *et al.*, 1999), as implemented in Network 4.6.1.3.

Dating and estimating rates

In order to estimate the nucleotide substitution rate of COI within the *Zygothrica* genus group, divergence time between some of their species were first estimated with the use of a lognormal relaxed clock (Drummond *et al.*, 2006), employing the newly obtained AMD sequences (Table S3) added to those generated by Robe *et al.* (2010a). In this case, an orthologous sequence of *Culex quinquefasciatus* (Genbank: XM001863772) was used as outgroup. The analysis was performed in Beast 1.8.0 (Drummond *et al.*, 2012), using the split between Hawaiian *Scaptomyza* and *Drosophila* species [23 Mya (Grimaldi, 1987)], and *Drosophila* and *Sophophora* subgenera [63 Mya (Tamura *et al.*, 2004)] as priors. The GTR+I+G nucleotide substitution model was applied to 50 million iterations, sampling every 5,000 and discarding the first 10% as burning. After, the results were evaluated in Tracer 1.6 and summarized in TreeAnnotator 1.8.0. The divergence times so obtained for the split between mycophylic species of the *Zygothrica* genus group were further used as priors in order to estimate the mean nucleotide substitution rate of COI, using a specific dataset obtained as stated above (Table S3). In this case, COI rate estimation was performed in Beast 1.8.0 using GTR+G, with 70 million of iterations sampling every 3,000 and burning the first 10%.

The divergence time estimated by Amd for the split between *Zygothrica* and *H. levigata* (see above) was also used as prior in another lognormal relaxed molecular clock run performed in Beast 1.8.0 with Hb sequences (Table S3) in order to estimate the divergence time between the three species of the *M. projectans* complex. In this case, the newly obtained Hb sequences were added to those generated by Robe *et al.* (2010b), and an orthologous sequence of *Musca domestica* (GenBank: Y13050) was used as outgroup. The run was performed using GTR+I+G with 50 million iterations sampling every 5,000 and burning the first 10% iterations.

The rate obtained for the COI mycophylic sequences ($5.828x10^{-9}$ per year) was finally used to estimate the fluctuation in population size over time using Bayesian skyline (BPS) (Drummond *et al*., 2005) and GMRF Bayesian skyride (GMRF) (Minin *et al.*, 2008), as implemented in Beast 1.8.0. For *M. projectans* and *M. projectans* affinis 1, GMRF was generated using HKY+I+G, with 40 million iterations, sampling every 5,000 and 1,000 respectively, with a burning of 10%. For *M. projectans* affinis 2 the BPS was generated using GTR+G+I, with 30 million iterations, sampling every 3,000 and burning the first 20%.

Morphometric analysis

For the morphometric analysis, the right wings of male specimens of the *M. projectans* complex (23 individuals of *M. projectans*, 13 of *M. projectans* affinis 1 and 19 of *M. projectans* affinis 2) was photographed in SteREO DiscoveryV20, with the program AxioVs40 V 4.8.1.0 (Carl Zeiss Imaging Solutions GmbH), with an increase of 37.5X. In order to obtain the coordinates, TPSdig2 (Rohlf, 2006) was used to mark ten landmarks chosen with basis on wings venation, representing bifurcations, intersections and tip veins (Fig. S1, Table S4). Partial and relative warps, centroid size and weight matrix were then calculated with TPSRelwn (Rohlf, 2007), and a graphic representation from shape variation was obtained. An analysis of variance (ANOVA) was performed to compare the centroid size of the wings of each species, and a canonical variate analysis (CVA) was performed to compare the relative warps of each species. ANOVA and CVA were performed in PAST 3 (Hammer *et al*., 2001).

Ecological niche modeling

The climatic variables [WorldClim 1.3 database (Hijmans *et al.*, 2005)] selected by Robe *et al*. (2014) [mean diurnal range (Bio-2), maximum temperature of

the warmest month (Bio-5), minimum temperature of the coldest month (Bio-6), precipitation seasonality (Bio-15), precipitation of the wettest quarter (Bio-16), precipitation of the driest quarter (Bio-17), precipitation of the warmest quarter (Bio-18) and precipitation of the coldest quarter (Bio-19)] with a resolution of 2.5-min (5 km^2) were used model potential distribution of each species of the *M. projectans* complex with basis on an environmental niche modeling (ENM) approach (Elith & Leathwick, 2009). As the cryptic diversity contained within the *M. projectans* complex was previously neglected, only the sampling points recorded by us (Table S2) were used in this analysis. The models were generated using maximum entropy machine learning method performed by MAXENT 3.3.3k (Phillips *et al.*, 2006) using 25% of the data as test and 75% as training or calibration information, in each of 50 bootstrap replicates. For these, in addition to the presence records, 10,000 backgrounds points were randomly chosen for each species. The maximum number of iterations was set to 500 and a default regularization parameter was used (Phillips & Dudik, 2008). The accuracy of the predictive distribution models was accessed through the AUC value (Area Under the Receiving Operating Curve).

The environmental suitability models so obtained were further used to measure Schoener's D (D) and Hellinger's (I) pairwise niche overlap in ENMTools 1.3 (Warren *et al.*, 2010), where 0 means no overlap and 1 indicates that niches are identical. The statistical significance of each pairwise niche overlap measure was tested using the identity test performed in the same program, with 200 pseudoreplicates. In this case, a conservative criteria was adopted, since null hypothesis was considered to be rejected only when both D and I comparisons were significant. Pairwise range overlaps (R) were also estimated using the binary models obtained by MAXENT, converted with the mean minimum training presence threshold. To test the correlation between niche and geographic measures, a Mantel test was performed in Past 3 with 5,000 permutations.

Results

Sampling and genetic diversity

We generated 103 new *Mycodrosophila* COI sequences, which ranged from 550pb to 855pb, and were added to the 32 *Mycodrosophila* sequences recorded by Machado *et al.* (in preparation). The analyses performed in order to detect Numts

allowed to reject their influence in species delimitation, since no frameshifts and stop codons were detected, and all the sequences that did not reject the null hypothesis of neutrality in codon-Z analysis and showed a dN/dS ratio near one, presented low K2P distance (d < 1.3%), indicating stochastic deviation in intraspecific variation.

When these sequences were analyzed in regard to species subdivision in a NJ phylogram (Fig. S2), we realized that they were compatible with the presence of at least ten Neotropical *Mycodrosophila* species, three of which were determined to the *M. projectans* morphotype and will be hereafter referred as the *M. projectans* complex. Moreover, cryptic diversity was also encountered for the *M. neoprojectans* and M001 morphotypes. The minimum K2P pairwise distance presented between these putative species was of the order of 4.6%.

Even with such a subdivision, the intraspecific distances ranged from 0% in several species to 5.8% in *M. projectans* (Table S2). The nucleotide diversity, number of polymorphic site and haplotype diversity were generally high, reaching values of 1.4%, 46 and 0.96 in *M. projectans* (Table S2). Signs of significant deviation of neutrality were found for different combinations of the three *M. projectans* complex species according to the Tajima's D, Fu and Li's D and/or Fu and Li's F. Diversity indices found for COII were even higher than those presented by COI (Table S2), although neither this dataset nor the concatenated COI+COII showed significant deviations from neutrality.

The phylogenetic tree generated with AMD (data not shown) supported the monophyly of *Mycodrosophila* (PP = 1.00), which was positioned inside the *Drosophila* genus (PP = 0.95). Although the monophyly of the *Zygothrica* genus group was not recovered, *Drosophila* was also paraphyletic in regard to *Zygothrica* and *Hirtodrosophila*. In fact, representatives of these two genera constituted a well supported clade (PP = 1.00), in which *Hirtodrosophila* is paraphyletic in regard to *Zygothrica* (PP = 1.00).

The *Mycodrosophila projectans* complex

Actual and potential distribution and niche overlap patterns

The three species of the *M. projectans* complex showed to be sympatric to an area encompassing the Southern region of Brazil (Fig. 1b, Table S2). ENM showed *M. projectans* affinis 1 as a very restrict species, with *M. projectans* more widely

distributed to the south and *M. projectans* affinis 2 further expanded to the north and west of the South American continent (Figure S3). Species of the complex were found exploring fungi of *Ganoderma*, *Polyporus* and *Trametes* (Table S5), and at least twice the three species were found in syntopy. In these cases, the triad was simultaneously collected using *Ganoderma* and another unidentified fungus. *M. projectans* and *M. projectans* affinis 2 were collected together exploring *Ganoderma* and *Polyporus* (Table S5).

Abiotic niche overlap measures generated based on ENM results were higher in the comparisons involving *M. projectans* affinis 1 e 2 (D = 0.87, I = 0.99), than when any of these is compared to *M. projectans* (D = 0.38, I = 0.69; D = 0.43, I = 0.73 respectively). Nevertheless, null hypothesis of niche equivalence was not rejected in any of these comparisons after a Bonferroni correction was applied. Range overlap measures were higher between *M. projectans* and *M. projectans* affinis 2 (R = 1), than with the former and *M. projectans* affinis 1 (R = 0.87) of with both affinis species (R = 0.94). No significant correlations were found between niche and range overlap measures among the three species after a Bonferroni correction.

Evolutionary patterns

Like the neutrality tests, COI median-joining analysis of the *M. projectans* complex (Fig. 2) suggested demographic expansion events since they showed some kind of star like pattern for *M. projectans*, *M. projectans* affinis 1 and *M. projectans* affinis 2, with central and more frequent haplotypes connected with peripheral and exclusive haplotypes through few mutational steps. However, there are some exceptions, like haplotypes 19 (Derrubadas/RS) and 28 (Pelotas/RS) in *M. projectans* and haplotypes 3 (Teodoro Sampaio/SP) and 6 (Melgaço/PA) in *M. projectans* affinis 2, that are separated by up to 6-13 mutational steps from their most close haplotype. Although the differentiation between species attained a minimum of 34 mutational steps for COI, intraspecific genetic flow seems to be maintained despite long distances, as suggested by the presence of few mutational steps between haplotypes encountered in the extremes of distribution (data not shown). Thus, in *M. projectans* affinis 2, for example, haplotypes that were collected about 3,000 Km apart, like haplotype 19 and haplotypes 4, 5 and 6, were differentiated by only 4-8 mutational steps. Likewise, at the same species, haplotypes differing by as much as 34 mutational steps were found in the same point. This pattern was

somewhat diluted in the concatenated network, due to reduced sampling, although differentiation between species was even more pronounced (Fig. S4).

As concerns the general patterns of genetic structure, there are some differences between the three *M. projectans* complex species. In this case, *M. projectans* seems to be the only exhibiting a geographic structure between the sampled biomes (Atlantic Forest and Pampa). So, in the COI network (Fig. 2), two haplogroups could be observed for this species: one, with the most common haplotype sampled on west (H8 and H23) and radiating to east, seems to be restricted to Atlantic forest (only one individual found in the border between Atlantic forest and Pampa); and the other almost entirely restricted to the Pampa biome, with several haplotypes being sampled in the border between Atlantic forest and Pampa and radiating to the south. Conversely, the COI networks of *Mycodrosophila projectans* affinis 1 and 2 (Fig. 2) present the most common haplotype originating from the west of Atlantic forest, from which it radiates haplotypes to the east in *M. projectans* affinis 1 and to east and north in Atlantic forest, northwest to Pampa and north to Amazonian biomes in *M. projectans* affins 2.

The divergence time estimated with the use of a 216 - 398 pb of a conserved portion of the Hb gene showed that *M. projectans* affinis 1 is the basal species in the complex, diverging approximately 16.5 million of years ago (6,9 - 29, 3), whereas *M. projectans* and *M. projectans* affinis 2 are sister species that diverged about 10.5 million of year ago (3,8 - 19). All these clades were supported by PP of 1. BSP and GMRF (Fig. S5) showed agreed with the neutrality tests and the starlike network pattern, suggesting all species experienced an expansion of population size in different moments. *Mycodrosophila projectans* affinis 1 showed a stable population size until 100 thousand year ago when an expansion have started. Yet *M. projectans* and *M. projectans* 2 experienced earlier expansions, initiated about 1 million year ago, and intensified at 650 thousand years ago for the second species.

Morphological differentiation

Although in Machado *et al.* (in preparation) it was not possible to assess the male aedeagus structure of *M. projectans* affinis 1, due to the sampling of only two females, that study found no differences between the male genitalia patterns found for *M. projectans* and *M. projectans* affinis 2. The sampling extension performed here revealed the presence of two male genitalia patterns within the *M. projectans*

complex (Fig. S6), with unique and distinguishable aedeagus structures found for *M. projectans* and *M. projectans* affinis 1, alghouth both patterns are encountered in *M. projectans* affinis 2. As external morphology was entirely cryptic between the three species and male genitalia provided inconclusive differentiation, wing morphometry was also tested as a potential morphological marker in species identification. Nevertheless, the ANOVA and CVA tests performed in order to detect differences related to the centroid size of wings and the scores of the relative warps, respectively (Table S6, Fig. 3) did not show any statistically significant difference between species. In fact, as a wide overlap of wing shape was detected between species (Fig. 3), wings also proven to be inadequate to distinguish species within the *M. projectans* complex.

Discussion

DNA barcode has demonstrated be a valuable tool to reveal cryptic diversity (Hebert *et al*., 2004), presenting a huge impact in groups with low knowledge or with hard taxonomy (Hebert *et al*., 2004; Clare *et al*., 2011; Crawford *et al*., 2012; Mutanen *et al.*, 2013; Gill *et al*., 2014). Machado *et al.* (in preparation) showed the great potential of this methodology in the *Zygothrica* genus group, revealing a wide variety of potential new species, some of which entirely cryptic in regard to each other. In *Mycodrosophila*, these authors suggested the presence of at least six putative new species, triplicating the number of species recorded in Brazil. Three of these were first identified as *M. projectans*, and could not be distinguished even at the aedeagus level. Here, we increased the sampling of these three putative new species, and added individuals assigned to more two putative new species of the genera. Together, Machado *et al.* (in preparation) and this paper increase more than 360% the species diversity of this genus in Neotropics, previously recorded as only three (Wheeler & Takada, 1963).

Interestingly, COI and COII genes showed levels of nucleotide and haplotype diversity considerable higher for most of the sampled *Mycodrosophila* species than to the previously studied Drosophilidae species (De Brito *et al.*, 2002; Hurtado *et al.*, 2004; Reed *et al.*, 2007; Mirol *et al.*, 2008; Moraes *et al*., 2009; Franco & Manfrin, 2013, De Ré *et al*., 2014). In fact, nucleotide diversity for *M. projectans stricto sensu* reached more than 10,000 times in the values found for *Drosophila maculifrons* (De

Ré *et al.*, 2014) with COI and 18,000 times for COII. Although, these high diversity indices could indicate Numts co-amplification in our database, tests performed to identify this artifact did not detect such an effect, so that this high diversity could be a characteristic of *Mycodrosophila* and possibly of other mycophylic species. Despite this high diversity, DNA barcode methodology was shown to be an efficient tool in species identification in this group (Machado *et al.*, in preparation).

Other remarkable characteristic of this genus is the recurrent sympatry and syntopy found between cryptic species. This property was first documented by Lacy (1982) for *M. claytone* through the analysis of 18 allozime loci. The author found two cryptic species, *M. claytone* A and B, which do not share any alleles in six loci occurring in syntopy. We found the same situation for *M. neoprojectans* affinis 1 and 2 and for *M. projectans* complex. The first species pair was collected only in Amazonian Biome, being recorded in sympatry and syntopy in only sampling point. However, it should be noted that the sampling realized in these areas were no extensive, so that both sympatry and syntopy could be more ubiquitous between both species. In the same way, the *M. projectans* complex species, which were more widely sampled, were frequently found in sympatry and syntopy. This makes this genus an interesting model to the study of speciation.

***Mycodrosophila projectans* complex**

Despite none morphological differentiation could be found between at least two of the three species of *M. projectans* complex in external morphology, male genitalia and wing shape and dimensions, the divergence time between the species was dated to the Miocene. *Mycodrosopila projectans* affinis 1 was the first lineage to diverge, at approximately 16.5 million years ago. In fact, our morphological identification distinguished correctly all male individuals of this species from the others. However, *M. projectans* affinis 2 that has diverged from *M. projectans stricto sensu* about 10.5 million years ago, has individuals *a priori* identified both as *M. projectans* affinis 1 and *M. projectans.* Yet the specimens of the last species always were identified as *M. projectans.*

Besides the morphological conservation, patterns of biotic and abiotic niche conservatism were also found for the *M. projectans* complex, in agreement with Robe *et al.* (2014), which demonstrated that niche conservatism seems a tendency in *Zygothrica* genus. This general pattern seems to apply for the *M. projectans*

complex, once despite the difference between amplitude of distribution, neither pairwise comparison rejected the null hypothesis of niche equivalency. This result was already expected once the three species are sympatric and syntopic, but ecological studies could certainly help to clarify in which way the species are using their resources, which could enable advances in speciation studies within this model group.

Another interesting result was obtained by the BSP and GMRF analysis, which showed signals of expansion on effective population size for the three studied species. Although these events were dated to the Pleistocene, they seemed to occur in different moments, with *M. projectans stricto sensu* and *M. projectans* affinis 2 starting their expansion between Calabrian and Middle Pleistocene (Cohen *et al.*, 2013) at approximately one million of years ago, and *M. projectans* affinis 1 starting its expansion in the Upper Pleistocene (Cohen *et al.*, 2013), at about 0.1 million of years ago. These results are concordant with phylogeographic studies performed with Brazilian cactophilic drosophilds, that also showed population expansion dated to the Pleistocene (Franco & Manfrin, 2013). This general pattern suggests that some drosophilids species have been subject of significant oscillations in population dynamics with the glacial cycles of this Epoch (Morales *et al.*, 2009; Franco & Manfrin, 2013). However, in others cases, the glaciations seem to have influenced population dynamics in the contrary way, a pattern recently supported for *D. maculifrons*, that presents an expansion signal after the last maximum glacial (De Ré *et al.*, 2014). Our results are also consistent with phylogeographic studies with other animal groups, where early Pleistocene seem to have more effect in population dynamics than the more recent glaciations (see reanalyses of Martins, 2011). However, although these last studies were also performed in Atlantic Forest, they showed a lineage breaking in the least of this region, whereas *M. projectans* complex seems to be more abundant and common to the west of this biome.

Acknowledgments

We are grateful to the authorities responsible for Brazilian Ecological Stations, Reserves or Parks and to the private properties owners that allowed us to perform Drosophilidae searches in their area. All these collections were authorized by the Brazilian Ministério do Meio Ambiente (MMA), in the form of the Sistema de

Autorização e Informação em Biodiversidade (SISBIO), and we thank them for their commitment to ethical guidelines and conservation efforts concerning Brazilian scientific activities. We would also like to thank to Dr. Eduardo Bernardi for the fungi identification and to M.Sc. João Pedro Junges dos Santos for his help in drosophilids identification. This study was supported by Universal-CNPq 14/2013, process numbers 471174/2013-0 and 472973/2013-4.

References

Bächli G, Vilela CR, Escher SA, Saura A. 2004. *The Drosophilidae (Diptera) of Fennoscandia and Denmark.* Bill.

Bächli G. 2015. TaxoDros: The database on taxonomy of Drosophilidae. Version 1.04. Available from: http://www.taxodros.uzh.ch/.

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

Carson HL. 1971. The ecology of *Drosophila* breeding sites. Harold L. Lyon Arboretum Lecture Number Two. The University of Hawaii, Honolulu; p. 1–27.

Clare EL, Lim BK, Fenton MB, Hebert PDN. 2011. Neotropical Bats: estimating species diversity with DNA Barcodes. *PLoS ONE* **6**: e22648.

Cohen KM, Finney SC, Gibbard PL, Fan J-X. 2013. The ICS International Chronostratigraphic *Chart. Episodes* **36**: 199-204.

Courtney SP, Kibota TT, Singleton TA. 1990. Ecology of mushroom-feeding Drosophilidae. *Advances in Ecological Research* **20**: 225–274.

Crawford AJ, Cruz C, Griffth D, Ross H, Ibáñez R, Lips KR, Driskell AC, Bermingham E, Crump P. 2012. DNA barcoding applied to *ex situ* tropical amphibian conservation programme reveals cryptic diversity in captive populations. *Molecular Ecology Resources* **13(6)**: 1005-18.

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

De Ré FC, Gustani EC, Oliveira APF, Machado LPB, Mateus RP, Loreto ELS, Robe LJ. 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.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005. *Molecular Biology and Evolution* **22**: 1185-1192

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. *PLoS Biology* **4**, e88.

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

Elith J, Leathwick JR. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review Of Ecology, Evolution, and Systematics* **40**: 677-697.

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(5)**: 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*.*

Gill BA, Harrington RA, Kondratieff BC, Zamudio KR, Poff NLR, Funk C. 2014. Morphological taxonomy, DNA barcoding, and species diversity in southern Rocky Mountain headwater streams. *Freshwater Science* 33(1): 288-301.

Gottschalk MS, Bizzo L, Döge JS, Profes MS, Hofmann PRP, Valente VLS. 2009. Drosophilidae (Diptera) associated to fungi: differential use of resources in anthropic and Atlantic Rain Forest areas. *Iheringia Série Zoologia* **99**: 442-448.

Grimaldi DA. 1987. Phylogenetics and taxonomy of *Zygothrica* (Diptera: Drosophilidae). *Bulletin of the American Museum of Natural History* **186**: 103–268.

Grimaldi DA. 1990. A phylogenetic revised classification of genera in the Drosophilidae (Diptera). *Bulletin of the American Museum of Natural History* **197**: 1–139.

Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica* **4(1)**: 9.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator. Proceedings of the National Academy of Sciences of the United States of America* ***101(41)****: 14812-14817.*

Hijmans RJ, Cameron SE, Parra JL. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965-1978.

Hurtado LA, Erez T, Castrezana S, Markow TA. 2004. Contrasting population genetic patterns and evolutionary histories among sympatric Sonoran Desert cactophilic *Drosophila. Molecular Ecology* **13***:* 1365–1375.

Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**: 111-120.

Lacy RC. 1982. Niche breadth and abundance as determinants of genetic variation in populations of mycophagous drosophilid flies (Diptera: Drosophilidae). *Evolution* **36(6)**: 1265-1275.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Machado S, dos 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.

Markow TA, O'Grady P. 2008. Reproductive ecology of *Drosophila*. *Functional Ecology* **22**: 747-759.

Martins FM. 2011. Historical biogeography of the Brazilian Atlantic forest and the Carnaval–Moritz model of Pleistocene refugia: what do phylogeographical studies tell us? *Biological Journal of the Linnean Society* **104**: 499–509.

McEvey SF, Polak M. 2005 *Mycodrosophila* (Diptera: Drosophilidae) of Fiji and Vanuatu With Description of Nine New Species. *Bishop Museum Occasional Papers* **84**: 35–67.

Minin VN, Bloomquist EW, Suchard MA. 2008. *Molecular Biology and Evolution* **25**: 1459-1471.

Mirol PM, Routtu J, Hoikkala A, Butlin RK. 2008. Signals of demographic expansion in *Drosophila virilis. BMC Evolutionary Biology* **8***:* 59.

Moraes EM, Yotoko KSC, Manfrin MH, Solferini VN, Sene FM. 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.

Mota NR, Robe LJ, Valente VLS, Budnik M, Loreto ELS. 2008. Phylogeny of the *Drosophila mesophragmatica* group (Diptera, Drosophilidae): an example of Andean evolution. *Zoological Science* **25**: 526–532.

Mutanen M, Kaila K, Tabell J. 2013. Wide-ranging barcoding aids discovery of one-third increase of species richness in presumably well-investigated moths. *Scientific Reports* **3(2901)**: 1-7.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological modelling* **190**: 231-259.

Phillips SJ, Dudik M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation*. Ecography* **31**: 161-175.

Powell JR. 1997. Progress and prospects in evolutionary biology: the *Drosophila* model. New York: Oxford University Press.

Reed LK, Nyboer M, Markow TA. 2007. Evolutionary relationships of *Drosophila mojavensis* geographic host races and their sister species *Drosophila arizonae. Molecular Ecology* **16**: 1007–1022.

Robe LJ, Loreto ELS, Valente VLS. 2010a. Radiation of the ,,*Drosophila*" subgenus (Drosophilidae, Diptera) in the Neotropics. *Journal of Zoological Systematics and Evolutionary Research* **48**: 310–321.

Robe LJ, Valente VLS, Loreto ELS. 2010b. Phylogenetic relationships and macro-evolutionary patterns within the Drosophila tripunctata "radiation" (Diptera: Drosophilidae). *Genetica* **138(7)**: 725-735.

Robe LJ, Machado S, Bartholomei-Santos ML. 2012. The DNA Barcoding and the caveats with respect to its application to some species of Palaemonidae (Crustacea, Decapoda). *Zoological Science* **29**: 714–724.

Robe LJ, Machado S, Bolzan AR, dos Santos JPJ, Valer F, dos Santos AP, Blauth ML, Gottschalk MS. 2014. Comparative ecological niche modeling and evolutionary ecology of Neotropical mycophagous Drosophilidae (Diptera) species. *Studies on Neotropical Fauna and Environment* **49(2)**: 79-94.

Rohlf FJ. 2006. tpsDig2, digitize landmarks and outlines, version 2.10. Dept Ecology & Evolution, St. Univ. New York, Stony Brook, New York.

Rohlf FJ. 2007. tpsRelw, relative warps analysis, version5/13/05. Dept Ecology & Evolution, St. Univ. New York, Stony Brook, New York.

Roque F, Tidon R. 2008. Eight new records of drosophilids (Insecta: Diptera) in the Brazilian savanna. *Drosophila Information Service* **91**: 94- 98.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annual Review of Entomology* **87**: 651-701

Staden R. 1996. The Staden sequence analysis package. *Molecular Biotechnology* **5(3)**: 233-41.

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

Tamura K, Subramanian S, Kumar S. 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Molecular Biology and Evolution* **21**:36–44.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**: 2731–2739.

Tatarenkov A, Kwiatowski J, Sharecky D, Barrío E, Ayala FJ. 1999. On the evolution of *Dopa decarboxylase* (*Ddc*) and *Drosophila* systematics. *Journal of Molecular Evolution*. **48**:445–462.

Throckmorton LH. 1975. The phylogeny, ecology and geography of Drosophila. In: King RC, ed. *Handbook of genetics*. New York: Plenum, 421–469.

Warren DL, Glor RE, Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33**: 607-611.

Wheeler MR, Takada H. 1963. A revision of the American species of *Mycodrosophila* (Diptera, Drosophilidae). *Annals of the Entomological Society of America* **56(3)**: 392-399.

Figures

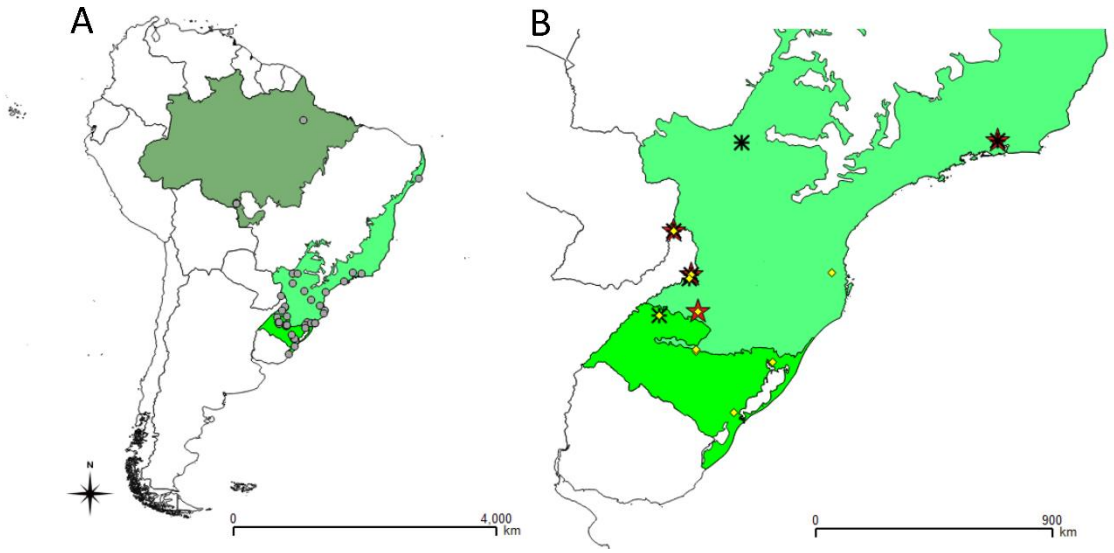

Figure 1: Map of sampling sites of the *Mycodrosophila* mycophylic species across Brazil. Brazilian biomes where collections were performed are highlighted. In A Sampling points are represented by circle; in B sampling sites of *M. projectans* complex are highlighted, with diamonds representing *M. projectans*, full stars *M. projectans* affinis 1 and line stars *M. projectans* affinis 2. *M. projectans* affinis 2 was also sampled in Melgaço/PA, placed within the Amazonian Biome (dark green in A), but it was omitted from map B.

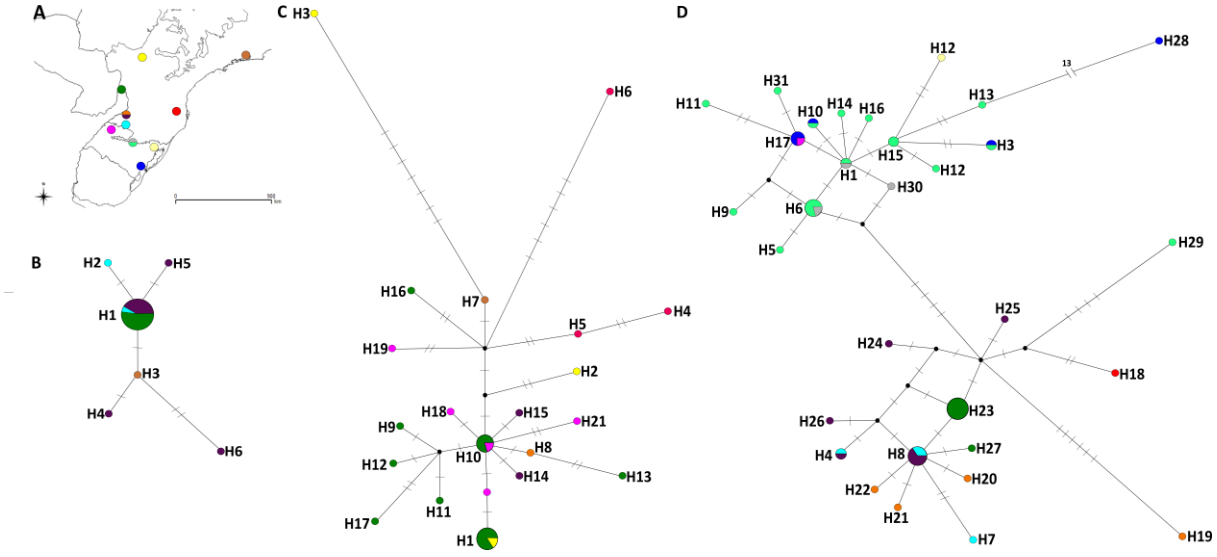

Figure 2: Median-joining network for COI haplotypes found within the *M. projectans* complex. A Map showing the sampling points where the species were found (circles), with limits between the biomes highlighted. Each sampling point was painted with different colors to locate the geographic origin of each haplotype. B *M. projectans* affinis 1, C *M. projectans* affinis 2 and D *M. projectans* networks. The size os the circles is proportional to the frequency of each haplotype, and sampling localities are color coded according to A. Black small circles represent median vectors. Dashes in the lines connecting different haplotypes represent the number of mutations between them. Dark pink in B represent a sampling point in north of Brazil (Melgaço/PA), that was omitted from the map A to facilitate visualization of the south area.

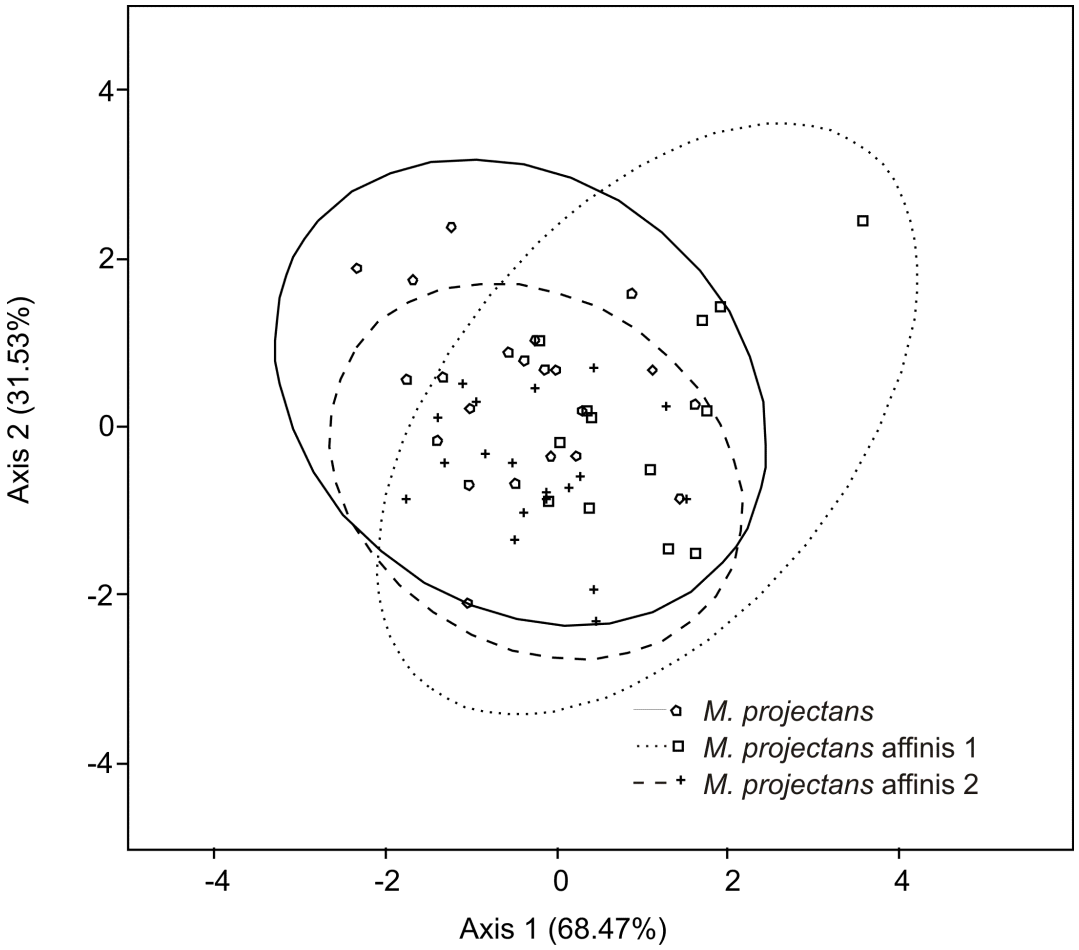

Figure 3: Cartesian plot from wings relative warps for *M. projectans* complex.

Supporting information

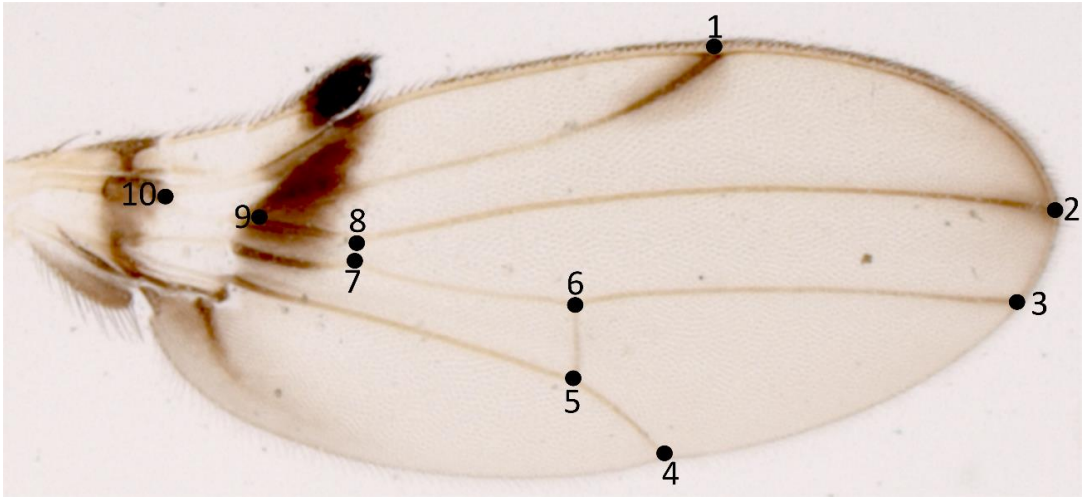

Figure S1: *M. projectans* complex right wing showing the chosen landmarks.

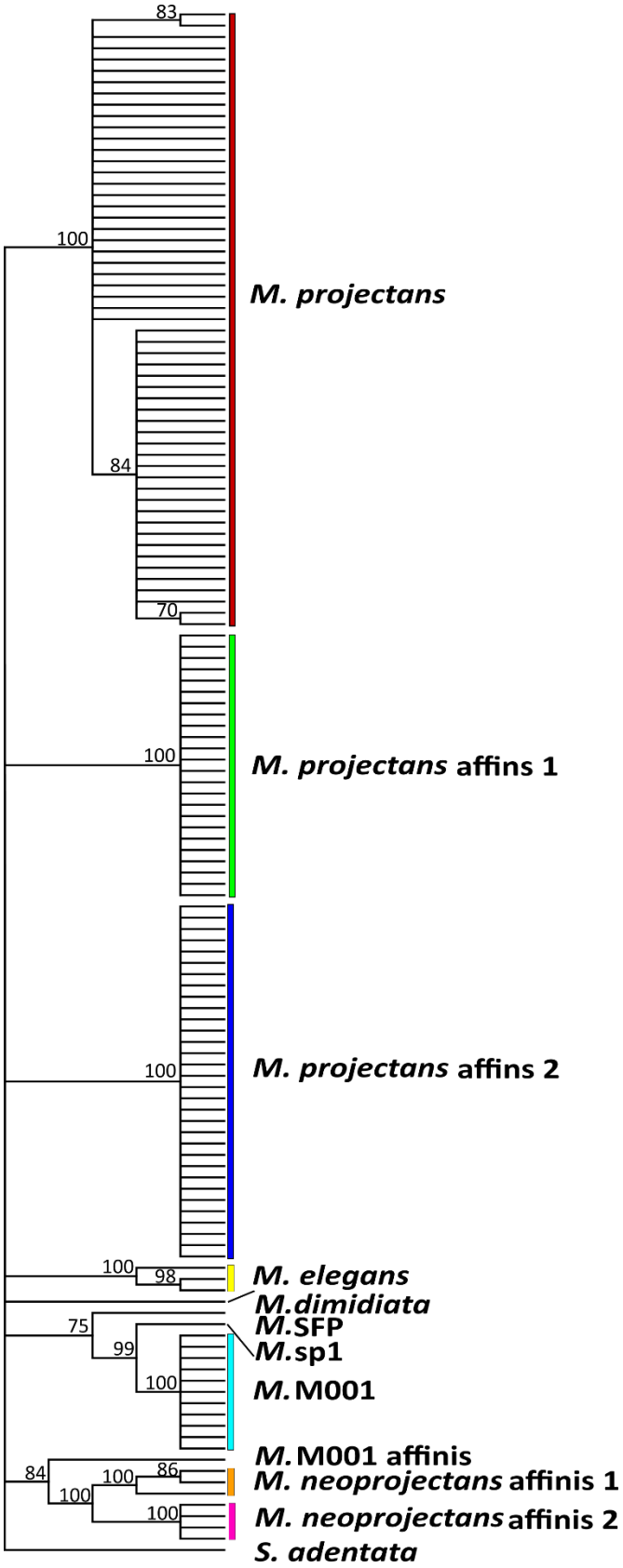

Figure S2: NJ phylogram of *Mycodrosophila* based on sequences of the COI gene. Numbers above the internal branches are bootstrap support.

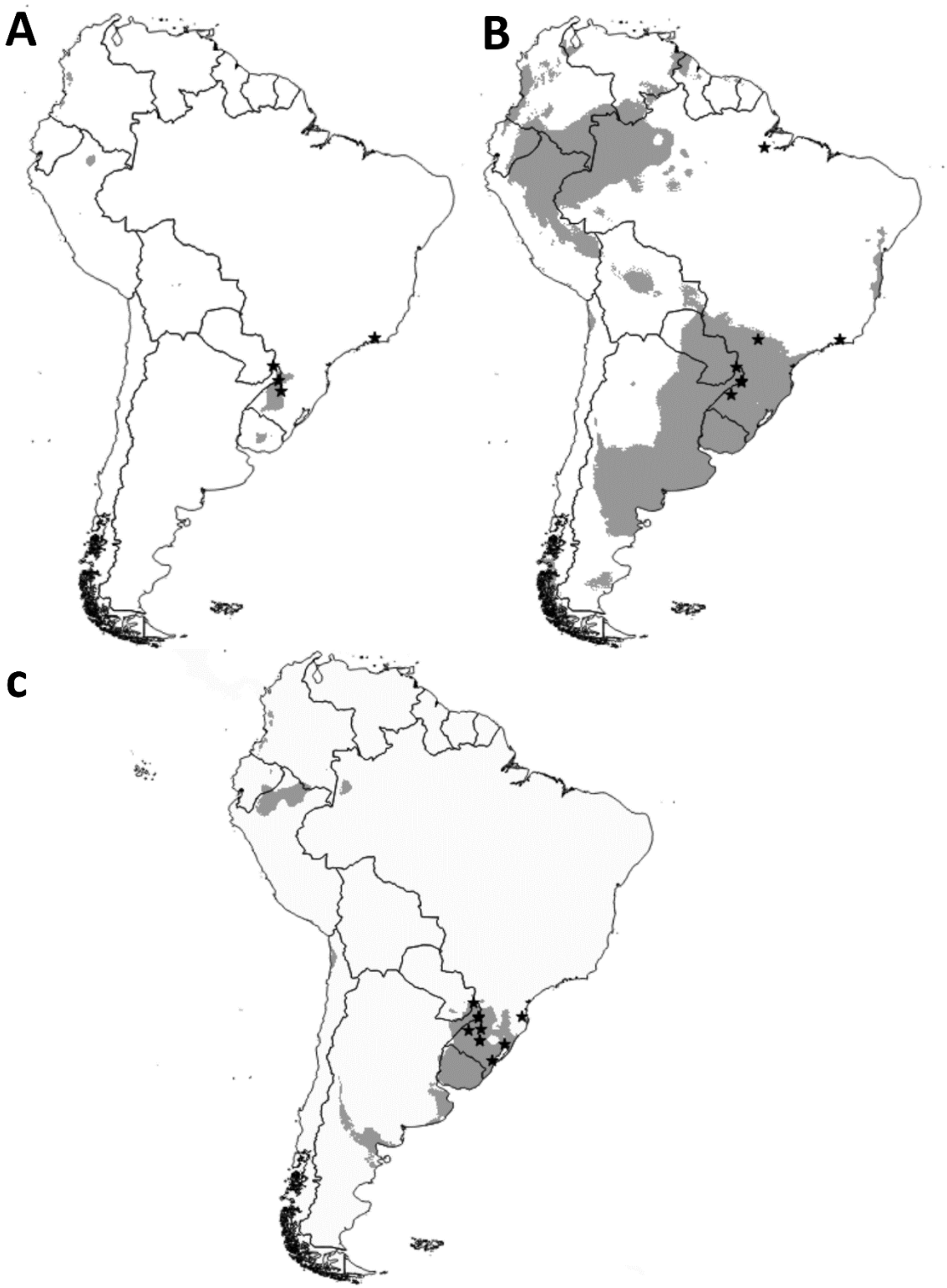

Figure S3: Environmental suitability distribution models predicted by MAXENT for *M. projectans* complex. A *M. projectans* affinis 1, B *M. projectans* affinis 2 and C *M. projectans*. Grey represent areas of potential presence after the applying the minimum training presence threshold. Stars represent the sampling points according to Tables S1 and S2.

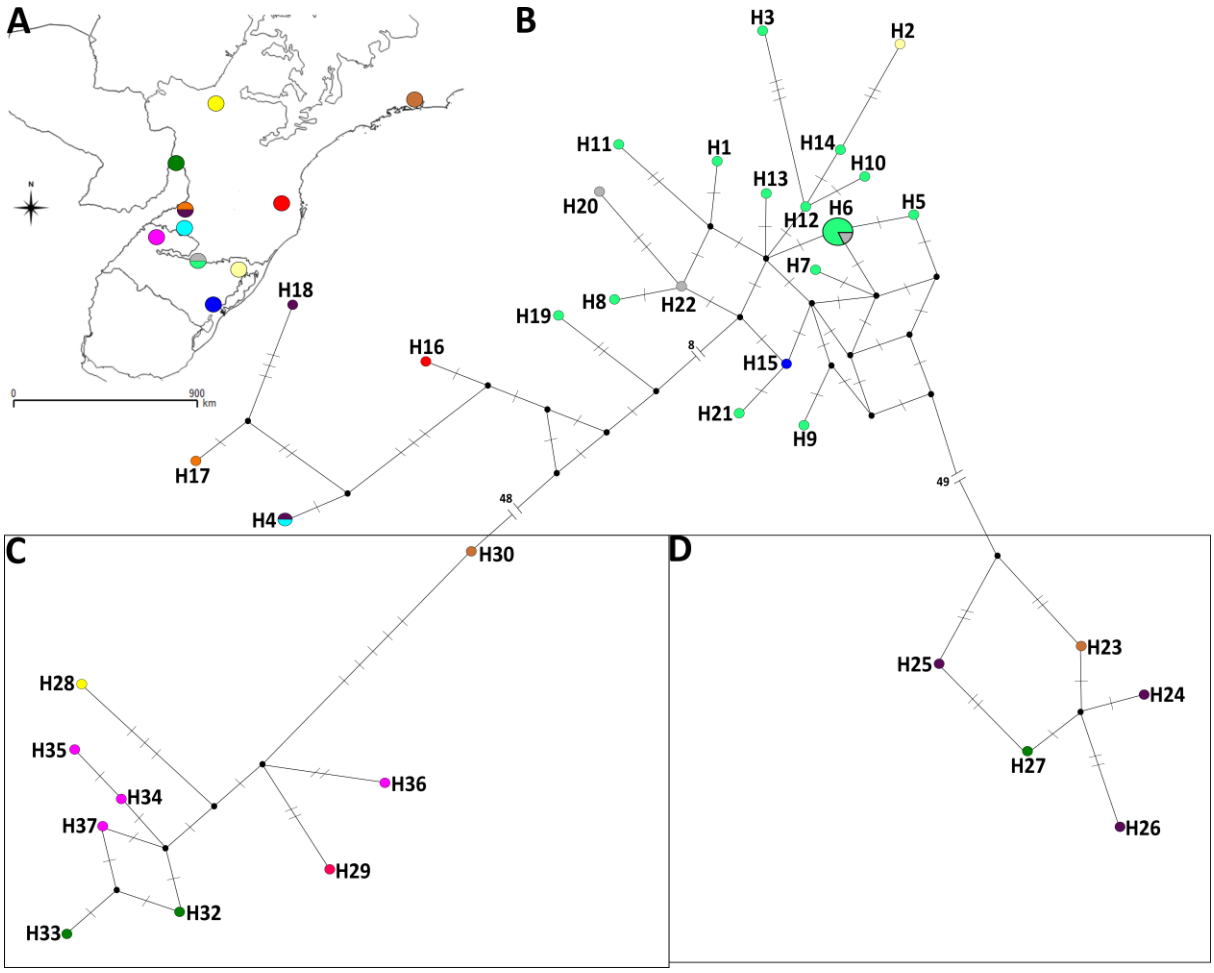

Figure S4: Median-joining network for COI+COII concatenated haplotypes found within the *M. projectans* complex. A Map showing sampling points where the species were found (circles), with limits between the biomes highlighted. Each sampling point was painted with different colors to locate the geographic origin of each haplotype. B *M. projectans*, C *M. projectans* affinis 1 and D *M. projectans* affinis 2 networks. The size os the circles is proportional to the frequency of each haplotype, and sampling localities are color coded according to A. Black small circles represent median vectors. Dashes in the lines connecting different haplotypes represent the number of mutations between them. Dark pink in C represents a sampling point in north of Brazil (Melgaço/PA), that was omitted from the map A to facilitate visualization of the south area.

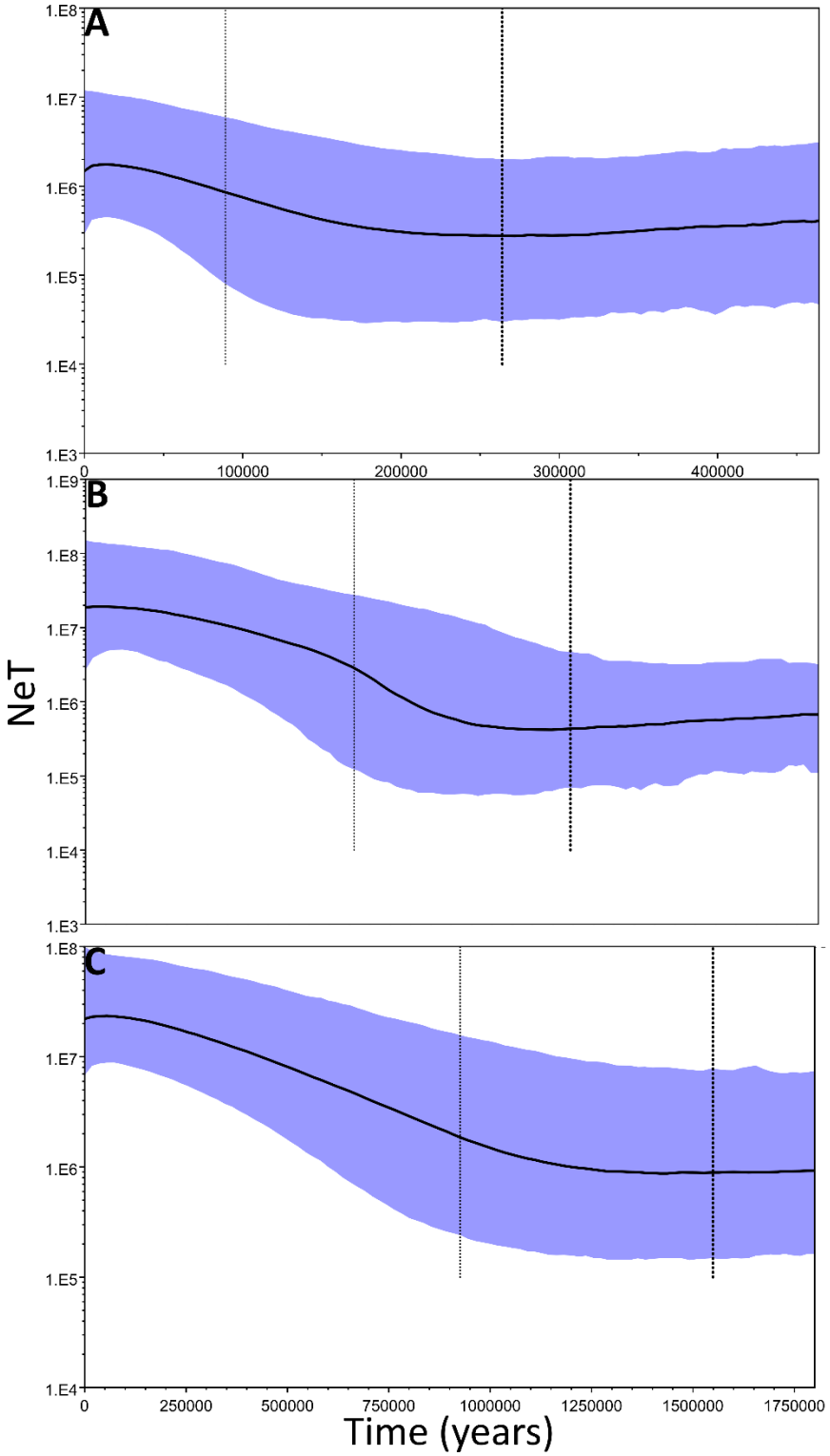

Figure S5: Plots showing the variation in effective population size (Ne) over time in the three *M. projectans* complex species according to the COI dataset. A and C GMRF Bayesian skyride obtained for *M. projectans* affinis 1 and *M. projectans*, respectively, and B Bayesian skyline of *M. projectans* affinis 2. Solid lines represent the estimated median Ne, and the solid area in blue indicates the 95% highest posterior probability interval.

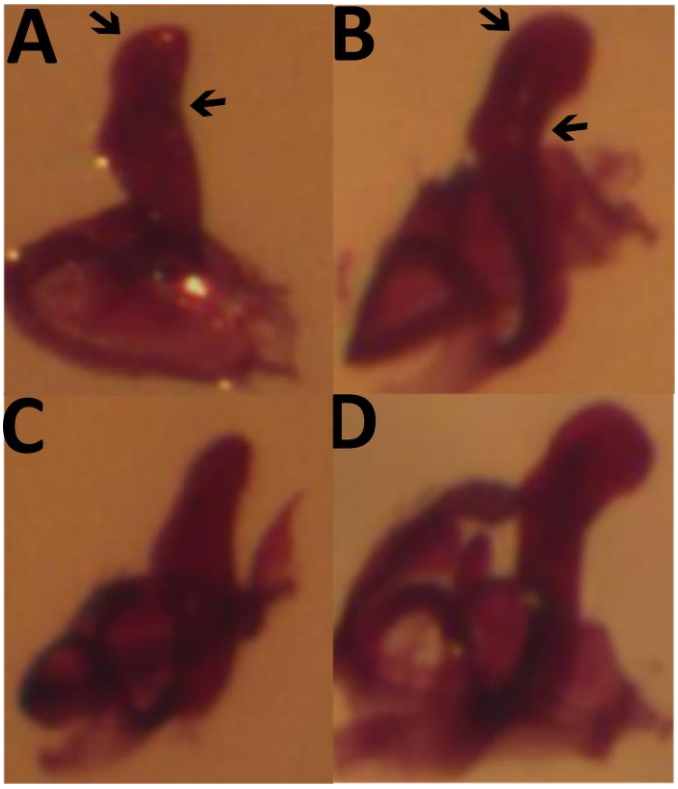

Figure S6: Male genitalia patterns found within the *M. projectans* complex. A *M. projectans*, B *M. projectans* affinis 1, C and D *M. projectans* affinis 2. The differences in aedeagous structure are indicated by arrows.

4. DISCUSSÃO

O grupo genérico *Zygothrica* proposto por GRIMALDI (1990a) é um grupo especioso que compreende aproximadamente 10% da diversidade de Drosophilidae (BÄCHLI, 2015), sendo composto tanto por espécies cosmopolitas, como por táxons restritos a determinadas regiões biogeográficas (GRIMALDI, 2010; VILELA E BÄCHLI, 2004; WHEELER E TAKADA, 1963). Entretanto, esse grupo tem sido alvo de poucos estudos, possivelmente devido (i) à dificuldade de manutenção de estoques em laboratório; (ii) à pouca efetividade dos métodos tradicionais de coleta na captura das mesmas, tornando necessária a realização de buscas ativas; (iii) à coleta ocasional e em baixo número registrada para muitas espécies; (iv) à dificuldade de identificação relacionada ao baixo número de taxônomos especialistas no grupo, ao difícil acesso à literatura (parte em alemão) e ao fato de que a maioria dos trabalhos descrevem apenas os machos. Estes problemas tornam-se ainda mais sérios na região Neotropical, onde existe um grande viés na distribuição taxonômica dos estudos relacionados à Drosophilidae (GOTTSCHALK *et al.*, 2008).

Um dos primeiros passos para superar esses entraves é a compreensão da diversidade de espécies do grupo, e o incentivo à difusão do conhecimento necessário para a correta identificação de suas espécies. Nesse sentido nós realizamos coletas em 45 sítios distribuídos nos biomas Amazônia, Mata Atlântica e Pampa no Brasil, dos quais 20 foram incorporados na análise de DNA *barcode* (Capítulo 3). Nesse trabalho, que incorporou 222 indivíduos, o ponto mais importante a ser ressaltado é, sem dúvida, a grande diversidade de espécies encontradas [56 espécies, das quais apenas 24 (39,3%) foram previamente descritas]. Dentre os fatores que contribuíram para essa alta frequência de espécies novas está o esforço de coleta, uma vez que esta tese é a primeira a ser realizada com uma ampla amostragem visando analisar drosofilídeos micofílicos.

A maior parte das espécies micofílicas potencialmente novas amostradas revelaram-se crípticas com relação a espécies previamente descritas. Tradicionalmente, o primeiro passo no reconhecimento e na identificação das espécies de drosofilídeos é baseado na morfologia externa, embora os padrões de coloração do tórax e abdômen sejam sabidamente uniformes entre alguns táxons e plásticos dentro de outros, dificultando sua utilização como carácter diagnóstico (GRIMADI, 2010). Dessa forma, a genitália masculina, principalmente a estrutura do edeago, tem sido considerada a estrutura diagnóstica mais importante para a

definição de espécies na família, sendo utilizada tanto para seu reconhecimento, quanto descrição e sinonímias (MEDEIROS E KLACZO, 2004; VILELA E BÄCHLI, 1990). Dessa forma, nós verificamos dois níveis de diferenciação entre as espécies crípticas por nós amostradas: aquelas crípticas em nível de morfologia externa, mas distinguíveis pela genitália masculina e aquelas crípticas devido à ausência de distinção inclusive em nível de genitália. Oito das 34 potencias espécies novas foram consideradas crípticas no segundo caso, embora as mesmas apresentassem níveis de diferenciação genética superiores a 5.5% para a sequência do marcador mitocondrial COI. Além disso, 16 das espécies potencialmente novas amostradas revelaram-se crípticas com relação a espécies previamente descritas apenas no primeiro nível, e dez foram distinguíveis pelo padrão de morfologia externa, não sendo consideradas crípticas.

Em nossa amostragem, todos os gêneros Neotropicais do grupo genérico *Zygothrica* foram coletados. O pequeno gênero Neotropical *Paraliodrosophila* (VILELA E BÄCHLI, 2007), foi representado por duas das cinco espécies descritas (*P. antennata* e *P. burlai*), sendo justamente essas duas com registros no Brasil (BÄCHLI, 2015). GOTTSCHALK *et al.* (2009) encontraram *P. antennata* emergindo de fungos de Marasmiaceae e Polyporaceae e nós encontramos as duas espécies do gênero sobrevoando ou repousando sobre os fungos, confirmando assim a associação de *Paraliodrosophila* com os corpos de frutificação de fungos macroscópicos como previamente sugerido VILELA E BÄCHLI (2007). A modelagem de nicho ecológico demonstrou que a temperatura máxima do mês mais quente é um importante fator limitante da distribuição de *P. antennata*, com a espécie sendo melhor adaptada a temperaturas mínimas intermediárias. Também é possível observar através do modelo gerado que esta espécie apresenta uma distribuição potencial ampla, ocupando praticamente toda a área da Mata Atlântica e parte da Amazônia, com registros já observados nestes dois biomas. Mesmo com esta ampla distribuição, nossos dados indicam que o fluxo gênico entre as populações é mantido ao longo de sua área de ocorrência, uma vez que as análises de *Neighbor Joining* (NJ) e *Generalized Mixed Yule Coalescent* (GMYC) consideram todos os espécimes analisados (originários dos estados de RO, SP, PR, RS) como uma única unidade evolutiva. Já no caso de *P. burlai*, a modelagem de nicho não foi realizada devido aos poucos registros de ocorrência da espécie (em apenas duas localidades nos estados do RJ e SP) quando o Capítulo 2 foi elaborado. Em nosso trabalho

expandimos a ocorrência dessa espécie para a região sul do Brasil, uma vez que ela foi ainda coletada em Florianópolis/SC.

A principal referência para as espécies Neotropicais do gênero cosmopolita *Mycodrosophila* (MCEVEY E POLAK, 2005; THROCKMORTON, 1975; WHEELER E TAKADA, 1963) é a revisão das espécies americanas de WHEELER E TAKADA (1963), com o relato da ocorrência de apenas seis espécies nessa região (*M. elegans*, *M. brunnescens*, *M. projectans*, *M. pseudoprojectans*, *M. neoprojectans* e *M. nigropleura*). Destas nós coletamos apenas *M. elegans* e *M. projectans*. Entretanto, foram encontradas evidências da existência de mais oito potenciais espécies novas (*Mycodrosophila* M001, *Mycodrosophila* affinis M001, *M. neoprojectans* affinis 1, *M. neoprojectans* affinis 2, *M. projectans* affinis 1 e *M. projectans* affinis 2, *Mycodrosophila* SFP, *Mycodrosophila* sp. 1), muitas das quais são bastante frequentes na região amostrada. Três destas potenciais espécies novas foram representadas unicamente por fêmeas (*M. projectans* affinis 1 no Capítulo 3, *M.* SFP e *Mycodrosophila* sp. 1 no Capítulo 4), mas seu *status* como espécie foi confirmado com marcadores moleculares, e corroborado no caso da primeira espécie, após a inserção de indivíduos machos no Capítulo 4.

Mycodrosophila elegans foi uma espécie pouco abundante nas amostragens realizadas. Ainda que apenas fêmeas tenham sido incluídas nas análises do Capítulo 3, machos também foram coletados, embora não tenha sido possível realizar a amplificação do marcador COI destes espécimes. Na verdade, nós percebemos que os drosofilídeos micofílicos com os quais trabalhamos apresentam problemas de amplificação até mesmo para genes mitocondriais, caso que parece ser mais marcante nas espécies de *Mycodrosophila*. Nós acreditamos que esse problema ocorra devido a co-extração de resíduos dos fungos ingeridos pelos espécimes, que possivelmente atuam inibindo as reações de polimerização em cadeia (PCR). Nós observamos esse resultado com extrações realizadas com diferentes métodos, e com kits fornecidos por diferentes empresas. Apesar de não termos realizado análises específicas, quantificações realizadas no aparelho Biodrop demonstram a impureza das extrações, com presença de contaminantes nas amostras, conforme detectado pela razão A260/A230.

De acordo com o modelo de distribuição potencial gerado para *M. elegans*, esta espécie é amplamente influenciada pelos extremos de temperatura, já que os fatores mais limitantes para sua distribuição são a temperatura mínima do mês mais

frio e a temperatura máxima do mês mais quente. Segundo o modelo, a espécie possui uma ampla distribuição no Neotrópico, com registros na América Central e na região amazônica da América do Sul, e distribuição potencial para o Pantanal e para parte da Mata Atlântica e do Pampa brasileiros. De fato, no Brasil ela parece estar restrita principalmente à região oeste, não adentrando muito na porção mais meridional da Mata Atlântica, exceto pela região de Foz do Iguaçu/PR (dados não mostrados).

Já para *M. projectans*, a outra espécie descrita de *Mycodrosophila* amostrada em nossas coletas, os registros disponíveis indicavam uma ampla distribuição ao longo da maior parte da região Neotropical (BÄCHLI, 2015) (Capítulo 3). Contudo, nossos dados indicam que o que vem sendo tratado como *M. projectans* representa, na verdade, um complexo de espécies muito similares em nível de morfologia externa, que não apresentam diferenciação no padrão de coloração e de morfometria de asas. Além disso, a genitália masculina não é capaz de fornecer uma distinção confiável entre as espécies, uma vez que apesar de *M. projectans* e *M. projectans* affinis 1 possuirem o edeago distintos, *M. projectans* affinis 2 apresenta indivíduos identificados com ambos os padrões. De fato, *M. projectans* affinis 2 apresenta a distribuição mais ampla, sendo registrada desde o Pampa até a Amazônia e sendo capaz de manter o fluxo gênico entre essas regiões uma vez apenas dez mutações separam os haplótipos mais distantes dos extremos de distribuição. Já *M. projectans* parece apresentar uma estruturação geográfica relacionada aos biomas Mata Atlântica e Pampa, conforme observado na rede de haplótipos com os dados do marcador molecular COI e *M. projectans* affinis 1 parece restrita à Mata Atlântica. Apesar dessas diferenças de distribuição, as três espécies apresentam um alto grau de simpatria, e até mesmo de sintopia. Esse padrão de distribuição parece ser comum para *Mycodrosophila*, uma vez que LACY (1982) observou a mesma situação com *M. claytone*. Este autor analisou 18 loci de alozimas e foi capaz de distinguir duas espécies crípticas, que não compartilhavam nenhum alelo em seis loci, mas coletadas sobre o mesmo fungo ao mesmo tempo.

Outra observação interessante quanto ao complexo *M. projectans*, diz respeito à abundância de espécimes coletados. Nas coletas realizadas na região leste do país, poucos indivíduos de cada espécie foram capturados, embora em coletas realizadas na região mais a oeste, a abundância tenha sido maior, com dezenas de espécimes coletados na região de Santa Maria/RS e centenas no

Parque Estadual do Turvo/RS e no Parque Nacional de Foz do Iguaçu/PR (dados não mostrados). Nós acreditamos que esse padrão possa estar relacionado a dois fatores não excludentes: (i) estocasticidade ambiental, explicada pela efemeridade dos fungos (COURTNEY *et al.*, 1990), que faz com que as populações se desloquem continuamente em busca desses recursos. Esse aspecto é parcialmente suportado pelos registros observados na região de Santa Maria/RS, onde coletas realizadas em diferentes momentos não encontraram níveis similares de abundância (dados não mostrados). Contudo, essa hipótese precisa ser testada através de coletas sistemáticas, num delineamento experimental voltado a esse fim; (ii) adaptação longitudinal às regiões do Pantanal, Amazônia e oeste da Mata Atlântica, com poucas populações/indivíduos chegando a região leste do Brasil, já que nosso esforço amostral foi concentrado na região leste da Mata Atlântica, com poucas coletas sendo realizadas na região oeste. Como os modelos gerados não incorporam dados de abundância das espécies coletadas, apenas sua presença em determinado local, podemos estar inferindo inadequadamente os fatores abióticos restritivos a distribuição das espécies. Dessa forma, para testar essa hipótese, novas coletas nessas regiões precisam ser realizadas, assim como novos modelos que diferenciem dados de abundância precisam ser gerados.

Com relação às outras potenciais espécies novas de *Mycodrosophila*, uma (*Mycodrosophila* M001) foi identificada através da morfologia externa e cinco através das análises moleculares (*Mycodrosophila* affinis M001, *Mycodrosophila neoprojectans* affinis 1 e 2, *Mycodrosophila* SFP, *Mycodrosophila* sp. 1). Neste caso, é interessante destacar que os mesmos padrões de simpatria e sintopia observados para o complexo *M. projectans* e para *M. claytone* (LACY, 1982) também são encontrados entre as espécies crípticas *M. neoprojectans* affins 1 e 2, embora as mesmas apresentem 5.5% de divergência para a região analisada do gene COI. Dentre as espécies novas de *Mycodrosophila*, cabe ressaltar ainda que *M.* M001 e *M. neoprojectans* affinis 1 encontraram-se em processo final de descrição por nosso grupo de pesquisa.

No que diz respeito a diversidade do marcador molecular COI, todas as espécies apresentaram valores de diversidade nucleotídica e haplotípica consideravelmente maior que outras espécies de drosofilídeos do gênero *Drosophila* (DE RÉ *et al.*, 2014). Dentre as espécies de *Mycodrosophila*, *M. projectans* foi a que apresentou os índices maiores, exceto para diversidade haplotípica onde *M.*

neoprojectans 1 e 2 apresentaram índice de 1. Já o macador molecular COII apresentou índices ainda maiores que o COI para o complexo *M. projectans*.

Diferentemente de *Mycodrosophila, Zygothrica* é essencialmente Neotropical (GRIMALDI, 1987; GRIMALDI E FENSTER, 1989; THROCKMORTON, 1975), sendo considerado um dos gêneros mais especiosos de Drosophilidae (GRIMALDI, 1987), e, dentre os táxons do grupo genérico *Zygothrica*, o mais estudado na sua região. Em nossa amostragem, encontramos 20% das 54 espécies descritas para o Brasil (GOTTSCHALK *et al.*, 2008), além de 13 potenciais espécies novas. Destas, entretanto, duas foram representadas apenas por fêmeas (*Z. atriangula* affinis H015 e *Z. poeyi* affinis), podendo pertencer a espécies com descrição exclusiva para machos. Dentre essas potenciais espécies novas, *Z. vittimaculosa* affinis já foi confirmada como uma espécie nova por outras metodologias e está sendo descrita por nosso grupo de pesquisa.

Zygothrica foi, junto de *Hirtodrosophila*, um dos gêneros que não apresentou um intervalo entre as distâncias intraespecíficas e interespecíficas (*Barcode gap*; MEYER E PAULAY, 2005). Isto ocorreu devido à alta distância intraespecífica encontrada em *Z. ptilialis* (5,3%) e à baixa divergência interespecífica registrada entre *Z. vittimaculosa* e *Z. vittimaculosa* affinis. No caso de *Z. ptilialis*, a alta divergência pode estar relacionada a padrões de estruturação geográfica na espécie ou a presença de diversidade críptica, não identificada devido ao número de indivíduos incorporados na análise. Através de uma análise filogeográfica será possível distinguir entre esses dois cenários. No caso de *Z. vittimaculosa* e *Z. vittimaculosa* affinis, os baixos valores de distância interespecífica devem estar relacionados à especiação recente, estimada através de um relógio molecular relaxado aplicado para sequências do gene AMD em 3,14 milhões de anos (dados não mostrados). Embora os padrões de diversidade destas espécies sejam também compatíveis com uma estruturação intraespecífica, análises morfológicas em nível de edeago e morfometria das asas suportam a existência das duas espécies-irmãs (FONSECA *et al.*, em preparação).

Das 11 espécies descritas de *Zygothrica* incorporadas na análise de DNA *barcode* no Capítulo 3, seis tiveram seus modelos de distribuição potencial gerados e analisados no Capítulo 2. Entretanto, como também observado em *M. projectans*, alguns dos modelos se mostraram posteriormente inadequados, devido a modificações no *status* das espécies realizadas após a descoberta de níveis

crípticos de diversidade. Esse é o caso dos espécimes identificados inicialmente como *Z. prodispar*, que foram separados em dois clados independentes pelo DNA *barcode*, revelando a presença de uma espécie críptica a nível de genitália (*Z. prodispar* affinis 2), além de uma espécie já identificada *a priori* como distinta (*Z. prodispar* affinis 1), que não foi incorporada na modelagem. *Zygothrica poeyi* também foi subdividida em duas espécies crípticas, *Z. laevifrons* 1 e 2. A mudança no uso do epíteto específico nesse caso ocorreu devido a observação de que *Z. laevifrons* é sinônimo sênior de *Z. poeyi* senso BURLA (1956) (GRIMALDI, 1990a), mas devido à inexistência da descrição da genitália masculina de *Z. poeyi* não foi possível determinar a identidade das espécies. Mais uma vez, uma terceira espécie, *Z. poeyi* affinis (representada por uma única fêmea), já havia sido identificada *a priori* como morfologicamente distinta, de forma que não foi incorporada na análise de modelagem. Tanto no caso de *Z. prodispar*, quando para *Z. laevifrons* (=*Z. poeyi*) análises morfológicas adicionais serão necessárias para a descrição e redescrição das espécies. Assim como para *Z. poeyi* affinis, *Z. atriangula* affinis H015 também foi representada por uma única fêmea em nossas amostragens e já havia sido identificada *a priori* como uma espécie distinta de *Z. atriangula*. Os espécimes identificados como *Z. zygopoeyi* nas análises de modelagem tiveram sua identidade revista, e no Capítulo 3 foram tratados como *Z. parapoeyi* affinis.

Apesar de apresentarem diferentes amplitudes de distribuição potencial, *Z. hypandriata*, *Z. orbitalis* e *Z. ptilialis* parecem ocorrer exclusivamente na região sul do continente americano, de forma que a temperatura mínima do mês mais frio e a precipitação do trimestre mais frio constituem os fatores limitantes em sua distribuição. Já *Z. bilineata* e *Z. dispar* parecem ter uma distribuição potencial bem ampla ao longo do continente, com a precipitação do trimestre mais úmido sendo determinante em ambos os casos.

Devido ao baixo número de pontos de ocorrência disponíveis na literatura e em nossas coletas, durante a elaboração do Capítulo 2, *Z. atriangula*, *Z. parapoeyi*, *Z. parvipoeyi*, *Z. virgatinigra*, *Z. vittinubila* e *Z. zygopoeyi* (=*Z.parapoeyi* affinis) apresentaram baixa performance nos modelos gerados, de forma que os mesmos não foram considerados. Por outro lado, *Z. bilineata* não foi incluída nas análises de DNA *barcode* (Capítulo 3) porque não foi possível amplificar o marcador COI para o espécime coletado e a mesma não foi reamostrada.

Com relação ao gênero cosmopolita *Hirtodrosophila*, encontramos, em nossas coletas, 31,25% (BÄCHLI, 2015; GOTTSCHALK *et al.*, 2008) das 16 espécies descritas para o Brasil (GRIMALDI, 2010; VILELA E BÄCHLI, 2004), além de 15 potenciais espécies novas. Dentre estas últimas, seis foram representadas apenas por fêmeas em nossas amostragens, sendo que quatro apresentaram características de morfologia externa distintas de espécies já descritas em análises realizadas *a priori*. Esse grande número de espécies potencialmente novas enfatiza um aspecto interessante: apesar de que o número de espécies descritas para a região Neotropical e para o Brasil (BÄCHLI, 2015; GOTTSCHALK *et al.*, 2008; VILELA E BÄCHLI, 2004) seja consideravelmente maior em *Zygothrica* do que em *Hirtodrosophila*, em nossas amostragens encontramos um número similar de espécies potencialmente novas para ambos os gêneros. Isto sugere que possivelmente o gênero *Hirtodrosophila* pode ser tão especioso quanto o gênero *Zygothrica* na região Neotropical, sem dúvida uma propriedade notável para dois gêneros irmãos (GRIMALDI, 1990a) que constituem grande parte do grupo genérico *Zygothrica*. A importância desta hipótese pode ser compreendida principalmente se considerarmos as evidências prévias de que *Hirtodrosophila* não seja monofilético (DA LAGE *et al.*, 2007; RUSSO *et al.*, 2013; VAN DER LINDE *et al.*, 2010; YASSIN, 2013), uma vez que os membros Neotropicais do gênero agrupam com outros gêneros representantes do grupo genérico *Zygothrica* como *Mycodrosophila* (DA LAGE *et al.*, 2007; RUSSO *et al.*, 2013; VAN DER LINDE *et al.*, 2010) e *Zygothrica* (Capítulo 4), enquanto os seus representantes não-Neotropicais agrupam independentemente (DA LAGE *et al.*, 2007; RUSSO *et al.*, 2013; VAN DER LINDE *et al.*, 2010).

Assim como em *Zygothrica*, o *Barcode gap* também não foi detectado em *Hirtodrosophila*. Isto ocorreu, principalmente, devido à alta divergência intraespecífica entre alguns espécimes de *H. gilva* (máximo de 5,1%) e à baixa divergência interespecífica entre *H. subflavohalterata* affinis 2 e *H. subflavohalterata* H002 (mínimo de 2,4%), entre *H. mendeli* affinis com *H. subflavohalterata* affinis 2 (mínimo de 2,7%) e com *H. subflavohalterata* affinis H002 (mínimo de 2,6%). No caso de *H. gilva*, novamente, esses padrões podem estar relacionados a padrões filogeográficos ou à presença de espécies crípticas não identificadas. Nos demais casos, os baixos níveis de divergência são compatíveis com especiação recente, já

que diferenças morfológicas tanto na morfologia externa como na genitália masculina foram detectadas *a priori* entre as diferentes espécies.

Hirtodrosophila levigata mostrou possuir uma espécie críptica, representada por uma única fêmea distinguida (*H. levigata* affinis 1). Contudo, como essa potencial espécie nova foi coletada no mesmo ponto em que outros indivíduos de *H. levigata*, o modelo gerado para esta espécie no Capítulo 2 permanece válido. Esse também é o caso de *H. mendeli* e *H. mendeli* affinis, exceto pelo fato de que a última foi identificada como potencial espécie nova através da morfologia das genitálias masculinas. Já no caso de *H. morgani*, foi identificada uma espécie distinguível pela genitália masculina (*H. morgani* affinis sul), que possivelmente apresenta pontos de simpatria com *H. morgani*, não alterando o modelo previamente gerado. Para as três espécies modeladas, a temperatura mínima do mês mais frio foi o fator limitante da distribuição. *Hirtodrosophila levigata* apresenta uma distribuição potencial mais restrita a região sul do continente americano, enquanto *H. morgani* apresenta uma distribuição potencial sobreposta a primeira, mas mais ampla, chegando ao sul da Amazônia; já *H. mendeli* apresenta uma amplitude de distribuição intermediária entre estes dois extremos.

Devido à baixa performance dos modelos, *H. gilva*, *H. subflavohalterata* e *H. subgilva* tiveram seus modelos de distribuição potencial omitidos. Contudo, cabe ressaltar, que *H. subgilva* foi posteriormente subdividida em duas espécies similares na morfologia externa, mas que são distinguíveis na morfologia de genitália masculina (*H. subgilva* affinis 1 e 2), além da espécie já descrita, também amostrada em nossas coletas. *Hirtodrosophila subflavohalterata*, por outro lado, se mostrou subdividida em seis espécies crípticas, sendo quatro distinguíveis através de genitália masculina e/ou padrão de morfologia externa (*H. subflavohalterata* affinis 1, 2, 3, H002) e duas indistinguíveis em qualquer destes níveis (*H. subflavohalterata* affinis 4 e 5, sendo a primeira representada apenas por fêmeas). Neste caso, entretanto, a espécie descrita, *H. subflavohalterata*, não foi amostrada.

No que diz respeito a aspectos filogenéticos do grupo genérico *Zygothrica*, GRIMALDI (1990a) realizou uma ampla revisão taxonômica, com cerca de 120 espécies pertencentes à maioria dos gêneros e subgêneros de Drosophilidae. Baseado na análise de parcimônia de 217 caracteres morfológicos, o autor afirma que, considerando os caracteres diagnósticos, “não há dúvida de que *Hirtodrosophila*, *Zygothrica*, *Mycodrosophila*, *Paramycodrosophila* e

Paraliodrosophila formam um grupo monofilético". Análises moleculares realizadas posteriormente (DA LAGE *et al.*, 2007; ROBE *et al.*, 2005; RUSSO *et al.*, 2013; VAN DER LINDE *et al.*, 2010) tendem a incorporar poucos representes do grupo genérico, mas mesmo assim, já há evidências refutando a análise morfológica de GRIMALDI (1990a), principalmente devido a polifilia de *Hirtodrosophila* (DA LAGE *et al.*, 2007; RUSSO *et al.*, 2013; VAN DER LINDE *et al.*, 2010; YASSIN, 2013).

Nossos dados, gerados com base em sequências do marcador nuclear AMD isoladas para nove espécies não suportam a monofilia do grupo genérico *Zygothrica*, apesar de não possuir suporte para refutá-la. *Mycodrosophila elegans*, *M. projectans* e uma espécie nova de *Mycodrosophila* registrada no município de São Francisco de Paula, formam um clado que confirma a monofilia do gênero. Por outro lado, embora espécies de *Hirtodrosophila* e *Zygothrica* formem um clado monofilético, as mesmas encontram-se misturadas no clado recuperado. Dessa forma, o clado formado por *H. levigata* encontra-se em polifilia com as espécies de *Zygothrica* e esse clado é irmão de outro mais basal formado por duas espécies de *Hirtodrosophila* (*H. morgani* e *H. pictiventris*). Através da incorporação de mais espécies e mais marcadores moleculares, trabalho que já vem sendo realizado por nosso grupo de pesquisa, será possível compreender melhor o cenário relacionado à evolução desse grupo de espécies.

Outro aspecto interessante relacionado à ecologia dos drosofilídeos micofílicos estudados é o seu comportamento gregário. Em nossas coletas, da mesma forma que em relatos anteriores (THROCKMORTON, 1975), observamos a coexistência de várias espécies sobre os mesmos fungos, que são, muitas vezes, colonizados por dezenas a centenas de indivíduos. Este padrão não foi, entretanto, observado, para espécimes de *Mycodrosophila* colonizando fungos do gênero *Ganoderma*, embora mais análises sejam necessárias para confirmar essa observação. O comportamento gregário, além da polifagia na utilização dos fungos, se deve, possivelmente, ao caráter efêmero e imprevisível dos recursos (COURTNEY *et al.*, 1990; JAENIKE,1978b).

THROCKMORTON (1975) sugeriu que a micofagia surgiu diversas vezes em Drosophilidae, de forma que, nesse cenário, o esperado seria que a divergência de nicho fosse o principal fator a explicar a evolução desses grupos (POWELL, 1997; ROBE *et al.*, 2010). Este padrão é, também, sugerido pela noção de THROCKMORTON (1975) de que a evolução da família teria se dado a partir de

rápidas e sucessivas radiações. Dentro do grupo genérico *Zygothrica*, entretanto, a avaliação dos padrões observados de sobreposição de nicho abiótico e a comparação das curvas de resposta individuais sugere a existência de fortes padrões de conservadorismo de nicho (Capítulo 2). De fato, dentre as 13 espécies avaliadas, apenas *Z. orbitalis* e *Z. prodispar* parecem apresentar sinais de divergência de nicho. A possibilidade de não-monofilia do grupo, discutida acima, traz, entretanto, novas interpretações a estes resultados.

Da mesma forma que no grupo genérico *Zygothrica*, observamos no Capítulo 1 que o padrão de conservadorismo de nicho parece explicar a evolução da linhagem *virilis-repleta* e do subgênero *Drosophila* como um todo. Contudo, a linhagem *virilis-repleta*, de fato, parece possuir características de uma radiação adaptativa, uma vez que sinais de divergência de nicho foram recuperados para esse grupo. Apesar do baixo número de espécies incorporadas (sete), as análises que suportam esses resultados são mais robustas que as utilizadas no grupo genérico *Zygothrica*, uma vez que a maioria das espécies do subgênero *Drosophila* utilizadas no trabalho possuem amostragens mais amplas que possibilitam gerar modelos de distribuição potencial mais acurados com relação às restrições abióticas das espécies, de forma a representar melhor seus nichos fundamentais. Além disso, aspectos bióticos relacionados ao uso de diferentes recursos vegetais (frutos, flores e fungos) forneceram um outro conjunto de dados, cujos padrões foram semelhantes àqueles recuperados para os dados climáticos. De modo geral, o conjunto de dados apresentados nos Capítulos 1 e 2, demostram que a evolução de Drosophilidae é mais complexa do que inicialmente hipotetizado por THROCKMORTON (1975), de forma que cada linhagem da família precisará ser individualmente analisada, de modo a fornecer *insights* quanto ao papel das radiações na evolução da família e definir se elas são de fato adaptativas ou meramente taxonômicas.

5. CONCLUSÕES

Com base nos resultados obtidos no presente estudo, pode-se concluir que:

- A hipótese de radiações adaptativas proposto por THROCKMORTON (1975) para explicar a evolução da família Drosophilidae não pode ser aplicada a toda família e deve ser avaliada para cada grupo, uma vez que encontramos evidências de divergência de nicho, que é compatível com o modo de evolução de uma radiação adaptativa, apenas na linhagem *tripunctata*.

- A modelagem de nicho de espécies do grupo genérico *Zygothrica* confirmou, tanto pelos fatores abióticos como pela análise de conservadorismo de nicho, o potencial da utilização desse grupo como bioindicador para avaliar mudanças climáticas.

- A metodologia do DNA *barcode* mostrou-se altamente eficiente na identificação de espécimes do grupo genérico *Zygothrica*, tanto pela concordância com a identificação taxonômica *a priori*, como por ressaltar a diversidade críptica não percebida pela taxonomia clássica.

- O gênero *Mycodrosophila* demostrou possuir uma alta variabilidade genética, além de possuir a capacidade de manter fluxo gênico a longas distâncias. Outro aspecto interessante é a recorrente simpatria e sintopia entre espécies irmãs crípticas, o que parece ser uma característica recorrente nesse gênero.

- *Mycodrosophila projectans* demonstrou ser um complexo de espécies de difícil diferenciação morfológica, sendo necessária a inclusão de marcadores moleculares para identificar a diversidade desse complexo. Esse cenário leva a questionamentos quanto à amplitude da distribuição atribuída para outras espécies micofílicas do grupo genérico *Zygothrica*, como *Z. dispar* e *Z. prodispar,* que também possuem, *a priori*, uma ampla distribuição pelos Neotrópicos.

6. PERSPECTIVAS

Em vista do conjunto de informações apresentados, pode-se resumir da seguinte forma as principais perspectivas relacionadas a presente tese:

- Descrição morfológica das potenciais espécies novas sugeridas pelo DNA *barcode;*

- Descrição das genitálias das fêmeas, uma vez que através do DNA *barcode* foi possível identificar a qual espécie de machos elas estão relacionadas;

- Relacionar as espécies de fungos onde os indivíduos foram coletados com as espécies identificadas, a fim de analisar a presença de generalismo ou especialização no uso de recursos.

- Realizar amostragens na região oeste dos biomas brasileiros (Pantanal e Amazônia), para testar a hipótese de que o complexo *M. projectans* apresenta distribuição e maior abundância nessa região.

- Testar a presença de conservadorismo de nicho no grupo genérico *Zygothrica* através de técnicas mais robustas, uma vez que suas relações filogenéticas, sua distribuição real e os seus padrões de utilização de recursos tenham sido melhor elucidados.

- Avaliar os padrões e processos relacionados à especiação dentro do complexo de espécies *M. projectans*.

- Estudar as oscilações demográficas sofridas por diferentes espécies micofílicas ao longo do tempo evolutivo, de forma a avaliar sua aplicação como bioindicador no estudo das alterações climáticas do Pleistoceno.

REFERÊNCIAS BIBLIOGRÁFICAS

BÄCHLI, G. TaxoDros: The database on taxonomy of Drosophilidae. Version 1.04. 2015. Available from: http://www.taxodros.uzh.ch/.

BARBOSA, F. G.; SCHNECK, F.; MELO, A. S. Use of ecological niche models to predict the distribution of invasive species: a scientometric analysis. **Braz. J. Biol.** 72(4):821-829, 2012.

BRYAN JR., E. H. Key of the Hawaiian Drosophilidae and drescription of new species. **Proc. Haw. Ent. Soc.** 1:25-42, 1938.

BURLA, H. Die Drosophilidengattung *Zygothrica* und ihre Beziehung zur *Drosophila*-Untergattung *Hirtodrosophila*. **Mitt. Mus. Naturkunde Berl., Zoolog. Reihe** 32(2):189–321, 1956.

BURLA, H.; BÄCHLI, G.; HUBER, H. *Drosophila* reared from the stinkhorn, *Phallus impudicus*, near Zurich, Switzerland. **J. Zoolog. Syst. Evol. Res.** 29:97-107, 1991.

BUTCHART, S.H.M. *et al*. Global biodiversity: indicators of recent declines. **Science** 328:1164-1168, 2010.

CBOL PLANT WORKING GROUP. A DNA barcode for land plants**. Proc. Natl. Acad. Sci. U.S.A.** 106:12794–12797, 2009.

CONSORTIUM FOR THE BARCODE OF LIFE. CBOL approves matK and rbcL as the BARCODE regions for Land Plants, 2009.

CLARE, E. L. *et al*. Neotropical Bats: estimating species diversity with DNA Barcodes. **PLoS ONE** 6(7):e22648, 2011.

COURTNEY, S. P.; KIBOTA, T. T.; SINGLETON, T. A. Ecology of Mushroom-feeding Drosophilidae. **Adv. Ecol. Res.** 20:225–274, 1990.

CRAWFORD, A. J. *et* al. DNA barcoding applied to ex situ tropical amphibian conservation programme reveals cryptic diversity in captive populations. **Mol. Ecol. Resour.** 13(6):1005-1018, 2012.

DA LAGE, J. L. *et al.* A phylogeny of Drosophilidae using the Amyrel gene: questioning the *Drosophila melanogaster* species group boundaries**. J. Zoolog. Syst. Evol. Res.** 45(1):47–63, 2007.

DA MATA *et al.* Invasive and flexible: niche shift in the *drosophilid Zaprionus indianus* (Insecta, Diptera). **Biol. Invasions** 12(5):1231-1241, 2010.

DAYRAT, B. Towards integrative taxonomy. **Biol. J. Linn. Soc.** 85:407-415, 2005.

DEAGLE, B. E.; TOLLIT, D. J. Quantitative analysis of prey DNA in pinniped faeces: potential to estimate diet composition? **Conserv. Genet.** 8:743–747, 2007.

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. **Biol. J. Linn. Soc. Lond.** 112:55-66, 2014.

DOGE, J. S.; VALENTE, V. L. S.; HOFMANN, P. R. P. Drosophilids (Diptera) from an Atlantic Forest Area in Santa Catarina, Southern Brazil. **Rev. Bras. Entomol.** 52(4): 615-624, 2008.

DURANDO, C. M. *et al.* Phylogenetic analysis of the repleta species group of the genus *Drosophila* using multiple sources of characters. **Mol. Phylogenet. Evol.** 16:296-307, 2000.

ELITH, J.; LEATHWICK, J. R. Species distribution models: ecological explanation and prediction across space and time. **Annu. Rev. Ecol. Evol. Syst**. 40:677-697, 2009.

FROTA-PESSOA, O. Sobre o subgênero *Hirtodrosophila*, com descrição de uma nova espécie (Diptera, Drosophilidae, *Drosophila*). **Rev. Bras. Entomol.** 5(4):469-483, 1945.

GARCIA, C. F. *et al*. Drosophilid assemblages at different urbanization levels in the city of Porto Alegre, State of Rio Grande do Sul, Southern Brazil. **Neotrop. Entomol.** 40(1):32-41, 2012.

GIAM, X. *et al*. Reservoirs of richness: least disturbed tropical forests are centres of undescribed species diversity. **Proc. R. Soc. Lond. B. Biol. Sci.** 279:67–76, 2012.

GLENNON, K. L.; RISSLER, L. J.; CHURCH, S.A. Ecogeographic isolation: a reproductive barrier between species and between cytotypes in *Houstonia* (Rubiaceae). **Evol. Ecol.** 26:909-926, 2012.

GONZALEZ-VARO, J. P.; ARROYO, J. M.; JORDANO, P. Who dispersed the seeds? The use of DNA barcoding in frugivory and seed dispersal studies. **Methods Ecol. Evol.** 5:806-814, 2014.

GOTTSCHALK, M. S. *et al*. Changes in Brazilian Drosophilidae (Diptera) assemblages across an urbanisation gradient. **Neotrop. Entomol.** 36(6):848-862, 2007.

GOTTSCHALK, M. S. Utilização de recursos tróficos por espécies Neotropicais de Drosophilidae (Diptera). D. Phil Thesis, Universidade Federal do Rio Grande do Sul, 2008.

GOTTSCHALK, M. S.; HOFMANN, P. R. P.; VALENTE, V. L. S. Diptera, Drosophilidae: historical occurrence in Brazil. **Check List** 4:485-518, 2008.

GOTTSCHALK, M. S. *et al*. Drosophilidae (Diptera) associated to fungi: differential use of resources in anthropic and Atlantic Rain Forest areas. **Iheringia Ser. Zool.** 99:442-448, 2009.

GRAHAM, C. H. *et al*. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. **Evolution** 58(8):1781-1793, 2004.

GRIMALDI, D. A.; JAENIKE, J. Competition in natural populations of mycophagous *Drosophila*. **Ecology** 65(4):1113-1120, 1984.

GRIMADI, D. Niche separation and competitive coexistence in mycophagous *Drosophila* (Diptera: Drosophilidae). **Proc. Entomol. Soc. Wash.** 87:498-511, 1985.

GRIMALDI, D. Phylogenetics and taxonomy of *Zygothrica* (Diptera: Drosophilidae). **Bull. Am. Mus. Nat. Hist.** 186:103-268, 1987.

GRIMALDI, D.; FENSTER, G. Evolution of extreme sexual dimorphisms: structural and behaviour convergence among broad-headed male Drosophilidae (Diptera). **Am. Mus. Novit.** 2939:1-25, 1989.

GRIMALDI, D. A. A phylogenetic, revised classification of the genera in the Drosophilidae (Diptera). **Bull. Am. Mus. Nat. Hist.** 197:1-13, 1990a.

GRIMALDI, D. A. Revision of *Zygothrica* (Diptera: Drosophilidae), Part II. The first African species, two New Indo-Pacific groups, and the *bilineata* and *samoaensis* species groups. **Am. Mus. Novit.** 2964:1–31, 1990b.

GRIMALDI, D. Description and immature stages of *Hirtodrosophila batracida* sp. n. (Diptera: Drosophilidae), a predator of frog embryos. **Insect Syst. Evol.** 25(2):129-136, 1994.

GRIMALDI, D. Drosophilidae (Small fruit flies, pomace flies, vinegar flies) IN: BROWN, B. V. *et al.* (eds.) Manual of Central American Diptera, Volume 2, 2010. xvi + 728 pp. NRC Research Press, Ottawa, Ontario, Canada.

GUISAN, A. *et al*. Using niche-based models to improve the sampling of rare species. **Conserv. Biol.** 20(2):501–511, 2006.

HAJIBABAEI, M.; SINGER, G. A.; HICKEY, D. A. Benchmarking DNA barcodes: An assessment using available primate sequences. **Genome** 49(7):851-854, 2006.

HATADANI, L. M. *et al.* Molecular phylogeny of the *Drosophila tripunctata* and closely related species groups (Diptera: Drosophilidae). **Mol. Phylogenet. Evol.** 51:595-600, 2009.

HEBERT, P. D. N. *et al*. Biological identification through DNA Barcodes. **Proc. R. Soc. Lond. B. Biol. Sci.** 270:313-321, 2003a.

HEBERT, P. D. N.; RATNASINGHAM, S.; DEWAARD, J. R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. **Proc. R. Soc. Lond. B. Biol. Sci.** 270:S96-S99, 2003b.

HEBERT, P. D. N. *et al*. Ten species in one: DNA Barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. **Proc. Natl. Acad. Sci. U.S.A.** 101(41):14812-14817, 2004.

HOLMES, B. H.; STEINKEB, D.; WARDA, R. D. Identification of shark and ray fins using DNA barcoding. **Fish. Res.** 95(2-3):280-288, 2009.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. Climate change 2007: synthesis report. Geneva (Switzerland): Intergovernmental Panel on Climate Change, 2007.

JAENIKE, J. Host selection by mycophagous *Drosophila*. **Ecology** 59(6):1286-1288, 1978a.

JAENIKE, J. Resource predictability and niche breadth in the *Drosophila quinaria* species group. **Evolution** 32(3):676-678, 1978b.

JAENIKE, J.; JAMES, A. C. Aggregation and the coexistence of mycophagous *Drosophila*. **J. anim. Ecol.** 60(3):913-928, 1991.

JAENIKE, J. Interactions between mycophagous *Drosophila* and their nematode parasites: from physiological to community ecology. **Oikos** 72(2):235-244, 1995.

JENNINGS, B. H. *Drosophila* - a versatile model in biology & medicine. **Materialstoday** 14(5):190-195, 2011.

JINBO, U.; KATO, T.; ITO, M. Current progress in DNA Barcoding and future implications for entomology. **Entomol. Sci.** 14:107-124, 2011.

JOLY, S. *et al*. Ecology in the age of DNA Barcoding: the resource, the promise and the challenges ahead. **Mol. Ecol. Resour.** 14(2):221-232, 2014.

JOO, S.; PARK, S. Identification of bird species and their prey using DNA Barcode on feces from Korean traditional village groves and forests (maeulsoop). **Animal Cells and Syst.** 16(6):488-497, 2012.

JUN, J. *et al*. Wildlife forensics using mitochondrial DNA sequences: Species identification based on hairs collected in the field and confiscated tanned Felidae leathers. **Genes Genomics** 33(6):721-726, 2011.

KASUYA, N. *et al.* Diversity and host associations of parasitoids attacking mycophagous drosophilids (Diptera: Drosophilidae) in northern and central Japan. **Entomol. Sci.** 16:227–234, 2013.

KELLERMANN, V. *et al.* Phylogenetic constraints in key functional traits behind species' climate niches: patterns of desiccation and cold resistance across 95 *Drosophila* species. **Evolution** 66:3377–3389, 2012a.

KELLERMANN, V. *et al.* Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically**. Proc. Natl. Acad. Sci. U.S.A.** 109(40):16228-16233, 2012b.

KIMURA, M. K. *et al*. Evidence for cryptic northern refugia in the last glacial period in *Cryptomeria japonica*. **Ann. Bot.** 114(8):1687-1700, 2014.

KOZAK, K. H.; WIENS, J. J. Does niche conservatism promotes speciation: a case study in north American salamanders. **Evolution** 60(12):2604-2621, 2006.

KRESS, W. J. *et al*. DNA barcodes for ecology, evolution, and conservation. **Trends Ecol. Evol.** 30(1):25-35, 2015.

KWIATOWSKI, J.; AYALA, F. J. Phylogeny of *Drosophila* and related genera: conflict between molecular and anatomical analyses. **Mol. Phylogenet. Evol.** 13(2):319-328, 1999.

LACY, R. C. Niche breadth and abundance as determinants of genetic variation in populations of mycophagous drosophilid flies (Diptera: Drosophilidae). **Evolution** 36(6):1265-1275, 1982.

LANDIM, M. I. P. F.; HINGST-ZAHER, E. Brazil's biodiversity crisis. **ICOM News** 63:4-15, 2010.

LEWINSOHN, T. M.; PRADO, P. I. How many species are there in Brazil? **Conserv. Biol.** 19:619-624, 2005.

LUO, A. *et al.* Potential efficacy of mitochondrial genes for animal DNA Barcoding: a case study using eutherian mammals. **BMC Genomics** 12:84, 2011.

MALOGOLOWKIN, C. Notas sobre "*Zygothrica dispar*". **Rev. Bras. Biol.** 12(4):455-457, 1952.

MARKOW, T. A.; O'GRADY, P. M. *Drosophila* biology in genomic age. **Genetics** 177:1269-1276, 2007.

MATA, T. A. *et al*. Invasive and flexible: niche shift in the drosophilid *Zaprionus indianus* (Insecta, Diptera). **Biol. Invasions** 12(5):1231-1241, 2010.

MCEVEY, S. F.; POLAK, M. *Mycodrosophila* (Diptera: Drosophilidae) of Fiji and Vanuatu with description of nine new species. **Bishop. Mus. Occas. Pap.** 84:35–67, 2005.

MEYER, C. P.; PAULAY, G. DNA barcoding: error rates based on comprehensive sampling. **PLoS Biol.** 3(12):e422, 2005.

MEDEIROS, H. F.; KLACZKO, L. B. How many species of *Drosophila* (Diptera, Drosophilidae) remain to be described in the forests of São Paulo, Brazil? Species lists of three forest remnants. **Biota Neotrop.** 4(1):1-12, 2004.

MILLER, K. B. *et* al. Association of insect life stages using DNA sequences: the larvae of *Philodytes umbrinus* (Motschulsky) (Coleoptera: Dytiscidae). **Syst. Entomol.** 30:499-509, 2005.

MITTERMEIER, R. A. Primate diversity and the tropical forest case studies from Brazil and Madagascar and the Importance of the megadiversity countries. In: WILSON, E. O.; PETER, F. M. (Eds.). **Biodiversity.** Washingtong: National Academies Press, 1988. pp. 145-157.

MITTERMEIER, R. A.; ROBLES-GIL, P.; MITTERMEIER, C. G. **Megadiversity: Earth's biologically wealthiest nations.** 1. ed. Monterrey: CEMEX, 1997.

MORALES-HOJAS, R.; VIEIRA, J. Phylogenetic patterns of geographical and ecological diversification in the subgenus *Drosophila*. PLoS ONE 7(11):e49552, 2012.

MUTANEN, M.; KAILA, K.; TABELL, J. Wide-ranging barcoding aids discovery of one-third increase of species richness in presumably well-investigated moths. **Sci. Rep.** 3(2901):1-7, 2013.

MYERS, N. *et al*. Biodiversity hotspots for conservation priorities. **Nature** 403:853-858, 2000.

O'GRADY, P.; MARKOW, T. Phylogenetic taxonomy in *Drosophila* - problems and prospects. **Fly** 3(1):1-5, 2009.

OLIVEIRA, D. C. S. G.; ALMEIDA, F. C.; O'GRADY, P. M. Monophyly, divergence times, and evolution of host plant use inferred from a revised phylogeny of the *Drosophila repleta* species group. **Mol. Phylogenet. Evol.** 64:533–544, 2012.

PANDEY, U. B.; NICHOLS, C. D. Human disease models in *Drosophila melanogaster* and the role of the fly in therapeutic drug discovery. **Pharmacol. Rev.** 63(2):411-436, 2011.

PETERSON, A. T.; SOBERÓN, J.; SÁNCHEZ-CORDERO, V. Conservatism of ecological niches in evolutionary time. **Science** 285:1265-1267, 1999.

PIPKIN, S. B.; RODRIGUEZ, R. L.; LEÓN, J. Plant host specificity among flower-feeding neotropical *Drosophila* (Diptera: Drosophilidae). **Am. Natur.** 100:135-156, 1966.

PLANAS, E. *et al*. Ecological niche and phylogeography elucidate complex biogeographic patterns in *Loxosceles rufescens* (Araneae, Sicariidae) in the Mediterranean Basin. **BMC Evol. Biol.** 14:1-17, 2014.

POPPE, J. L.; VALENTE, V. L.; SCHMITZ, H. J. Structure of Drosophilidae assemblage (Insecta, Diptera) in Pampa Biome (São Luiz Gonzaga/RS). **Pap. Avulsos Zool.** 52(16):185-195, 2012.

POPPE, J. L. *et al.* High diversity of Drosophilidae (Insecta, Diptera) in the Pampas Biome of South America, with descriptions of new *Rhinoleucophenga* species. **Zootaxa** 3779(2):215–245, 2014.

POWELL, J. R. Progress and prospects in evolutionary biology: the *Drosophila* model. New York: Oxford University Press 1997.

PRAMUAL, P.; WONGPAKAM, K. Association of black fly (Diptera: Simuliidae) life stages using DNA Barcode. **J. Asia Pac. Entomol.** 17(3):549-554, 2014.

PYRON, A.; BURBRINK, F.T. Lineage diversification in a widespread species: roles for niche divergence and conservatism in the common kingsnake, *Lampropeltis getula*. **Mol. Ecol.** 18:3443-3457, 2009.

RASO, L. *et* al. Molecular identification of adult and juvenile Linyphiid and Theridiid spiders in alpine glacier foreland communities. **PLoS ONE** 9(7):e101755, 2014.

RATNASINGHAM, S.; HEBERT, P. D. N. BOLD: the Barcode of Life Data system (www.barcodinglife.org). **Mol. Ecol. Notes** 7:355-364, 2007.

REMSEN, J.; O'GRADY, P. Phylogeny of Drosophilinae (Diptera: Drosophilidae), with comments on combined analysis and character support. **Mol. Phylogenet. Evol.** 24:249-264, 2002.

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. **Mol. Phylogenet. Evol.** 36:623-640, 2005.

ROBE, L. J.; VALENTE, V. L. S.; LORETO, E. L. S. Phylogenetic relationships and macro-evolutionary patterns within the *Drosophila tripunctata* "radiation" (Diptera: Drosophilidae). **Genetica** 138:725-735, 2010a.

ROBE, L. J.; LORETO, E. L. S.; VALENTE, V. L. S. Radiation of the ,,*Drosophila*" subgenus (Drosophilidae, Diptera) in the Neotropics. **J. Zoolog. Syst. Evol. Res**. 48(4):310-321, 2010b.

ROBERTS, D. B. *Drosophila melanogaster*: the model organism. **Entomol. Exp. Appl.** 121:93-103, 2006.

ROQUE, F.; TIDON, R. Eight new records of drosophilids (Insecta; Diptera) in the Brazilian savanna. **Dros. Inf. Serv.** 91:94- 98, 2008.

ROUQUETTE, J.; DAVIS, A. J. *Drosophila* species (Diptera: Drosophilidae) oviposition patterns in fungi: the effects of allospecifics, substrate toughness, ovipositor structure and degree of specialization. **Eur. J. Entomol.** 100:351-355, 2003.

RUSSO, C. M. *et al*. Phylogenetic analysis and a time tree for a large drosophilid data set (Diptera: Drosophilidae). **Zool. J. Linn. Soc.** 169:765–775, 2013.

SANCHES, A. *et al*. Illegal hunting cases detected with molecular forensics in Brazil. **Investig. Genet.** 3(1):17, 2012.

SEIFERT, K. A. *et al.* Prospects for fungus identification using CO1DNA barcodes, with *Penicilliumas* a test case. **Proc. Natl. Acad. Sci. U.S.A.** 104:3901–3906, 2007.

SEIFERT, K. A Progress towards DNA Barcoding of fungi. **Mol. Ecol. Resour.** 9:83–89, 2009.

SHATZ, A. J. *et al*. Characterizing the potential distribution of the invasive Asian longhorned beetle (*Anoplophora glabripennis*) in Worcester County, Massachusetts. **Appl. Geogr.** 45:259-268, 2013.

STUMP, A. D. *et al.* Distribution and mechanism of α-amanitin tolerance in mycophagous *Drosophila* (Diptera: Drosophilidae). **Environ. Entomol.** 40(6):1604-1612, 2011.

SVENNING, J.-C. *et al*. Applications of species distribution modeling to paleobiology. **Quat. Sci. Rev.** 30:2930-2947, 2011.

TATARENKOV, A.; AYALA, F. J. Phylogenetic relationships among species groups of the *virilis-repleta* radiation of *Drosophila*. **Mol. Phylogenet. Evol.** 21(2):327-331, 2001.

THROCKMORTON, L. H. The phylogeny, ecology and geography of *Drosophila*. IN: KING, R. C. (ed), Handbook of genetics. New York: Plenum, 1975.

TODA, M. J.; KIMURA, M. T.; TUNO, N. Coexistence mechanisms of mycophagous drosophilids on multispecies fungal hosts: aggregation and resource partitioning. **J. Anim. Ecol.** 68:794-803, 1999.

TUNO, N. *et al.* Tolerance of *Drosophila* flies to ibotenic acid poisons in mushrooms. **J. Chem. Ecol.** 33:311–317, 2007.

VALADÃO, H.; HAY, J. D. V.; TIDON, R. B. Temporal dynamics and resource availability for drosophilid fruit flies (Insecta, Diptera) in a gallery forest in the Brazilian Savanna. **Int. J. Ecol.** 2010:1-7, 2010.

VAN DER LINDE, K. V. D.; HOULE, D. A supertree analysis and literature review of the genus *Drosophila* and closely related genera (Diptera, Drosophilidae). **Insect Syst. Evol.** 39:241-267, 2008.

VAN DER LINDE, K. V. D. *et al*. A supermatrix-based molecular phylogeny of the family Drosophilidae. **Genet. Res. (Camb).** 92:25-38, 2010.

VENCES, M. Comparative performance of the 16S rRNA gene in DNA Barcoding of amphibians. **Front. Zool.** 2:5, 2005.

VILELA, C. R.; BÄCHLI, G. Taxonomic studies on Neotropical species of seven genera of Drosophilidae (Diptera). **Mitt. Schweiz. entomol. Ges.** 63:1-332, 1990.

VILELA, C. R.; BÄCHLI, G. On the identities of nine Neotropical species of *Hirtodrosophila* (Diptera, Drosophilidae). **Mitt. Schweiz. entomol. Ges.** 77:161–195, 2004.

VILELA, C. R.; BÄCHLI, G. Three new species of Neotropical Drosophilidae (Diptera). **Mitt. Schweiz. entomol. Ges.** 78:143-163, 2005.

VILELA, C. R.; BÄCHLI, G. (2007) Revision of the Neotropical genus Paraliodrosophila (Diptera, Drosophilidae). **Mitt. Schweiz. entomol. Ges.** 80:291–317, 2007.

WHEELER, M. R.; TAKADA, H. A revision of the American species of *Mycodrosophila* (Diptera, Drosophilidae). **Ann. Entomol. Soc. Am.** 56(3):392-399, 1963.

WILL, K. P.; MISHLER, B. D.; WHEELER, Q. D. The perils of DNA Barcoding and the need for integrative taxonomy. **Syst. Biol.** 54:844–851, 2005.

YAMASHIDA, S.; HIJII, N. The role of fungal taxa and developmental stage of mushrooms in determining the composition of the mycophagous insect community in a Japanese forest. **Eur. J. Entomol.** 104:225-233, 2007.

YASSIN A. Phylogenetic classification of the Drosophilidae Rondani (Diptera): the role of morphology in the postgenomic era. **Syst. Entomol.** 38:349-364, 2013.

YODER J. B. *et al.* Ecological opportunity and the origin of adaptive radiations. **J. Evol. Biol.** 23:1581-1596, 2010.

YOROZUYA, H. Effects of parasitoids on a mycophagous drosophilid community in northern Japan and an evaluation of the disproportionate parasitism hypothesis. **Entomol. Sci.** 9:13–22, 2006.

YOROZUYA, H. Effects of nematode parasitism on mycophagous drosophilid species in northern Japan. **Entomol. Sci.** 12:363–369, 2009.

YOTOKO, K. S. C. *et al.* A molecular study of the systematics of the *Drosophila tripunctata* group and the *tripunctata* radiation. **Mol. Phylogenet. Evol.** 28:614–619, 2003.

XIA, Y. *et al.* COI is better than 16S rRNA for DNA Barcoding Asiatic salamanders (Amphibia: Caudata: Hynobiidae). **Mol. Ecol. Resour.** 12:48–56, 2012.

ZEALE, M. R. K. *et* al. Taxon-specific PCR for DNA Barcoding arthropod prey in bat faeces. **Mol. Ecol. Resour.** 11:236-244, 2011.

ANEXOS

ANEXO A – Nota técnica

Título:

An efficient and cheap entomological aspirator to collect mycophylic and anthophilic adult Drosophilidae flies

Artigo publicado no periódico:

***Drosophila* Information Service – DIS**

ISSN:

0070-7333

can be derived from a single injected embryo. Injection of a mixture of transgenic plasmids therefore provides a fast and cheap method of generating multiple transformed fly lines with a relatively small number of microinjections.

The co-injection method requires that transformed progeny be individually genotyped, but this does not slow down the crosses, as red-eyed flies are genotyped after mating. More importantly, the transgenes of all transformed lines should be sequenced, to rule out human error and acquired mutations, regardless of the method of injection. For example, of the 18 flies genotyped here, one showed evidence of mutations in the "A" transgene that was not present in the injected "A" DNA, or in any other fly bearing the same transgene. This may be the result of a PCR amplification error, but alternatively it may reflect a DNA mutation occurring before or after transgene insertion. Mutations aside, without genotyping it is impossible to rule out the possibility that the DNA or the flies could have been mislabeled, either in the lab or by the injection company.

The results presented here almost certainly underestimate the frequency of independent transformation events in different germline cells within a multiplex-injected embryo, for two reasons. First, only three progeny were selected for genotyping; sequencing of additional red-eyed progeny could only have increased the count of embryos giving rise to progeny bearing all three transgenes. Second, only three transgenic vectors were co-injected; it is possible that, for example, the three "B" progeny of injected embryo #3 represent three independent transformation events within that embryo. Taking this into account, it is possible that the number of co-injected plasmids could be increased significantly, further reducing the number of embryos to be injected.

References: Bischof, J., R.K. Maeda, M. Hediger, F. Karch, and K. Basler 2007, Proc. Natl. Acad. Sci. U.S.A. 104: 3312–3317; Boy A.L., Z. Zhai, A. Habring-Müller, Y. Kussler-Schneider, P. Kaspar, and I. Lohmann 2010, Genesis 48: 452–456; Gloor, G.B., C.R. Preston, D.M. Johnson-Schlitz, N.A. Nassif, R.W. Phillis, W.K. Benz, H.M. Robertson, and W.R. Engels 1993, Genetics 135: 81–95; Ramos, A.I., and S. Barolo 2013, Philos. Trans. R. Soc. Lond. B 368: 2013-18; Rubin, G., and A. Spradling 1982, Science 218: 348–353; Swanson, C.I., T. Hinrichs, L.A. Johnson, Y. Zhao, and S. Barolo 2008, Gene 408: 180-186.

An efficient and cheap entomological aspirator to collect mycophylic and anthophilic adult Drosophilidae flies.

Machado, Stela[1], João Pedro Junges dos Santos[1], Lizandra Jaqueline Robe[1,2], and Elgion Lucio da Silva Loreto[1]. [1]Universidade Federal de Santa Maria (UFSM), Programa de Pós Graduação em Biodiversidade Animal (PPGBA), Avenida Roraima, 1000, CEP 97105-900 - Santa Maria, RS - Brazil, Phone: +55 55 3220-8912; [2]Universidade Federal do Rio Grande (FURG), Programa de Pós-Graduação em Biologia de Ambientes Aquáticos Continentais (PPGBAC), Instituto de Ciências Biológicas – ICB, Av. Itália, km 08, Campus Carreiros, CEP 96203-900 - Rio Grande, RS – Brazil, Phone: +55 53 3293-5128; Corresponding author: Lizandra Jaqueline Robe, lizbiogen@gmail.com.

Introduction

Traditionally, the methodology used to collect drosophilids in Brazil relies on flies' attraction to traps baited with resources, principally fermented fruits like banana (Tidon and Sene, 1988; Medeiros and Klaczko, 1999). However, this collection method attracts mainly frugivorous species of the genus *Drosophila* (Gottschalk *et al.*, 2008), providing a biased sample of subjacent biodiversity, once species with other feeding preferences are rarely recorded. In fact, *Drosophila* encompasses almost 60% of the 304 reported Brazilian drosophilid species, being followed by far by the mycophylic *Zygothrica* (with 54 species) and *Hirtodrosophila* genera (with only 16 species) (Gottschalk *et al.*, 2008).

Species of *Hirtodrosophila*, *Mycodrosophila*, *Paraliodrosophila*, and *Zygothrica* encompass the putatively monophyletic *Zygothrica* genus group (Grimaldi, 1990), which presents different degrees of association with macroscopic fungi. As only part of these species use fungi as resources for feeding or

oviposition (Courtney *et al*., 1990), we generically named them "mycophylic" species instead of "mycophagous", which is more frequently used. In this case, there seems to be a negative correlation between collection frequency and specialization level in the use of macroscopic fungi as feeding, breeding, or oviposition sites. *Zygothrica* is just the more generalist genus (Courtney *et al.*, 1990; Grimaldi, 1987) and the best recorded mycophylic taxon (Gottschalk *et al.*, 2008), although resource specialization may not be the only factor responsible for this scenario. In fact, reduced or skewed sampling and high levels of unregistered diversity seems to be the case for both, mycophilic (Bolzan, 2011) and anthophilic (Schmitz, 2010) species, which seem to be much more diverse in the Neotropics than previously reported.

The traditional methods used to collect mycophylic drosophilid species are entomological nets, mouth aspiration, or through the collection and storage of resources until adult eclosion (Markow and O'Grady, 2006; Gottschalk *et al.*, 2009; Robe *et al.*, 2014). This last method also seems to be widely used in the collection of antophylic drosophilid species (Vilela, 1984; dos Santos and Vilela 2005; Robe *et al.*, 2013). However, these sampling methodologies have important limitations: entomological nets are frequently hampered by fungi or flower disposition (that sometimes block net passage), and tend to be inefficient when the number of available specimens is low; mouth aspiration adds a health risk, once the collector could aspirate potentially harmful fungus spores or flower pollen, and it is also very inefficient. Once only few flies could be collected before breath breaking. Collection and storage of resources is also a biased sampling strategy, since some species can use fungi/flowers for purposes other than oviposition and these will not be collected at all. So, we developed a cheap entomological aspirator in order to make the collection of mycophagous and/or antophylous drosophilids more safe and efficient.

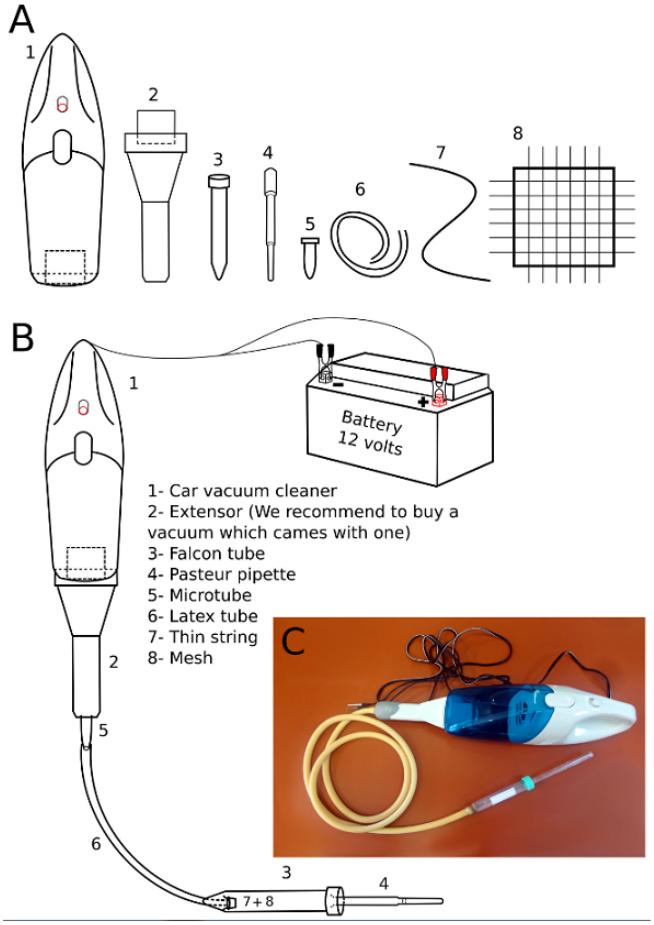

Figure 1. Entomological aspirator scheme showing the required components (A), the assembly design (B), and the final equipment photography (C).

Aspirator Design

The aspirator construction is based on the use of a car vacuum cleaner (preferentially with an extensor), which is coupled to a 1.5 ml microtube (or other small rigid tube), approximately 60-120 cm of a latex or plastic tube with a diameter around 0.5 cm (the first is preferred due to its higher flexibility), a thin string and a fine mesh, a 15 ml Falcon tube, a plastic Pasteur pipette (or other rigid tube as a piece of glass pipette or a Bic pen), epoxy adhesive Araldite and Durepox, two alligator battery clips, and a motorcycle battery (Figure 1).

First of all, the Falcon lid needs to be cut and connected with Araldite to a rigid tube provided, for example, by a Pasteur pipette. Both, the Falcon tube and the Pasteur pipette should have their extremities cut. In parallel, one end of the latex or plastic tube needs to be covered by a mesh with the use of a thin string. This region of the latex or plastic tube should then be passed through the cut end of the

Falcon tube in a way that the mesh is placed within the tube. Araldite glue should be used to connect these pieces firmly, without leaving any air passage. The 1.5 ml microtube needs also to be cut at both of its extremities, and its major diameter end needs to be connected to the vacuum cleaner extensor or, in its absence, directly to the car vacuum cleaner with the use of Durepox. The minor diameter extremity of the cut microtube should be firmly connected to the latex tube. The vacuum cleaner plug to a car's cigarette lighter needs finally to be changed to alligator clips or electrical plugs in order to connect the entire equipment to the motorcycle battery (Figure 1). In order to fasten and easily transport this manufactured entomological aspirator, it is important to leave the extensor unconnected to the main piece, so that the equipment is mounted in the field and readily connected to the battery.

Advantages and Disadvantages

The manufactured entomological aspirator presented here is easily constructed, transported, and handled. Besides, it is safer than a mouth aspirator, allowing the sampling of adult flies in spaces difficult to access with the use of entomological nets. According to flies' availability, it allows effective capture of a great number of species and specimens (tens to hundreds) in a short period of time. This equipment was used by Robe *et al.* (2014) in the sampling of mycophilic drosophilid species across different Brazilian biomes, where it was shown to be highly efficient, leading to the capture of more than 300 individuals encompassing 22 species (besides approximately 180 individuals belonging to undescribed species) in no more than 45 hours (15 collection sites with only 2-4 hours of active search + aspiration activity). The only disadvantages of the equipment refer to battery weights and the need for battery charging before going to the field, although the time of duration of battery generally compensates.

Acknowledgments: We are grateful to the Brazilian Funding Agency Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support (Universal-CNPq 14/2013, process number 471174/2013-0).

References: Bolzan, A.R., 2011, DNA barcode de drosofilídeos micófagos pertences aos gêneros *Hirtodrosophila*, *Mycodrosophila* e *Zygothrica*, unpublished dissertation; Courtney, S.P., T.T. Kibota, and T.A. Singleton 1990, Adv. Ecol. Res. 20: 225–274; Dos Santos, R.C.O, and C.R. Vilela 2005, Rev. Bras. Entomol. 49(4): 544-551; Grimaldi, D.A., 1987, Bull. Am. Mus. Nat. Hist. 186: 1-268; Grimaldi, D.A., 1990, Bull. Am. Mus. Nat. Hist. 197: 1–139; Gottschalk, M.S., P.R.S. Hofmann, and V.L.S. Valente 2008, Check list 4(4): 485-518; Gottschalk, M.S., L. Bizzo, J.S. Döge, M.S. Profes, P.R.S. Hofmann, and V.L.S. Valente 2009, Iheringia 99(4): 442-448; Machado, S., J.P.J. dos Santos, P.M. Fonseca, A.R. Bolzan, J. David, E.L.S. Loreto, M.S. Gottschalk, and L.J. Robe, Mycophilic drosophilids: DNA barcoding as a way of overcoming the taxonomic impediment, unpublished article; Markow, T.A., and P. O'Grady 2006, Drosophila: *A Guide to Species Identification and Use*. Academic Press (Elsevier), London; Medeiros, H.F, and L.B. Klaczko 1999, Dros. Inf. Serv. 82: 100-102; Schmitz, H.J., 2010, Genética, Ecologia e Evolução de drosofilídeos (Insecta, Diptera) associados a flores, unpublished thesis; Robe, L.J., S. Machado, A.R. Bolzan, J.P.J. dos Santos, F.B. Vales, A.P. dos Santos, M.L. Blauth, and M.S. Gottschalk 2014, Stud. Neotrop. Fauna E. 49(2): 79-94; Tidon, R., and F.M. Sene 1988, Dros. Inf. Serv. 67: 90; Vilela, C.R., 1984, Revta Brasil. Zool. 2(2): 63-69.

ANEXO B – Material suplementar do Capítulo 3

Table S1. List of sampling sites of the *Zygothrica* genus group mycophylic species, with their respective geographic co-ordinates.

	Sampling cities	Specifications	Coordinates	Biome
1	Melgaço (PA)	Floresta Nacional de Caxiuanã	-51.457129, -1.737762	Amazonian Forest
2	Colorado do Oeste (RO)	Private property I	-60.577239, -13.105933	Amazonian Forest
3	Colorado do Oeste (RO)	Private property II	-60.59025, -13.010472	Amazonian Forest
4	Colorado do Oeste (RO)	Private property III	-60.595756, -13.111058	Amazonian Forest
5	Teodoro Sampaio (SP)	Parque Estadual Morro do Diabo	-52.175194, -22.624639	Atlantic Forest
6	Diamante do Norte (PR)	Estação Ecológica do Caiuá	-52.873833, -22.607778	Atlantic Forest
7	Tuneiras do Oeste (PR)	Reserva Biológica das Perobas	-52.818611, -23.886222	Atlantic Forest
8	Florianópolis (SC)	Morro da Lagoa da Conceição	-48.475833, -27.590833	Atlantic Forest
9	Bossoroca (RS)	Private property	-54.982892, -28.525614	Pampa
10	Ivoti (RS)	Cascata São Miguel	-51.117222, -29.583611	Atlantic Forest
11	Pejuçara (RS)	Balneário do Picolé	-53.665278, -28.386667	Atlantic Forest
12	Pelotas (RS)	Horto botânico UFPEL	-52.431944, -31.816111	Pampa
13	Rio Grande (RS)	Estação Ecológica do Taim	-52.535205, -32.538438	Pampa
14	Santa Maria (RS)	Bosque UFSM	-53.713056, -29.72194	Pampa
15	Santa Maria (RS)	Jardim Botânico (UFSM)	-53.726111, -29.717222	Pampa
16	Santa Maria (RS)	Morro do Elefante	-53.720484, -29.678583	Atlantic Forest
17	Santa Maria (RS)	Cascata do Mezzomo	-53.725556, -29.587778	Atlantic Forest
18	Santa Maria (RS)	São Marcos	-53.691554, -29.678847	Atlantic Forest
19	Santiago (RS)		-54.8575, -29.214167	Pampa
20	Viamão (RS)	Parque Saint'Hilaire	-51.101944, -30.088056	Pampa

Table S2. Species collected in each point numbered according to Table S1.

Species	N	M	F	MID %	Sampling points
H. gilva (Burla 1956)	6	3	3	9.10	2, 3, 6
H. levigata (Burla 1956)	18	13	5	12.86	7, 8, 10, 11, 16, 17, 20
H. mendeli (Mourão, Gallo & Bicudo 1965)	9	4	5	10.20	6, 17, 19
H. morgani (Mourão, Gallo & Bicudo 1967)	7	4	3	12.70	4, 6, 9, 10
H. subgilva (Burla 1956)	3	2	1	11.90	8, 17
H. trilineata (Chung 1960)	*	*	*	12.60	*
H. levigata affinis	1	0	1	12.86	11
H. mendeli affinis	2	2	0	10.20	6
H. morgani affinis sul	6	3	3	7.70	12, 13, 17, 19
H. subflavohalterata affinis 1	3	2	1	9.20	2, 4
H. subflavohalterata affinis 2	4	2	2	2.80	2, 4
H. subflavohalterata affinis 3	2	2	0	9.50	1
H. subflavohalterata affinis 4	1	0	1	9.20	1
H. subflavohalterata affinis 5	9	4	5	5.50	2, 3, 4, 5
H. subflavohalterata affinis H002	9	4	5	2.80	5, 6, 10, 11, 14, 16
H. subgilva affinis 1	1	1	0	7.70	19
H. subgilva affinis 2	3	1	2	9.10	17
H. H007	1	0	1	11.80	17
H. H010	1	0	1	8.70	8
H. H012	1	0	1	10.60	20
H. H013	1	0	1	11.50	8
M. dimidiata (Loew 1862)	*	*	*	9.30	*
M. elegans (Wheeler & Takada 1963)	3	0	3	11.10	1, 5, 6
M. projectans (Sturtevant 1916)	10	4	6	10.00	11, 14, 16, 20
M. M001	5	2	3	14.00	13, 14, 15
M. affinis M001	1	1	0	14.00	1
M. neoprojectans affinis 1	3	2	1	5.50	1, 3
M. neoprojectans affinis 2	4	3	1	5.50	1, 2, 3
M. projectans affinis 1	2	0	2	10.00	11
M. projectans affinis 2	4	3	1	10.90	1, 5
Pa. antennata (Wheeler 1957)	10	5	5	11.50	3, 4, 5, 6, 11, 15, 19
Pa. burlai (Vilela & Bächli 2007)	2	2	0	9.80	8
Zy. atriangula (Duda 1927)	2	2	0	7.50	1
Zy. dispar (Wiedemann 1830)	6	2	2	9.40	6, 7, 8, 10
Zy. hypandriata (Burla 1956)	7	6	1	9.70	16, 17, 20
Zy. orbitalis (Sturtevant 1916)	13	9	4	10.60	5, 6, 8, 10, 12, 17
Zy. parapoeyi (Burla 1956)	2	2	0	7.30	10, 15
Zy. parvipoeyi (Burla 1956)	3	2	1	10.70	17
Zy. prodispar (Duda 1925)	2	2	0	9.10	6, 12
Zy. ptilialis (Burla 1956)	11	9	2	9.70	3, 7, 8, 10, 12, 16, 17, 20
Zy. virgatinigra (Burla 1956)	1	1	0	11.80	1
***Zy. vittimaculosa* (Burla 1956)**	1	0	1	3.70	10

***Zy. vittinubila* (Burla 1956)**	1	1	0	11.00	7
Zy. apopoeyi affinis	8	7	1	12.50	7, 8
Zy. atriangula affinis H015	1	0	1	7.50	10
Zy. festiva affinis	1	1	0	9.80	10
Zy. fuscina affinis	2	1	1	12.80	15
Zy. gracilipoeyi affinis	1	1	0	7.20	3
Zy. laevifrons 1	5	3	2	7.80	8, 10, 16, 20
Zy. laevifrons 2	3	2	1	7.80	2, 7
Zy. parapoeyi affinis	4	2	2	7.30	7
Zy. poeyi affinis	1	0	1	11.80	8
Zy. prodispar affinis 1	4	2	2	9.10	7, 10
Zy. prodispar affinis 2	2	1	1	9.00	10
Zy. vittimaculosa affinis	3	#	#	3.70	18, 20
Zy. H009	4	1	2	10.80	8

Notes: * Sequence downloaded from GenBank; # Sexing was not performed. N = number of specimens, M = number of males, F = number of females, MID% Minimum interspecific distance presented by each sampled species (in %).

Table S3. Bold access numbers for the newly obtained COI sequences. (Table omitted).

Table S4. Genbank access numbers from the downloaded COI Drosophilinae sequences (Table omitted).

Table S5. K2P summary distances for each evaluated taxonomic category.

Taxonomic category		Min distance (%)	Max distance (%)	Mean distance (±std. err.) (%)	TRR[1]	BG[2]
Hirtodrosophila genus	Intraespecific	0	5.1	0.72 (0.26)	17.4	-2.7
	Interespecific	2.4	18.6	12.5 (1.38)		
	Interespecific to cryptic species	2.4	17	12.31 (1.37)		
Mycodrosophila genus	Intraespecific	0	3.3	1.27 (0.38)	10.2	1.3
	Interespecific	4.6	17	12.9 (1.44)		
	Interespecific to cryptic species	4.6	15.3	12.77 (1.43)		
Paraliodrosophila genus	Intraespecific	0	1.5	0.37 (0.17)	38.3	12
	Interespecific	13.6	14.8	14.18 (1.62)		
Zygothrica genus	Intraespecific	0	5.3	0.77 (0.26)	17.5	-1.6
	Interespecific	3.7	21.7	13.48 (1.48)		
	Interespecific to cryptic species	3.7	17.7	13.01 (1.43)		
Zygothrica generic group	Intraespecific	0	5.3	0.77 (0.27)		
	Interespecific congeneric	2.4	21.7	12.99 (1.43)		
	Interespecific intergeneric	7.4	20	14.15 (1.51)		

[1] Taxonomic resolution ratio (quotient between mean interspecific congeneric divergences and mean intraspecific divergence)
[2] Barcoding gap (subtraction of minimum interspecific distance from maximum intraspecific divergence)

Table S6. Threshold performance comparison.

Threshold	False negative*	False positive*	Error rate (%)	Origin
1%	3	5	3.60	BOLD System Threshold (Ratnasingham & Hebert, 2007)
1.15-1.5%	3	4	3.15	optimization based on ABGD estimation (Puillandre *et al*., 2012)
1.92%	16	4	9.01	optimization based on the “localMinima” function of Spider (Brown *et al*., 2012)
3%	25	0	11.26	standard proposed by Hebert *et al*. (2003a)
5.25%	29	0	13.06	optimization based on Lefébure *et al*. (2006)

* Estimations performed in ABGD.

Table S7. Generalized Mixed Yule Coalescent (GMYC) analysis cluster subdivision. (Table omitted).

ANEXO C – Material suplementar do Capítulo 4

Table S1. List of sampling sites of the *Zygothrica* genus group mycophylic species, with their respective geographic co-ordinates.

	Sampling cities	Specifications	Longitude	Latitude	Bioma
1	Bossoroca (RS)	Particular property	-54.982892	-28.525614	Pampa
2	Canela (RS)	Floresta Nacional de Canela	-50.886389	-29.3	Atlantic Forest
3	Canguçu (RS)	Particular property	-52.781582	-31.482258	Pampa
4	Chuí (RS)	Particular property	-53.435216	-33.657111	Pampa
5	Derrubadas (RS)	Parque Estadual do Turvo - Site I	-53.964439	-27.245034	Atlantic Forest
6	Derrubadas (RS)	Parque Estadual do Turvo - Site II	-53.881841	-27.138175	Atlantic Forest
7	Dom Pedro de Alcantara (RS)	Particular property	-49.853779	-29.368181	Atlantic Forest
8	Horizontina (RS)	Particular property	-54.302035	-27.632631	Atlantic Forest
9	Ivoti (RS)	Particular property	-51.117222	-29.583611	Atlantic Forest
10	Jaguari (RS)	Particular property Reserva Cerro Chapadão	-54.703365	-29.469931	Atlantic Forest
11	Pejuçara (RS)	Particular property	-53.665278	-28.386667	Atlantic Forest
12	Pelotas (RS)	Horto Botânico UFPel	-52.431944	-31.816111	Pampa
13	Pelotas (RS)	Morro Redondo	-52.630971	-31.55035	Pampa
14	Piratini (RS)	RPPN Minas do Paredão	-52.952263	-30.932599	Pampa
15	Rio Grande (RS)	Estação Ecológica do Taim	-52.535205	-32.538438	Pampa
16	Santa Maria (RS)	Particular property I	-53.725556	-29.587778	Atlantic Forest
17	Santa Maria (RS)	Bosque UFSM	-53.713056	-29.72194	Pampa
18	Santa Maria (RS)	Jardim Botânico UFSM	-53.726111	-29.717222	Pampa
19	Santa Maria (RS)	Morro do Elefante	-53.720484	-29.678583	Atlantic Forest
20	Santa Maria (RS)	Particular property	-53.726697	-29.678583	Atlantic Forest
21	Santa Maria (RS)	Particular property II	-53.691554	-29.678847	Atlantic Forest
22	Santiago (RS)	Particular property	-54.8575	-29.214167	Pampa
23	Santiago (RS)	Particular property in Ernesto Alves	-54.713508	-29.252444	Pampa
24	São Francisco de Paula (RS)	Floresta Nacional de São Francisco de Paula	-50.386389	-29.422778	Atlantic Forest
25	Torres (RS)	Particular property	-49.763905	-29.369366	Pampa
26	Viamão (RS)	Parque Saint'Hilaire	-51.101944	-30.088056	Pampa
27	Blumenau (SC)	Parque Nacional da Serra do Itajaí	-49.083333	-27.05	Atlantic Forest
28	Florianópolis (SC)	Morro da Lagoa da Conceição	-48.475833	-27.590833	Atlantic Forest
29	Florianópolis (SC)	Parque Municipal da Lagoa do Peri	-49.833333	-29.383333	Atlantic Forest
30	Garopaba (SC)		-48.638173	-28.105941	Atlantic Forest
31	Três Barras (SC)	Floresta Nacional de Três	-50.303709	-26.219199	Atlantic

		Barras			Forest
32	Diamante do Norte (PR)	Estação Ecológica do Caiuá	-52.873833	-22.607778	Atlantic Forest
33	Foz do Iguaçu (PR)	Parque Nacional do Iguaçu	-54.478745	-25.620647	Atlantic Forest
34	Guaraqueçaba (PR)	APA de Guaraqueçaba	-48.3	-25.2	Atlantic Forest
35	Guarapuava (PR)		-51.298331	-25.063492	Atlantic Forest
36	Tuneiras do Oeste (PR)	Reserva Biológica das Perobas	-52.818611	-23.886222	Atlantic Forest
37	Salesópolis (SP)	Estação Biológica de Boracéia	-45.868313	-23.631432	Atlantic Forest
38	Teodoro Sampaio (SP)	Parque Estadual Morro do Diabo	-52.175194	-22.624639	Atlantic Forest
39	Itatiaia (RJ)	Parque Nacional de Itatiaia	-44.6	-22.45	Atlantic Forest
40	Nova Iguaçu (RJ)	Parque Nacional do Tinguá	-43.433333	-22.566667	Atlantic Forest
41	Colorado do Oeste (RO)	Particular property I	-60.577239	-13.105933	Amazon Rainforest
42	Colorado do Oeste (RO)	Particular property II	-60.59025	-13.010472	Amazon Rainforest
43	Colorado do Oeste (RO)	Particular property III	-60.595756	-13.111058	Amazon Rainforest
44	Melgaço (PA)	Floresta Nacional de Caxiuanã	-51.457129	-1.737762	Amazon Rainforest
45	Maceio (AL)		-35.703274	-9.607393	Atlantic Forest

Table S2. Summary information and essencial statistics for each sampled species. Species collected in each point were numbered according to Table S1.

	Species	N	M	F	MI%	MX%	MD% (±)	π%(SD%)	H	Hd (SD)	S	Tajima's D	Fu and Li's D	Fu and Li's F	Sampling points	Sequences origineted by
COI	***M. dimidiata*** (Loew 1862)	1	-	-	-	-	-	-	-	-	-	-	-	-		GenBank: EU493682.1
	M. elegans (Wheeler & Takada 1963)	3	-	3	0	1.5	1 (0.3)	0.963 (0.454)	2	0.667 (0.314)	9	-	-	-	32, 38, 44	Machado et al. (in preparation)
	M. projectans (Sturtevant 1916)	55	35	20	0	5.8	1.5 (0.3)	1.418 (0.114)	31	0.957 (0.014)	46	-1.18365	**-3.14451**	**-2.88897**	1, 5, 6, 11, 12, 17, 19, 26, 27, 33	Machado et al. (in preparation) and in this manuscript
	M. projectans affinis 1	24	18	6	0	1.1	0.1 (0)	0.109 (0.044)	6	0.380 (0.125)	6	**-1.8104**	-2.4431	**-2.62117**	6, 11, 33, 40	Machado et al. (in preparation) and in this manuscript
	M. projectans affinis 2	32	26	6	0	2.3	0.8 (0.2)	0.749 (0.112)	21	0.923 (0.036)	37	**-1.94121**	***-3.30493***	***-3.37137***	1, 5, 6, 33, 38, 40, 44	Machado et al. (in preparation) and in this manuscript
	M. M001	11	6	5	0	1.3	0.3 (0.1)	0.201 (0.061)	5	0.709 (0.137)	5	-1.4646	-1.4445	-1.63404	15, 17, 18	Machado et al. (in preparation) and in this manuscript
	M. M001 affinis	1	1	0	-	-	-	-	-	-	-	-	-	-	44	Machado et al. (in preparation)
	M. neopreojectans affinis 1	3	2	1	0.1	0.8	0.6 (0.2)	0.558 (0.223)	3	1 (0.272)	6	-	-	-	44	Machado et al. (in preparation)
	M. neopreojectans affinis 2	4	3	1	0.2	0.8	0.4 (0.2)	0.37 (0.1)	4	1 (0.117)	4	-0.78012	-0.78012	-0.78052	41, 42, 44	Machado et al. (in preparation)
	M. SFP	1	-	1	-	-	-	-	-	-	-	-	-	-	24	This manuscript
	M. sp1	1	-	1	-	-	-	-	-	-	-	-	-	-	31	This manuscript
COII	***M. projectans*** (Sturtevant 1916)	27	12	15	0	9.2	1.4 (0.3)	2.626 (0.625)	9	0.684 (0.094)	16	-0.72659	-0.6094	-0.75611	6, 11,12, 17, 19, 26, 27	This manuscript
	M. projectans affinis 1	5	2	3	0	13.9	5.6 (0.9)	5.152 (2.461)	4	0.900 (0.161)	41	-1.21704	-1.17971	-1.128391	6, 33, 40	This manuscript
	M. projectans affinis 2	10	8	2	0.5	15.3	4.5 (0.6)	4.151 (1.071)	5	0.822 (0.097)	7	-0.46971	-0.12602	-0.23643	1, 6, 33, 38, 40, 44	This manuscript
COI + COII	***M. projectans*** (Sturtevant 1916)	27	12	15	0	3.9	1.3 (0.2)	1.457 (0.228)	23	0.972 (0.025)	49	-0.7719	-1.19637	-1.24783	6, 11, 12, 17, 19, 26, 27	This manuscript
	M. projectans affinis 1	5	2	3	0.2	4.4	2.1 (0.3)	1.952 (0.827)	5	1 (0.126)	45	-1.15378	-1.11965	-1.21921	6, 40	This manuscript
	M. projectans affinis 2	10	8	2	0.4	6.4	2.1 (0.2)	0.867 (0.111)	10	1 (0.045)	20	-0.90337	-0.82139	-0.949	1, 6, 33, 38, 40, 44	This manuscript

N = number of specimens, M = number of males, F = number of females, MI% minimum intraspecific distance in porcentage, MX% maximum intraspecific distance in porcentage, MD% mean intraspecific distance in porcentange with standard error, π nucleotide diversity, H number of haplotypes , Hd haplotype diversity, S variable sites. In bold significance of p<0.05 and in bold and italic p<0.02.

Table S3. Zygothrica genus group species with the respective number of individuals amplified to each gene.

Species	AMD	COI	Hb
H. levigata	2	19	3
H. morgani	1	-	-
M. elegans	1	3	-
M. projectans	1	18	4
M. projectans affinis 1	-	-	4
M. projectans affinis 2	-	-	2
M. SFP	1	1	-
Z. orbitalis	1	13	-
Z. ptilialis	3	11	1
Z. vittimaculosa	3	6	-
Z. vittimaculosa affinis	2	2	-

Table S4. Descrition of wing landmarks used in morphometric analysis.

Landmark	Description
1	Intersection between R2+3 and C veins
2	Intersection between R4+5 and C veins
3	Intersection between M and C veins
4	Intersection between CuA and C veins
5	Intersection between Cu and dM-Cu veins
6	Intersection between M and dM-Cu veins
7	Intersection between M and R-M veins
8	Intersection between R4+5 and R-M veins
9	Second bifurcation of R vein
10	First bifurcation of R vein

Table S5. Relationship of *M. projectans* complex species found in fungus. Sampling point number refer to Table S1.

Sampling point	Fungus subpoint	Fungus identification	Species captured
1	SLG6	*Ganoderma* sp.	*M. projectans*, *M. projectans* affinis 2
5	T6	*Polyporus* sp. 2	*M. projectans*, *M. projectans* affinis 2
6	Y4	*Ganoderma* sp. 3	*M. projectans*, *M. projectans* affinis 1, *M. projectans* affinis 2
6	Y5	Not identified	*M. projectans*
27	PNSI10	*Ganoderma sp1*	*M. projectans*
33	P6	Not identified	*M. projectans, M. projectans* affinis 1, *M. projectans* affinis 2
33	P25	*Polyporus* sp. 1	*M. projectans* affinis 2
33	F7	*Trametes* sp. 1	*M. projectans* affinis 2
33	F7'	*Ganoderma* sp. 4	*M. projectans* affinis 2
33	F5	Not identified	*M. projectans* affinis 2
33	P25	*Polyporus* sp. 1	*M. projectans* affinis 2
33	P23	Not identified	*M. projectans*
33	P22	Not identified	*M. projectans*
33	F8'	Not identified	*M. projectans*
33	F2	Not identified	*M. projectans* affinis 1
33	P13	*Trametes* sp. 2	*M. projectans* affinis 1
40	TIN8	*Ganoderma* sp.	*M. projectans* affinis 1
40	TIN3	Not identified	*M. projectans* affinis 2
44	PE	Not identified	*M. projectans* affinis 2

Table S6. ANOVA and CVA of wing variation between *M. projectans* complex.

	Sum of sqrs	df	Mean square	F	P
Between groups	5.50E+31	2	2.75E+31	1	0.3723
Within groups	1.39E+33	51	2.73E+31		
Total	1.45E+33	53			

Levene's test for homogeneity, from means P: 0.7497
Levene's test, from medians P: 0.8141