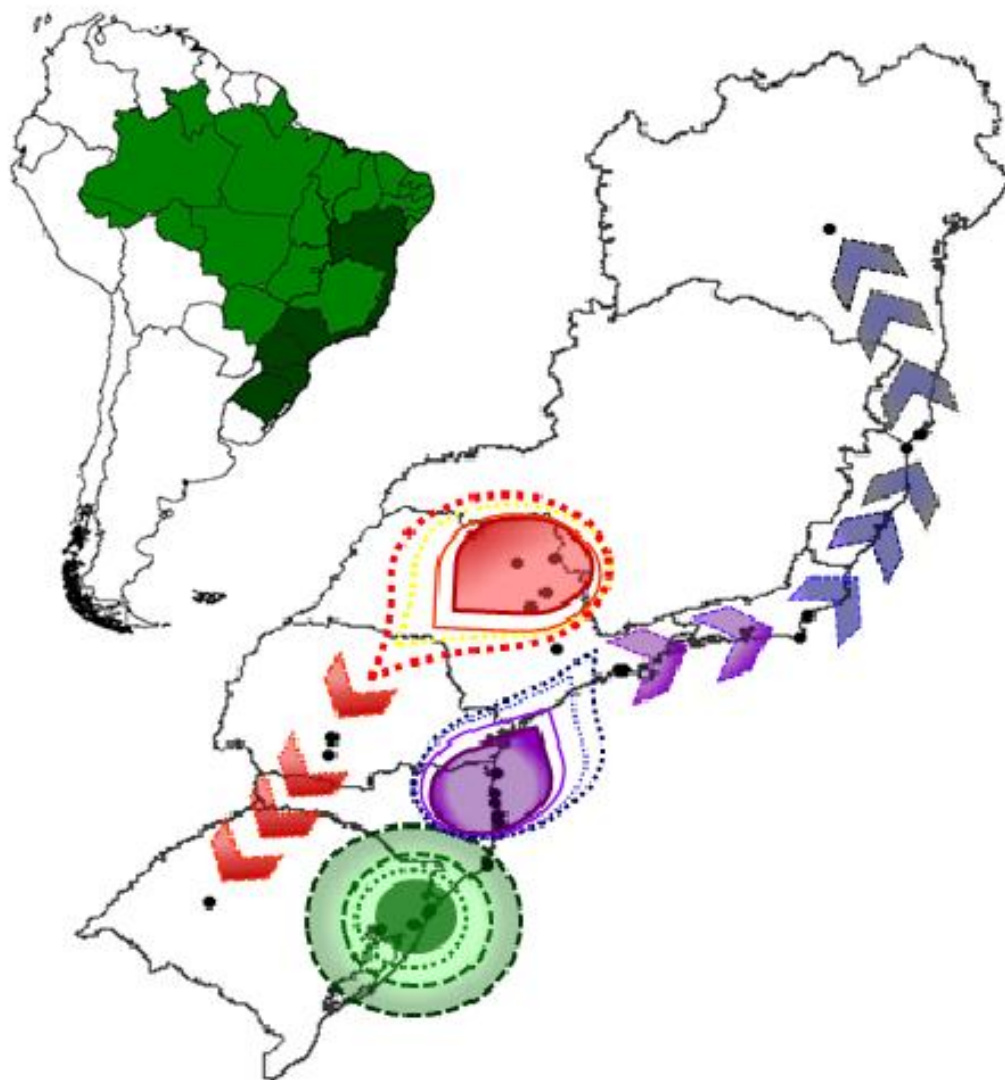


UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO  
DEPARTAMENTO DE GENÉTICA

História demográfica e estrutura de populações para a espécie cactófila  
*Drosophila meridionalis*



DORA YOVANA BARRIOS LEAL

Ribeirão Preto, SP  
2013

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

**História demográfica e estrutura de populações para a espécie cactófila  
*Drosophila meridionalis***

**DORA YOVANA BARRIOS LEAL**

Tese apresentada ao Departamento de Genética da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Ciências.

Área de Concentração: Genética.

**Orientador: Prof<sup>a</sup> Dr<sup>a</sup> Maura Helena Manfrin**

**Ribeirão Preto, SP  
2013**

Autorizo a divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada à fonte.

### ***FICHA CATALOGRÁFICA***

Barrios-Leal, Dora Yovana

História demográfica e estrutura de populações para a espécie cactófila *Drosophila meridionalis*. Dora Yovana Barrios Leal; orientadora: Maura Helena Manfrin. Ribeirão Preto, 2013

105 f.

Tese de Mestrado, apresentada à Faculdade de Medicina de Ribeirão/USP.  
Área de Concentração: Genética.

1. Filogeografia.
2. Gene *period*.
3. Gene *COI*.
4. *Drosophila meridionalis*.

Este trabalho foi realizado com auxílio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPQ, através do Programa de Estudantes-Convênio de Pós-Graduação PEC-PG. Número do processo: 190431/2010-5 Edital/Chamada: PEC-PG 2010.

## FOLHA DE APROVAÇÃO

Dora Yovana Barrios Leal

Título da Tese: **História demográfica e estrutura de populações para a espécie cactófila *Drosophila meridionalis***

Tese apresentada ao Departamento de Genética da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Ciências.

Área de Concentração: Genética.

Aprovado em:

Banca Examinadora:

Prof.(a) Dr.(a) \_\_\_\_\_

Instituição: \_\_\_\_\_ Assinatura \_\_\_\_\_

Prof.(a) Dr.(a) \_\_\_\_\_

Instituição: \_\_\_\_\_ Assinatura \_\_\_\_\_

Prof.(a) Dr.(a) \_\_\_\_\_

Instituição: \_\_\_\_\_ Assinatura \_\_\_\_\_

Dedico  
A minha mãe e meu pai,  
Às minhas irmãs e aos meus irmãos,  
Às minhas sobrinhas Lina e Ana Sofia e aos meus sobrinhos Francisco, Edwin, Eric, Juan  
Esteban, JuanJose e Junior,  
Ao meu Aron.

## AGRADECIMENTOS

- À Profa. Dra. Maura Helena Manfrin (FFCLRP/USP) pela única e grande oportunidade de realizar o mestrado, pela contribuição para o meu crescimento intelectual e pessoal com sua orientação competente, incentivo, apoio, confiança e amizade. Agradeço a sua orientação imprescindível para que eu pudesse concluir com sucesso esta tese.
- Ao Prof. Dr. Ademilson Espencer Egea Soares (FMRP/USP) pela oportunidade e confiança e pelo excelente trabalho que realiza na coordenação do Departamento de Genética.
- À Dra. Érica C. C. Silva, por gentilmente disponibilizar algumas sequências utilizadas no presente trabalho.
- À doutoranda Taís Carmona Lavagnini, minha irmã adotiva no Brasil, pelo companheirismo, pelas inúmeras correções do português não somente deste trabalho, pelas discussões sobre biologia, pela motivação e ajuda em diversas etapas ao longo destes dois anos. Obrigada, somente por tudo!!!!
- A doutoranda Gislaine Angélica Rodrigues Silva pela sua amizade, companheirismo e suporte nestes dois anos.
- A doutoranda Natácia Evangelista de Lima pelo companheirismo e revisão da redação de partes dessa tese.
- Às amigas e os amigos do Laboratório de Genética Evolutiva, Camila Borgonove, Camila Kokudai Santos, Rafaela Rossetti e Mateus Henrique Santos pela convivência agradável, suporte, companheirismo e amizade.
- Ao técnico Paulo Ricardo Epifânio pela sua boa vontade e paciência.
- À doutoranda Cíntia Graziela Santos pelo compartilhamento do conhecimento em morfometria.
- A minha família pelo incentivo e apoio sempre.
- A minha família brasileira adotiva, Marli Carmona Lavagnini e Antonio Carlos Lavagnini pelo apoio e ajuda.
- Ao Alejandro Grajales pela ajuda na Colômbia e pelo cuidado de Aron.
- A meus amigos colombianos na Colômbia, Harold, Leo, Sergio, Gloria, Alejandro, Marcela, Luís, e Sugui pela amizade destes longos anos, apoio e incentivo sempre. Em especial a Leo, por os bate-papos noturnos e pelas ajudas bioinformáticas e em outros temas.
- A meus amigos colombianos no Brasil, Claudia Charria, Lina Perilla, Alejandro Parra, Sergio Urquizo e em especial a Andrés Aristízabal.

- Omar Martinez, Vanessa Bonatii e Claudineia Pereira pela sua amizade e ajuda em vários momentos, em especial em meu processo de adaptação no Brasil.
- Ao Departamento de Genética da FMRP/USP, seus funcionários e secretárias pela ajuda com as pendências burocráticas e pela disponibilização dos recursos e espaço físico para a realização deste trabalho.
- Ao CNPQ-Programa PEC-PG pela concessão da bolsa de mestrado.
- Aos órgãos FAPESP, CNPq, FAEPA e USP, pelos auxílios indiretos.
- A todos meu muito obrigada.



**Todo es posible. Lo imposible simplemente nos lleva más tiempo.**  
**Tudo é possível. O impossível só leva mais tempo.**  
Dan Brown-Fortaleza Digital.

## RESUMO

BARRIOS-LEAL, D. Y. **História demográfica e estrutura de populações para a espécie cactófila *Drosophila meridionalis*** 2013. 105f. Tese de Mestrado. Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo.

*Drosophila meridionalis* é uma espécie endêmica da América do Sul, sendo amplamente distribuída na Costa Atlântica do Brasil. Com o objetivo de elaborar uma hipótese filogeográfica para esta espécie foram obtidas sequências do gene nuclear *period* e do gene mitocondrial *COI*. Foram calculados os índices de diversidade nucleotídica e realizados os testes: AMOVA, testes de neutralidade, a *Mismatch Distribution*, *Bayesian Skyline Plot*, *NCPA*. Foram obtidas três redes pelo gene *COI*, denominadas A (populações do interior), B (populações do litoral sul) e C (populações do litoral sudeste e oriental) e uma única rede obtida para o gene *period*, esta rede divide as populações em dois grupos sendo o primeiro congruente com a rede A e o segundo compreendendo as redes B e C, do gene *COI*. A AMOVA mostrou uma estruturação alta e significativa entre as populações do interior e o litoral para os dois genes ( $\Phi_{ct}=0,72$  gene *COI*;  $\Phi_{ct}=0,70$  gene *period*), que pode ser explicada pela presença de barreiras geográficas, como a Serra do Mar. Eventos de expansão populacional e de fluxo gênico restrito com isolamento por distância foram detectados nas populações do litoral e o interior respectivamente. A expansão da área de ocorrência de *D. meridionalis* provavelmente teve início com as populações do litoral do Rio Grande de Sul, em direção ao litoral de Santa Catarina com posterior colonização a longa distância dos estados de São Paulo, Rio de Janeiro e Bahia. Migrações assíncronas de indivíduos de populações litorâneas de São Paulo e Santa Catarina provavelmente colonizaram o interior de São Paulo, e a partir destas populações, se iniciara uma expansão populacional em direção ao sul pelo interior, colonizando o Paraná e Rio Grande do Sul. A análise bayesiana (*MCCT*) indicou que o tempo do ancestral comum mais recente ( $T_{MRCA}$ ) para todos os haplótipos de *D. meridionalis* é de 81.700 anos atrás, data que marca a separação das populações do interior e do litoral aproximadamente no final do Pleistoceno. Eventos similares têm sido sugeridos para explicar a distribuição geográfica de espécies do cluster *D. buzzatii*, que ocorrem em simpatria em grande parte com populações de *D. meridionalis*. Esta espécie, como as espécies do cluster *D. buzzatii*, apresentou indicativos de flutuações demográficas, podendo estar associadas à expansão e contração da distribuição da vegetação xerofítica, durante as oscilações paleoclimáticas do Pleistoceno.

## ABSTRACT

BARRIOS-LEAL, D. Y. **Demographic history and population structure for the cactophilic species *Drosophila meridionalis*** 2013. 105f. Tese de Mestrado. Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo.

*Drosophila meridionalis* is an endemic species of South America, being widely distributed in the Atlantic Coast of Brazil. Aiming to develop a phylogeographic hypothesis for this species, sequences of mitochondrial *COI* and nuclear *period* genes were obtained. The diversity indexes, AMOVA, neutrality tests, Mismatch Distribution, Bayesian Skyline Plot and NCPA were calculated. We obtained three networks for the *COI* gene, denominated A (inland populations), B (south coast populations) and C (eastern and southeastern coast populations) and a single network obtained for the *period* gene, this network divides the population into two groups, being the first congruent with the network A of the *COI* gene and the second comprising the networks B and C of the *COI* gene. The AMOVA results, showed a high and significant structuring among inland and coastal populations, for both genes ( $\Phi_{ct}=0,72$  *COI* gene;  $\Phi_{ct}=0,70$  *period* gene), that can be explained by the presence of geographical barriers, such as Serra do Mar. Population expansion events and restricted gene flow with isolation by distance events were detected in coastal and inland populations respectively. The expansion of the area of occurrence of *D. meridionalis* probably was initiated with the populations of the coast of Rio Grande do Sul, towards the coast of Santa Catarina with subsequent long-distance colonization of the states of São Paulo, Rio de Janeiro and Bahia. Asynchronic migrations of individuals from coastal populations of São Paulo and Santa Catarina probably colonized the inland of São Paulo, and from these populations, a population expansion towards the south through the inland was initiated, colonizing the states of Paraná and Rio Grande do Sul. The bayesian analysis (MCCT) indicated that the time of the most recent common ancestor ( $T_{MRCA}$ ) for all haplotypes of *D. meridionalis* is from 81,700 years ago, a date that marks the separation of inland and coastal populations approximately at the end of the Pleistocene. Similar events have been suggested to explain the geographic distribution of species of the cluster *D. buzzatii*, occurring in sympatry largely with populations of *D. meridionalis*. This species, as the cluster *D. buzzatii* species, presented indicatives of demographic fluctuations, which can be associated with the expansion and contraction of the distribution of xerophytic vegetation, during the paleoclimatic fluctuations of the Pleistocene.

## LISTA DE FIGURAS

**Figura 1.** Distribuição geográfica e representação diagramática dos cinco diferentes tipos (I-V) de cariótipos de *Drosophila meridionalis*; 2507 (Angra dos Reis - RJ); B97 (Arraial do Cabo - RJ); B51 (Seabra - BA); B53 (Milagres - BA) A96 (Serra do Cipó - MG); B93 (Itu - SP); B22 (La Vina - ARG); B31 (Puerto Tirol - ARG). Modificado de: (BAIMAI; SENE; PEREIRA, 1983). 24

**Figure 2.** Estrutura do gene *period* em *Drosophila melanogaster*. Os “introns” são representados por linhas e os exons por retângulos. O domínio PAS (região C2) está colorida com a cor cinza para os exons 3 e 4. No exon cinco estão indicadas às regiões não conservadas n2, n3 e n4, interespaçadas pelas regiões conservadas C3, C4 e C5 (COLOT; HALL; ROSBASH, 1988). A faixa listrada indica a região n2, formada pela região repetitiva Treonina-Glicina. O retângulo negro representa o domínio inibitório CCID (região C3 até região C4)(CHANG; REPERT, 2003). Abaixo da estrutura do gene, uma barra negra indica a região do exon cinco que foi analisada no presente trabalho. Tomada de (FRANCO et al., 2010). ..... 32

**Figura 3.** Mapa topográfico gerado com auxílio do programa DIVA-GIS v. 7.5.0, disponível em [www.diva-gis.org](http://www.diva-gis.org), com as localizações geográficas das populações de *Drosophila meridionalis* analisadas neste trabalho. A localidade de cada código esta descrita na Tabela 1. As divisões das populações nos grupos do interior e litoral que constam no mapa foram consideradas durante apresentação dos resultados e discussão. \* VIA e N45 são populações incluídas no grupo do litoral..... 38

**Figura 4.** Rede A de haplótipos não enraizada contendo 21 haplótipos do gene mitocondrial *COI* de populações de *Drosophila meridionalis*. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. Haplótipos hipotéticos, não amostrados, são representados por pequenos círculos pretos. As cores dentro de cada haplótipo representam as diferentes populações, sendo essas as mesmas apresentadas na Figura 1..... 52

**Figura 5.** Rede B de haplótipos não enraizada contendo 20 haplótipos do gene mitocondrial *COI* de populações de *Drosophila meridionalis*. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. Haplótipos hipotéticos, não amostrados, são representados por pequenos círculos pretos. As cores dentro de cada haplótipo representam as diferentes populações, sendo essas as mesmas apresentadas na Figura 1..... 54

**Figura 6.** Rede C de haplótipos não enraizada contendo 14 haplótipos do gene mitocondrial *COI* de populações de *Drosophila meridionalis*. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. Haplótipos hipotéticos, não amostrados, são representados por

pequenos círculos pretos. As cores dentro de cada haplótipo representam as diferentes populações, sendo essas as mesmas apresentadas na Figura 1..... 55

**Figura 7.** Distribuição dos 55 haplótipos do gene mitocondrial *COI* das 25 populações de *Drosophila meridionalis*. As populações estão representadas por círculos, cujo tamanho é proporcional ao número de indivíduos que os possuem. As cores na gama azul, roxo e cinza representam as populações da rede C, as cores da gama verde representam as populações da rede B e as cores da gama amarela e vermelha representam as populações da rede A..... 57

**Figura 8.** Representação gráfica das diferenças par a par entre as sequências (Mismatch Distribution) do gene mitocondrial *COI* dos indivíduos de *Drosophila meridionalis*. A: Indivíduos alocados no clado 1-11 da rede A. B: Indivíduos alocados no clado 3-3 da rede A. C: Indivíduos agrupados na rede B de haplótipos. D: Indivíduos agrupados na rede C de haplótipos. Linhas mais fortes representam as diferenças par a par observadas, enquanto linha mais tênue representa as diferenças par a par esperadas com eventos de expansão populacional. 63

**Figura 9.** Reconstrução da história demográfica, da espécie *Drosophila meridionalis* a partir das sequências do gene mitocondrial *COI*. A curva negra representa a variação na média do tamanho efetivo populacional, em relação ao tempo em anos atrás. As linhas azuis representam intervalos de confiança (95%). A linha vermelha vertical tracejada indica o tempo para o ancestral comum mais recente da amostra ( $T_{MRCA}$ ) da espécie. O tamanho populacional é medido em unidades de  $\log N_e \cdot \tau$ . ..... 64

**Figura 10.** Relações entre as linhagens e o  $T_{MRCA}$  de 192 indivíduos em 25 populações de *Drosophila meridionalis* baseado em análises de coalescência bayesiana (Maximum Clade Credibility Tree) das sequências do gene mitocondrial *COI*. O número superior corresponde ao suporte estatístico do nó (probabilidade a posteriori) e os números sublinhados correspondem ao  $T_{MRCA}$  com seus intervalos de confiança (95%). A escala de tempo está em milhões de anos..... 65

**Figura 11.** Rede de haplótipos não enraizada contendo 30 haplótipos do gene nuclear *period* de populações de *Drosophila meridionalis*. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. Haplótipos hipotéticos, não amostrados, são representados por pequenos círculos pretos. As cores dentro de cada haplótipo representam as diferentes populações, sendo essas as mesmas apresentadas na Figura 1..... 68

**Figura 12.** Distribuição dos 30 haplótipos do gene nuclear *period* nas 22 populações de *Drosophila meridionalis*. As populações estão representadas por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. As cores na gama azul, roxo, e cinza e verde representam as populações do litoral e as cores da gama amarela e vermelha representam as populações do interior. .... 69

**Figura 13.** Representação gráfica das diferenças par a par entre as sequências (Mismatch Distribution) do gene nuclear *period* dos indivíduos de *Drosophila meridionalis*. A: Indivíduos alocados no clado 3-3 B: Indivíduos agrupados na rede total de haplótipos. Linhas mais fortes representam as diferenças par a par observadas, enquanto linha mais tênue representa as diferenças par a par esperadas com eventos de expansão populacional..... 74

**Figura 14.** Modelo de probabilidade de presença atual da espécie *Drosophila meridionalis*, gerada pelo programa MaxEnt. Os diferentes tons de cinza representam a probabilidade da ocorrência da espécie..... 75

**Figura 15.** Modelo de probabilidade de presença da espécie *Drosophila meridionalis* para o último período glacial (21.000 anos atrás), gerada pelo programa MaxEnt, utilizando as projeções do CCSM. Os diferentes tons de cinza representam a probabilidade da ocorrência da espécie. .. 76

**Figura 16.** Modelo de probabilidade de presença da espécie *Drosophila meridionalis* para o último período glacial (21.000 anos atrás), gerada pelo programa MaxEnt, utilizando as projeções do MIROC. Os diferentes tons de cinza representam a probabilidade da ocorrência da espécie. 76

## LISTA DE TABELAS

- Tabela 1.** Localidades das populações de *Drosophila meridionalis* amostradas. Para cada localidade estão relacionados: CG = Coordenadas Geográficas, CL = Código de Localidade, N1 = Amostras do gene mitocondrial COI, N2 = Amostras do gene nuclear *period*. ..... 37
- Tabela 2.** Índices de diversidade genética e haplótipos de populações de *Drosophila meridionalis*, utilizando sequências do gene mitocondrial COI..... 50
- Tabela 3.** Resultados significativos da NCPA para a rede A, obtida a partir de amostras do gene mitocondrial COI de populações de *Drosophila meridionalis*, de acordo com a chave de inferências filogeográficas de Templeton (2011), disponível em <http://darwin.uvigo.es/software/geodis.html> ..... 51
- Tabela 4.** Resultados significativos do agrupamento de clados não enraizados da rede C, obtida a partir de amostras do gene mitocondrial COI de populações de *Drosophila meridionalis*, de acordo com a chave de inferências filogeográficas de Templeton (2011), disponível em <http://darwin.uvigo.es/software/geodis.html> ..... 56
- Tabela 5.** Análise de Variância Molecular (AMOVA), para as populações de *Drosophila meridionalis*, baseadas nas sequências do gene mitocondrial COI. .... 58
- Tabela 6.** Valores de  $F_{ST}$  calculados entre pares de populações de *Drosophila meridionalis*, para o gene mitocondrial COI. Os asteriscos representam os valores de p significativos ( $p \leq 0.05$ ). .... 59
- Tabela 7.** Teste de Mantel para as redes de haplótipos das populações de *Drosophila meridionalis*, baseadas nas sequências do gene mitocondrial COI. .... 61
- Tabela 8.** Resultados dos testes de neutralidade ( $F_s$  de Fu, D de Tajima e R2) e *Mismatch Distribution* para populações de *Drosophila meridionalis*, utilizando sequências do gene mitocondrial COI..... 62
- Tabela 9.** Índices de diversidade genética e haplótipos de populações de *Drosophila meridionalis*, utilizando sequências do gene nuclear *period*. .... 66
- Tabela 10.** Resultados significativos do agrupamento de clados não enraizados da rede de haplótipos, obtida a partir de amostras do gene nuclear *period* de populações de *Drosophila meridionalis*, de acordo com a chave de inferências filogeográficas de Templeton (2011), disponível em <http://darwin.uvigo.es/software/geodis.html>..... 67
- Tabela 11.** Análise de Variância Molecular (AMOVA) para as populações de *Drosophila meridionalis*, baseadas nas sequências do gene nuclear *period*. .... 70

<b>Tabela 12.</b> Teste de mantel para a rede de haplótipos das populações de <i>Drosophila meridionalis</i> , baseadas nas sequências do gene nuclear <i>period</i> .....	70
<b>Tabela 13.</b> Valores de $F_{ST}$ calculados entre pares de populações de <i>Drosophila meridionalis</i> , para o gene <i>period</i> . Os asteriscos representam os valores de p significativos ( $p \leq 0.05$ ). .....	71
<b>Tabela 14.</b> Resultados dos testes de neutralidade ( $F_s$ de $F_u$ , $D$ de Tajima e $R_2$ ) e <i>Mismatch Distribution</i> para populações de <i>Drosophila meridionalis</i> , aplicados às sequências do gene nuclear <i>period</i> . .....	73



## SUMÁRIO

I. INTRODUÇÃO .....	19
1.0 Considerações iniciais.....	19
1.1 Floresta Atlântica e sua regionalização ecológica e biogeográfica. ....	20
1.2 Modelo Biológico: <i>Drosophila meridionalis</i> (grupo <i>D. repleta</i> , subgrupo <i>D. mulleri</i> ). ....	23
1.3 Teoria da Coalescência e Filogeografia.....	26
1.4 Marcadores Moleculares.....	30
1.4.1 Gene Citocromo Oxidase I ( <i>COI</i> ). ....	30
1.4.2 Gene nuclear <i>period</i> . ....	31
II. OBJETIVOS .....	35
III. MATERIAL E MÉTODOS.....	36
1.0 Material Biológico.....	36
1.1 Obtenção de sequências.....	38
1.2 Isolamento, amplificação e sequenciamento do gene mitocondrial <i>COI</i> .....	39
1.3 Isolamento, amplificação e sequenciamento do gene nuclear <i>period</i> .....	40
2.0 Análises das sequências.....	41
2.1 Inferências filogeográficas e estrutura populacional. ....	41
2.2 Testes de Neutralidade e a <i>Mismatch Distribution</i> . ....	44
2.3 Inferências Bayesianas.....	47
3.0 Modelagem de distribuição de espécies.....	48
IV. RESULTADOS .....	49
1.0 Análises das sequências do gene mitocondrial <i>COI</i> . ....	49
1.1 <i>Nested Clade Phylogeographic Analysis</i> (NCPA).....	51
1.2 AMOVA, testes de Mantel, testes de neutralidade e <i>Mismatch Distribution</i> .....	57
1.3 Análises Bayesianas.....	64
2.0 Análises das sequências do gene nuclear <i>period</i> . ....	65
2.1 <i>Nested Clade Phylogeographic Analysis</i> (NCPA).....	67
2.2 AMOVA, testes de Mantel, testes de neutralidade e <i>Mismatch Distribution</i> .....	69
3.0 Modelagem de Nicho Ecológico.....	74
V. DISCUSSÃO.....	77

1.0 Gene mitocondrial <i>COI</i> : variabilidade e análises filogeográficas para <i>Drosophila meridionalis</i> .....	77
1.1 Rede A.....	78
1.2 Rede B.....	79
1.3 Rede C.....	80
2.0 Gene nuclear <i>period</i> : variabilidade e análises filogeográficas para <i>Drosophila meridionalis</i> .....	82
3.0 Hipótese filogeográfica para <i>Drosophila meridionalis</i> .....	85
4.0 Análises Bayesianas.....	88
5.0 Modelagem de Nicho Ecológico.....	90
VI CONCLUSÕES.....	92
VII. REFERÊNCIAS BIBLIOGRÁFICAS.....	93

## I. INTRODUÇÃO

### 1.0 Considerações iniciais.

No processo de evolução das espécies estão envolvidos aspectos da ecologia e da história natural, assim como eventos históricos. Neste processo, entre as questões históricas pode-se descrever como os eventos históricos influenciaram a distribuição geográfica e genética das populações, determinando certamente a partir deste ponto as potencialidades evolutivas e adaptativas das populações de uma espécie. Em termos históricos, as influências dos ciclos glaciais e interglaciais do Pleistoceno, que começaram há aproximadamente 2,5 milhões de anos (HEWITT, 2011), tem sido muito discutido como um fator influente na distribuição geográfica atual de grupos animais e vegetais (CARNAVAL; HICKERSON; HADDAD, 2009; PENNINGTON et al., 2000; PRADO; GIBBS, 1993; THOMÉ et al., 2010).

A América do Sul é caracterizada pela combinação de diversidade climática e ambiental que permitem a grande diversificação da sua cobertura florestal, englobando desde as florestas úmidas até regiões semi-áridas (HUECK, 1972). Entre as duas principais florestas úmidas da América do Sul, Floresta Amazônica e Floresta Atlântica, há uma faixa de vegetação sazonalmente seca e aberta, orientada na direção nordeste-sudoeste, conhecida como Floresta Tropical Sazonalmente Seca (FTSS) (AB´SABER, 2003; HUECK, 1972). Tem sido proposto que tanto as florestas úmidas quanto a vegetação sazonalmente seca tiveram sua distribuição modificada durante o período Quaternário, como consequência das mudanças climáticas globais (PENNINGTON et al., 2000; PRADO; GIBBS, 1993).

A hipótese mais tradicional sugere que, durante as condições frias e secas dos períodos glaciais nos Neotrópicos, foi favorecida a expansão geográfica da vegetação xerofítica sobre as florestas úmidas, que foram reduzidas para locais chamados de refúgios do Quaternário (HAFFER, 1969) e durante os períodos interglaciais, com condições quentes e úmidas, a situação seria revertida (AB´SABER, 1977; PENNINGTON et al., 2000). Diversos estudos apoiam as suposições citadas acima, incluindo estudos geomorfológicos (AB´SABER, 1977; WANG et al., 2004), palinológicos (DE OLIVEIRA; BARRETO; SUGUIO, 1999; LEDRU, 1993), espeleológicos (AULER et al., 2004; WANG et al., 2004), paleomodelagem (CARNAVAL;

HICKERSON; HADDAD, 2009; CARNAVAL; MORITZ, 2008), e estudos biogeográficos (CARNAVAL; BATES, 2007; MORAES et al., 2009; PENNINGTON et al., 2000; QUIJADA-MASCAREÑAS et al., 2007).

Neste sentido, para compreender a dinâmica populacional e a atual distribuição geográfica e genética das espécies neotropicais que tiveram associação ecológica obrigatória com domínios úmidos ou secos, é crucial considerar o impacto destas flutuações climáticas cíclicas na história destas espécies.

### **1.1 Floresta Atlântica e sua regionalização ecológica e biogeográfica.**

A Floresta Atlântica originalmente estende-se por 1.300.000km<sup>2</sup> ao longo do litoral brasileiro até o Paraguai e a Argentina. Atualmente é aceita a classificação de que a Floresta Atlântica brasileira “*sensu lato*” é a área que abrange todas as fisionomias florestais que ocorrem no corredor seco do leste que passa pelo nordeste, sul e sudeste do Brasil, leste do Paraguai e nordeste de Argentina. As fisionomias florestais são: as florestas ombrófila densa, ombrófila aberta, ombrófila mista, a floresta estacional decídual e semidecídual, os mangues, as restingas e os campos de altitude (LEDRU et al., 2007; OLIVEIRA-FILHO; FONTES, 2000).

Na região próxima ao litoral, está a floresta ombrófila densa, caracterizada pela ausência de estação seca. As encostas do Planalto Atlântico, conhecidas genericamente como Serra do Mar ao sul e Serra da Mantiqueira mais ao interior, retêm boa parte da umidade vinda do oceano, provocando as chuvas orográficas, principais mantenedoras das condições climáticas dessa floresta densa. Adjacentes à floresta ombrófila existem ainda ecossistemas mais relacionados às características físicas típicas de ambientes costeiros, como os manguezais e as restingas. Em direção ao interior, observa-se a redução da umidade dos seus ecossistemas, existindo um gradiente fisionômico da floresta estacional semidecidual, que aos poucos se torna na zona de contato com o domínio do Cerrado. Mais ao sul, principalmente no estado do Paraná, a floresta do interior (conhecida como Floresta Mista) também exhibe características distintas, como, por exemplo, a presença marcante do pinheiro-do-Paraná (*Araucaria angustifolia*), espécie cujo gênero é característico de regiões mais frias ou temperadas da América do Sul. Mesmo nas encostas voltadas para o mar, a zonação altitudinal das serras também é refletida em fitofisionomias distintas, observando-se desde as florestas de planície, passando para

Submontanha e Montanha; esta última acima de 800m de altitude, e que em muitos pontos transforma-se em vegetações abertas, como campos de altitude.

A heterogeneidade de ambientes, relacionada às características geomorfológicas, ao longo da área de distribuição da Floresta Atlântica, pode ser congruente com a diferenciação populacional e distribuição de diferentes grupos taxonômicos, auxiliando no entendimento dos processos envolvidos.

O mecanismo mais comumente utilizado para explicar a diversificação na Floresta Atlântica, assim como no continente sul-americano em geral, é o isolamento em áreas de hábitat estável, durante as flutuações climáticas do Quaternário – hipótese de refúgios do Pleistoceno. Esta hipótese sugere que ciclos sucessivos do clima durante o Quaternário (especialmente durante o Pleistoceno no LGM- *Last Glacial Maximum*) originaram a retração ou a expansão dos limites geográficos das espécies permitindo a diversificação destas, dirigido pelos ciclos (HAFFER, 1969). Propõe ainda que as extensas áreas da floresta úmida existiam como refúgios durante os períodos secos e frios (glaciais) no final do Terciário e do Quaternário, e deram origem a muitas espécies existentes e subespécies de plantas e animais (HAFFER; PRANCE, 2001). Assim, altos níveis de endemismo e diversidade genética evidenciaram áreas de alta estabilidade ou floresta no passado (refúgios), e em contraste a baixa diversidade e endemismo são assinaturas moleculares de expansão recente em áreas instáveis (de não refúgio) (CARNAVAL; MORITZ, 2008).

Os modelos florestais baseados em registros de pólen sugerem uma grande área de estabilidade florestal histórica no corredor central no estado da Bahia e refúgios pequenos localizados ao longo da costa brasileira, além de uma área ao norte do rio Paraíba, chamado “refúgio Pernambuco” e outros prováveis fragmentos pequenos ao sul do Rio Doce. Estes refúgios correspondem aos centros atuais de endemismo em vários táxons e os padrões de diversidade do mtDNA em um subconjunto de espécies examinadas (CARNAVAL; HICKERSON; HADDAD, 2009; CARNAVAL; MORITZ, 2008).

Em diversos grupos taxonômicos têm encontrado padrões filogeográficos que são congruentes com as predições da hipótese de refúgios do Pleistoceno na Floresta Atlântica (CABANNE et al., 2008; D’HORTA et al., 2011; PAVAN et al., 2011). Entretanto, evidências

limitadas foram encontradas, de que a hipótese do refúgio teve um papel na diversificação das espécies do grupo de espécies de *Rhinella crucifer* (THOMÉ et al., 2010).

Entre outras hipóteses para explicar a origem da diversidade da Floresta Atlântica encontra-se:

-A hipótese neotectônica: sugere que a elevação da costa leste brasileira no Terciário teria produzido modificações geográficas, hidrográficas e climáticas na região da Floresta Atlântica, o que pode ter causado a fragmentação da Floresta com divergência na fauna associada (GRAZZIOTIN et al., 2006). Os padrões filogenéticos e as distribuições geográficas em peixes sugerem uma íntima associação com a história geológica da margem continental passiva da América do Sul, desde o Cretáceo até os dias atuais (RIBEIRO, 2006). As falhas geológicas ativas, muitas das quais apresentam evidência de ruptura na superfície no Quaternário, podem fornecer explicações alternativas para a diversificação da biota da Floresta Atlântica (RICCOMINI; ASSUMPÇÃO, 1999)

-Rios como barreiras: Os sistemas fluviais desempenham um papel importante nos processos evolutivos relacionados com a diversificação biológica, já que podem funcionar como barreiras primárias ou secundárias para o fluxo gênico, modelando a distribuição atual da biota e diferenciando linhagens e espécies. Os principais rios na diversificação da Floresta Atlântica do Brasil são: São Francisco, Jequitinhonha, Doce e Paranapanema (DANTAS MENDONÇA; CABANNE; SANTOS RODRIGUES, 2011). Estudos no grupo *Rhinella crucifer*, apresentam linhagens divergentes espacialmente concordantes com Rio Doce (THOMÉ et al., 2010). A história evolutiva do gênero *Calyptommatus* (lagarto) sugere um padrão complexo e alternativo de cladogênese envolvendo áreas arenosas entre as margens opostas do rio São Francisco, sugerindo fortemente que o rio levou à especiação (SIEDCHLAG et al., 2010).

- Gradientes ecológicos: A Floresta Atlântica está circundada por biomas de vegetação aberta (as FTSS e o Cerrado) e devido à existência de uma transição gradual do clima quente e úmido da Floresta Atlântica para o clima seco dos biomas de vegetação aberta, muitos organismos associados com as Florestas úmidas também são encontrados e misturados nos biomas abertos. Tanto a Floresta Atlântica como as regiões vizinhas, apresentam características diferentes, e espera-se, portanto, que os táxons divergissem entre as regiões, como resposta às

pressões seletivas ecológicas. No entanto, apesar do cenário ser pouco provável, estudos tem abordado o problema da divergência de espécies em diferentes gradientes ecológicos neste bioma (DANTAS MENDONÇA; CABANNE; SANTOS RODRIGUES, 2011).

Portanto, uma visão realista do processo de diversificação da Floresta Atlântica deve considerar os múltiplos mecanismos, em múltiplos *taxa*, operando em diferentes escalas temporais e espaciais.

### **1.2 Modelo Biológico: *Drosophila meridionalis* (grupo *D. repleta*, subgrupo *D. mulleri*).**

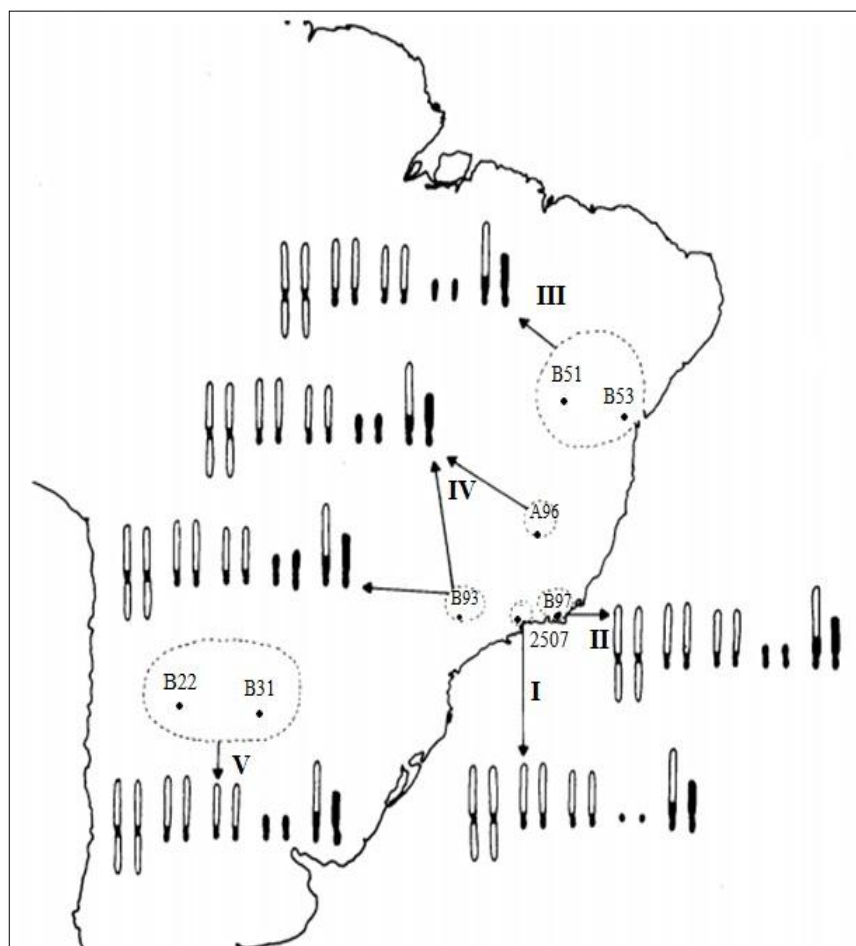
O grupo *repleta* é classificado como um dos maiores do gênero *Drosophila* com aproximadamente 100 espécies e é naturalmente endêmica do continente Americano. Esse grupo é composto por seis subgrupos: *Drosophila mulleri*, *Drosophila hydei*, *Drosophila mercatorum*, *Drosophila repleta*, *Drosophila inca* e *Drosophila fasciola* (RAFAEL; ARCOS, 1989; VILELA, 1983; WASSERMAN, 1982), que, por sua vez, estão subdivididos em complexos e *clusters*.

O subgrupo *Drosophila mulleri* tem aproximadamente 50 espécies (DURANDO et al., 2000; WASSERMAN, 1982). Os estudos citológicos de cromossomos politênicos, morfologia da genitália e sequências de DNA têm sido usados para inferir relações filogenéticas entre as espécies do subgrupo *D. mulleri* sendo base para separar vários complexos e *cluster* de espécies (OLIVEIRA et al., 2005; VILELA, 1983; WASSERMAN, 1982).

O complexo *D. meridina*, descrito dentro do subgrupo *D. mulleri* é composto por três espécies alopátricas: *D. meridiana*, encontrada no México e Estado Unidos; *D. promeridiana*, na Colômbia e Venezuela; e *D. meridionalis*, no Brasil, Paraguai e Argentina (VILELA, 1983; WASSERMAN, 1992). *Drosophila meridionalis* (WASSERMAN, 1962) é uma espécie cactófila, que tem sido associada com os cactos *Opuntia ficus-indica*, *Opuntia vulgaris*, *Cereus fernambucesis*, *Cereus peruvianus*, (PEREIRA; VILELA; SENE, 1983) *Cereus hindmannianus*, *Opuntia monacantha* (RUIZ et al., 2000) e *Pilosocereus machrisii* (COSTA; KUHN; SENE, 2000)

Estudos baseados em cariótipos de *D. meridionalis* de diferentes áreas de sua distribuição geográfica indicam que há diferenciação populacional (Figura 1) (BAIMAI; SENE; PEREIRA,

1983; BAIMAI, 1998). A diferenciação cariotípica ocorre principalmente devido às alterações que envolvem os grandes blocos de heterocromatina constitutiva presentes no cromossomo seis e/ou microcromossomos, sugerindo uma origem alopatrica dos cinco cariótipos descritos. Esta estruturação, de acordo com os autores pode ser resultado de uma história evolutiva concordante com a dinâmica das oscilações climáticas no continente sul americano durante o Último Máximo Glacial, onde as mudanças climáticas parecem ter levado a uma distribuição fragmentada da vegetação e particularmente das populações de cactos.



**Figura 1.** Distribuição geográfica e representação diagramática dos cinco diferentes tipos (I-V) de cariótipos de *Drosophila meridionalis*; 2507 (Angra dos Reis - RJ); B97 (Arraial do Cabo - RJ); B51 (Seabra - BA); B53 (Milagres - BA) A96 (Serra do Cipó - MG); B93 (Itu - SP); B22 (La Vina - ARG); B31 (Puerto Tirol - ARG). Modificado de: (BAIMAI; SENE; PEREIRA, 1983)



Análises da corte sonora de indivíduos de *D. meridionalis* em populações do sul e sudeste do Brasil mostraram que, apesar das diferenças cariotípicas das populações, não há variação significativa para a maioria dos parâmetros da corte sonora, sugerindo que as populações constituem apenas uma espécie com uma variabilidade cromossômica interpopulacional (COSTA; KUHN; SENE, 2000).

A espécie *D. meridionalis* e as espécies do cluster *D. buzzatii* são cactófilas, apesar da ausência de monofilia entre ambos os grupos. Sendo assim, as populações de *D. meridionalis* geralmente ocorrem em simpatria com as populações das espécies cluster *D. buzzatii*, muitas vezes com adultos emergindo do mesmo cladódio em decomposição (SENE; PEREIRA; VILELA, 1982; VILELA, 1983).

A distribuição geográfica das diferentes espécies cactófilas associada ao seu cacto hospedeiro é um ponto de vista evolutivo relevante, pois, é através desta forte associação que estas moscas têm sido capazes de invadir regiões áridas, desertos e semidesertos do Novo Mundo (OLIVEIRA et al., 2005). As espécies cactófilas têm especificidade ecológica, devido à oviposição e ao desenvolvimento larval das espécies que ocorrem exclusivamente em tecidos de cactos em decomposição (PEREIRA; VILELA; SENE, 1983) e por consequência estas espécies são um excelente modelo biológico para estudos evolutivos (MANFRIN; SENE, 2006).

As principais características da especificidade biológica destas *Drosophila* com os cactos são: 1. A distribuição geográfica das espécies cactófilas é determinada pela ocorrência de espécies de cactáceas na América do Sul, devido a esta associação, sua distribuição geográfica compreende a diagonal seca, que inclui os biomas da Caatinga, Cerrado e Chaco, que se situam entre a Floresta Amazônica e a Floresta Atlântica (MANFRIN; SENE, 2006); 2. Os diferentes cactos apresentam diferentes composições químicas e microflora (STARMER et al., 1986), o que produz ambientes ecológicos distintos para os quais as *Drosophila* cactofílicas devem estar adaptadas (MATZKIN, 2008); 3. Eventos demográficos de expansão populacional e os processos de diferenciação podem estar relacionados a eventos não sincrônicos de expansão e retração da vegetação xerófita na América do Sul, durante alterações climáticas causadas por eventos paleoclimáticos (DE BRITO; MANFRIN; SENE, 2002; FRANCO; MANFRIN, 2013; MORAES et al., 2009; SENE; PEREIRA; VILELA, 1988).

Sendo assim, o estudo filogeográfico e da estruturação populacional da espécie cactófila *D. meridionalis* pode contribuir para entender a influência das flutuações climáticas do Quaternário na distribuição e composição da diversidade genética dos táxons neotropicais e também a corroborar os padrões biogeográficos existentes na América do Sul para outras espécies do gênero *Drosophila*.

### 1.3 Teoria da Coalescência e Filogeografia.

A teoria da coalescência, desenvolvida por Kingman (1982), é um modelo retrospectivo da genética de populações, que traça todos os alelos de um gene dentro de uma amostra populacional até um ancestral comum mais recente (MRCA - *Most Recent Common Ancestor*). Kingman (1982) descreveu a evolução visualizando-a para trás no tempo, com base nos pressupostos do tamanho populacional constante e acasalamento aleatório. A coalescência não modela a mutação, no entanto, dado um modelo mutacional, os dois podem ser combinados para expressar a probabilidade de um coalescente com mutações (TEMPLETON, 2006). A teoria da coalescência contribuiu com a fusão conceitual e analítica das áreas da genética de populações e da filogenética, sendo a árvore de genes o foco de estudo em ambas as áreas. No entanto, a filogenia tradicional busca estimar uma árvore e usá-la para deduzir as relações evolutivas, enquanto a genética de populações tradicional observa a árvore como um resultado aleatório de um processo genético populacional (NIELSEN; BEAUMONT, 2009). O enfoque na história da espécie, sobretudo no passado biogeográfico, é a grande diferença entre a abordagem filogeográfica e a abordagem da genética de populações, que fornece apenas informações sobre a estrutura das populações (KNOWLES, 2009; KNOWLES; MADDISON, 2002).

A filogeografia é definida como o estudo dos princípios e os processos que determinam a distribuição geográfica de linhagens genealógicas (AVISE, 2000, 2009; AVISE et al., 1987). A distribuição espacial e temporal da variabilidade genética de uma espécie é influenciada por eventos recorrentes, tais como a mutação, o fluxo gênico, a deriva genética e a seleção natural e eventos históricos, como a expansão da área de ocorrência e a divisão populacional (AVISE, 2000; AVISE et al., 1987). A análise e interpretação da distribuição das linhagens requerem o processamento conjunto de informações de uma série de disciplinas, incluindo a sistemática

filogenética, a genética de populações, a etologia, a demografia, a paleontologia, a geologia e a geografia histórica (AVISE, 2000).

Os estudos filogeográficos fornecem informações detalhadas de uma espécie, sobre como os eventos geológicos, as influências ambientais e os fatores geográficos interagem com os aspectos da ecologia e da história natural da espécie (KNOWLES, 2009). Quando essas relações são integradas com outros tipos de informação (por exemplo, dados fenotípicos), a informação demográfica detalhada extraída a partir dos dados genéticos pode proporcionar um cenário para uma série de perguntas evolutivas e ecológicas que não poderiam ser resolvidas sem um conhecimento da história filogeográfica (KNOWLES, 2009; TEMPLETON, 2009).

A análise filogeográfica básica geralmente consiste na estimativa de uma árvore ou de uma rede. Os ramos da árvore são relacionados e podem ser interpretados como evidência de ocorrência específica de eventos demográficos históricos em um contexto geográfico (AVISE, 2000; AVISE et al., 1987; NIELSEN; BEAUMONT, 2009). A filogeografia tem em geral duas linhas (*Nested Clade Phylogeographic Analysis* e filogeografia estatística) que diferem quanto à forma de analisar as inferências (HEY; MACHADO, 2003). Em ambas as linhas, as inferências de eventos e processos associados à distribuição dos organismos é feita depois de estimar uma árvore de genes ou uma rede de haplótipos (NIELSEN; BEAUMONT, 2009; TEMPLETON, 2009).

Uma destas linhas é representada pela *Nested Clade Phylogeographic Analysis* (NCPA), que quantifica as associações entre a geografia e uma dada árvore de haplótipos (TEMPLETON; ROUTMAN; PHILLIPS, 1995; TEMPLETON, 2006). A NCPA possui limitações, entre as quais se destaca a sensibilidade à resolução genética dos marcadores para detectar eventos ou processos evolutivos recentes e as interpretações biológicas limitadas em virtude do tamanho da amostra e do número de localidades amostradas (TEMPLETON, 2004b). No entanto, as principais críticas à NCPA baseiam-se na observação de um excesso de inferências falso-positivas de até 70% em avaliações do método usando dados simulados de populações panmíticas, em que a história demográfica e a estrutura populacional eram previamente conhecidas. As inferências falso-positivas mais comuns foram: “fluxo gênico restrito com isolamento por distância” e “expansão

populacional contínua” (KNOWLES, 2008; KNOWLES; MADDISON, 2002; PANCHAL; BEAUMONT, 2007).

Entretanto Templeton (2008) avaliou a taxa de inferências equivocadas pela NCPA com 150 controles positivos, que abrangia uma ampla gama de organismos, distribuições espaciais e desenhos de amostragem, e todos compartilhavam duas características: eram eventos evolutivos verdadeiros e desenhos de amostragem reais. Esses controles positivos mostraram que a ocorrência de falsos positivos e negativos, em populações reais, é muito menor do que em populações simuladas. O autor também mostrou que essa alta taxa de inferências errôneas, apresentada nos estudos de simulações prévios (KNOWLES; MADDISON, 2002; PANCHAL; BEAUMONT, 2007), era resultado de artefatos, devido às premissas não realísticas dos modelos matemáticos utilizados nas simulações, sugerindo que muitas vezes eles não são coincidentes com a complexidade das populações naturais (TEMPLETON, 2004b, 2008). Templeton (2008) argumenta que os falsos positivos observados podem ser minimizados com a correção do nível de significância com testes múltiplos de Bonferroni, que consiste na divisão dos valores de P dos cladogramas significativos pelo número de testes realizados. Os falsos positivos e falsos negativos também são minimizados com uma validação cruzada, na qual uma inferência deve ser confirmada por outro conjunto de dados (múltiplas regiões de DNA) ou por uma subamostra dos dados originais (TEMPLETON 2006, 2008, 2009, 2010).

As inferências da NCPA dividem-se em três grupos: o isolamento por distância, a fragmentação alopátrica e a expansão de distribuição (TEMPLETON, 2006), que podem ser testadas independentemente com outras análises (GARRICK et al., 2008). Assim, a fragmentação alopátrica poderia ser testada com métodos filogenéticos (Máxima-Verossimilhança, Neighbour-Joining e Parcimônia), mas se a fragmentação é um processo recente pode tratar-se de uma estruturação das populações, e isto poderia ser verificado com uma Análise de Variância Molecular (AMOVA) (EXCOFFIER; SMOUSE; QUATTRO, 1992). A expansão da distribuição das populações pode ser testada usando a *Mismatch Distribution Analysis* (ROGERS; HARPENDING, 1992) e os testes de neutralidade. Por último a inferência do isolamento por distância, sugere que a associação entre a variação genética e a geografia (TEMPLETON, 2006) pode ser verificada com o teste de Mantel.

A outra linha é o enfoque da “filogeografia estatística” que estima a inferência demográfica levando em conta a estocasticidade dos processos genéticos (KNOWLES; MADDISON, 2002). Esta abordagem é baseada na coalescência e na variância mutacional permitindo estimar parâmetros genéticos populacionais e testar hipóteses. Inclui métodos que calculam a probabilidade dos dados, usando abordagens baseadas em verossimilhança ou inferência bayesiana. Os programas computacionais usados para estimar as inferências filogeográficas (revisar (EXCOFFIER; HECKEL, 2006)), embora sejam computacionalmente intensivos fornecem estimativas detalhadas do passado, como o tamanho populacional, o crescimento, a subdivisão, os padrões de fluxo gênico e o tempo de divergência (KNOWLES, 2009; NIELSEN; BEAUMONT, 2009).

A função de verossimilhança é a probabilidade de se obter os dados (ou qualquer função proporcional a esta probabilidade). Os “dados” referem-se aos tipos de variantes genéticas e as suas frequências em uma amostra. Em uma notação matemática, a probabilidade é dada por  $p(X | \Theta)$ , onde  $X$  são os dados,  $\Theta$  é um vetor que contém todos os parâmetros de interesse, e “|” é lido como “probabilidade condicional”, indicando que a probabilidade calculada é para um determinado valor dos parâmetros. Os parâmetros podem ser: a taxa de migrantes, o tamanho efetivo populacional, e/ou as taxas de crescimento populacional. Esses parâmetros podem ser estimados, por exemplo, com “*maximum likelihood*” (NIELSEN; BEAUMONT, 2009). Na inferência bayesiana, a distribuição *a priori* ( $p(\Theta)$ ) dos parâmetros é assumida o que permite o cálculo de uma distribuição *a posteriori*  $p(\Theta | x)$  dos parâmetros, que é a distribuição condicional dos parâmetros dado os dados (BEAUMONT; RANNALA, 2004; NIELSEN; BEAUMONT, 2009).

As duas abordagens são baseadas em Monte Carlo via Cadeias de Markov (MCMC, do inglês *Markov Chain Monte Carlo*). Assim a MCMC é um tipo especial de processo estocástico que se baseia em simulações (Cadeia de Markov), para estudar as propriedades de uma distribuição de probabilidade que não pode ser tão facilmente estudada usando métodos analíticos. As Cadeias de Markov são um modelo apropriado para modelar sequências de variáveis aleatórias (como são os pares de bases de nucleotídeos no DNA), onde a distribuição de probabilidades de estados futuros é completamente determinada pelo estado atual em qualquer

ponto da cadeia. A ideia básica da integração de Monte Carlo é gerar números aleatórios que são simulados pela Cadeia de Markov (BEAUMONT; RANNALA, 2004).

Os padrões e as datações de eventos históricos de grupos taxonômicos é o principal interesse nas análises filogeográficas, mas uma única análise não pode fornecer inferências para todas estas perguntas concomitantemente, e cada abordagem possui suas premissas e limitações (GARRICK et al., 2008; NIELSEN; BEAUMONT, 2009). Vários estudos nos últimos anos em diferentes grupos taxonômicos têm usado as duas abordagens para fazer inferências filogeográficas, por exemplo, no sapo *Leptodactylus validus* (CAMARGO; HEYER; DE SÁ, 2009), nos peixes *Gymnothorax undulatus* e *G. flavimarginatus* (REECE et al., 2010), no leão marinho *Eumetopias jubatus* (PHILLIPS et al., 2011).

#### **1.4 Marcadores Moleculares.**

O entendimento da história evolutiva das populações vem aumentando rapidamente pelo uso dos dados moleculares, que tem revelado a distribuição da variação genética dentro e entre populações (AVISE, 2009; ZINK; BARROWCLOUGH, 2008). Nas últimas décadas marcadores como as isoenzimas e as enzimas de restrição foram substituídos por sequências de DNA mitocondrial (mtDNA), cloroplastidial (cpDNA) e nuclear (nuDNA) (ZINK; BARROWCLOUGH, 2008). Diversos estudos publicados obtiveram conclusões sobre a história populacional, os padrões de fluxo gênico, a estrutura genética, as reconstruções das relações filogenéticas, com base na variação das sequências de mtDNA, cpDNA e nuDNA (AGARWAL; SHRIVASTAVA; Á, 2008; AVISE, 2009; BEAUMONT et al., 2010; ZINK; BARROWCLOUGH, 2008)

##### **1.4.1 Gene Citocromo Oxidase I (COI).**

Ao longo do tempo evolutivo do mtDNA, este genoma citoplasmático perdeu a maior parte do seu DNA para o núcleo e atualmente é representado em células animais por cerca de 37 genes todos ligados ao longo de uma molécula circular fechada, com aproximadamente 17.000 pares de nucleotídeos de comprimento (AVISE, 2009). O mtDNA é altamente compacto, isto significa que não possui introns ou longas extensões intergênicas de sequência espaçadoras não codificantes que caracterizam os genomas nucleares. Apresentam também uma suposta

neutralidade e sua acumulação de substituições de nucleotídeos é mais rápida do que o DNA de um gene nuclear típico (AVISE, 2009; HICKERSON et al., 2010; ZINK; BARROWCLOUGH, 2008).

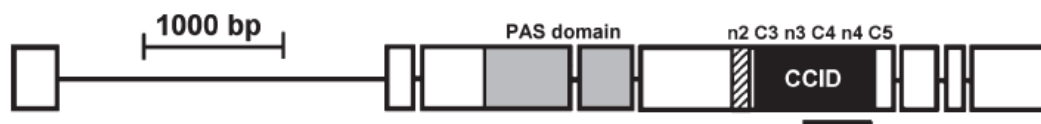
O mtDNA apresenta transmissão materna da molécula, o que representa um caráter não recombinante transmitido por meio da fêmea. Assim os indivíduos retém em seu mtDNA a história matrilinea, fato que faz dessa molécula um importante marcador para o estabelecimento de genealogias de genes e para o estudo filogeográfico (AVISE, 2000, 2009). Outra característica importante é o tamanho populacional efetivo menor que o mtDNA apresenta em relação aos genes nucleares. De acordo com Templeton (2006), o tempo de coalescência é dado por  $2xN_e$ , em que  $x$  é o nível de poliploidia e  $N_e$  é o tamanho efetivo populacional. Desse modo, para genes autossômicos ( $x=2$ ) e ligados ao cromossomo X ( $x=1.5$ ), o tempo de coalescência é  $4N_e$  e  $3N_e$ , respectivamente. Para o mtDNA ( $x=1$ ), o tempo de coalescência é  $2N_e/2$  ou simplesmente  $N_e$  (tamanho populacional das fêmeas). Dessa maneira, os tempos de coalescência para os genes mitocondriais são cerca de três a quatro vezes menores em comparação aos genes ligados ao X e autossômicos, respectivamente (TEMPLETON, 2006).

O gene mitocondrial *Citocromo Oxidase I (COI)* codifica a subunidade I do citocromo C, a qual é parte de um complexo gênico codificante de proteínas envolvidas na cadeia respiratória (SIMON et al., 1994). Esse gene é um marcador molecular informativo para estudos filogenéticos (FRANCO et al., 2010; GLEASON; GRIFFITH; POWELL, 1998; GOTO; KIMURA, 2001; LEWIS; BECKENBACH; MOOERS, 2005; MANFRIN; BRITO; SENE, 2001; NAGARAJA; RANGANATH, 2004; YANG et al., 2012), populacionais e filogeográficos em insetos (DE BRITO; MANFRIN; SENE, 2002; FRANCO; MANFRIN, 2013; MORAES et al., 2009; PFEILER; MARKOW, 2011). O gene mitocondrial *COI* é considerado um dos genes relevantes para identificação de animais pelo chamado código de barras (HEBERT; RATNASINGHAM; DEWAARD, 2003) e tem sido intensamente usado para este propósito em insetos (JANZEN et al., 2009; LINARES et al., 2009; SMITH et al., 2012).

#### **1.4.2 Gene nuclear *period*.**

O gene *period* está ligado ao cromossomo X, é um dos quatro loci em *Drosophila melanogaster* que influenciam seu ritmo biológico. Nesta espécie o gene *period* é composto por

cinco éxons (Figura 1) e codifica a proteína PER, que possui cerca de 1200 aminoácidos (CITRI et al., 1987). *Period* e *Timeless* são os genes atualmente aceitos para o modelo molecular do relógio biológico central de *D. melanogaster* (KOSTÁL; ZÁVODSKÁ; DENLINGER, 2009).



**Figure 2.** Estrutura do gene *period* em *Drosophila melanogaster*. Os “introns” são representados por linhas e os éxons por retângulos. O domínio PAS (região C2) está colorida com a cor cinza para os éxons 3 e 4. No éxon cinco estão indicadas às regiões não conservadas n2, n3 e n4, interespaçadas pelas regiões conservadas C3, C4 e C5 (COLOT; HALL; ROSBASH, 1988). A faixa listrada indica a região n2, formada pela região repetitiva Treonina-Glicina. O retângulo negro representa o domínio inibitório CCID (região C3 até região C4) (CHANG; REPPERT, 2003). Abaixo da estrutura do gene, uma barra negra indica a região do éxon cinco que foi analisada no presente trabalho. Tomada de (FRANCO et al., 2010).

O gene ortólogo do *period* de *D. melanogaster* tem sido clonado em diferentes organismos, incluindo a mosca doméstica *Musca domestica* (PICCIN et al., 2000), a mosca varejeira *Lucilia cuprina* (WARMAN et al., 2000), a mosca *Ceratitis capitata* (MAZZOTTA et al., 2005), em *Drosophila pseudoobscura* e *Drosophila virilis* (COLOT; HALL; ROSBASH, 1988) e em mamíferos (TEI et al., 1997; SHEARMAN et al., 1997; SUN et al., 1997).

O gene *period* controla o ritmo circadiano de acasalamento em *D. melanogaster* e *D. simulans* (SAKAI; ISHIDA, 2001; TAUBER et al., 2003) e está relacionado a outras características ligadas ao sucesso reprodutivo, tais como: a fecundidade em *D. melanogaster* (BEAVER et al., 2002, 2003), capacidade reprodutiva (diminuição do depósitos de ovos) no gafanhoto do deserto *Schistocerca gregaria* (TOBBACK et al., 2011) e tempo de duração da cópula em *D. melanogaster* (BEAVER; GIEBULTOWICZ, 2004). Além de *Drosophila*, foi demonstrado que os níveis de expressão do gene *period* podem causar isolamento reprodutivo pré-zigótico em moscas *Bactrocera cucurbitae* (Diptera: Tephritidae), por acarretarem alocria



e a alteração nos períodos do dia em que essas moscas estão receptivas à cópula (MIYATAKE et al., 2002).

O gene *period* possui um efeito pleiotrópico no ritmo do IPI (*interpulse interval*), um dos componentes do padrão acústico da corte sonora de machos de *Drosophila melanogaster* (KYRIACOU; HALL, 1980) e de outras espécies de *Drosophila* (LAGISZ et al., 2012). Yu et al. (1997) encontraram que a variação da região repetitiva de Thr-Gly do gene *period*, a qual direciona a variação no comprimento da sequência de proteína *per*, é fonte fundamental da variabilidade no comportamento. Como a corte sonora é um aspecto importante do comportamento sexual de *Drosophila* (GLEASON, 2005; YU et al., 1987) e de outros insetos (LINS et al., 2012), é possível que variabilidade na região repetitiva de Thr-Gly, contribua com aspectos críticos de isolamento reprodutivo entre espécies (KONOPKA; KYRIACOU; HALL, 1996; KYRIACOU; HALL, 1980; YU et al., 1987), por esta razão o gene *period* pode ser considerado um candidato que contribua diretamente para a especiação.

A parte central do gene *period* codifica uma longa extensão interrompida de treonina-glicina (Thr-Gly) em pares que tem implicação na evolução dos ritmos circadianos e ultradianos (SAWYER et al., 2006). Além disso, estas regiões variáveis dão origem à diversidade dentro de uma população, gerando diversificação estrutural e funcional do gene (COLOT; HALL; ROSBASH, 1988).

Estudos revelam que extensas regiões fora do domínio PAS do gene *period* (Figura 1) evoluem rapidamente e possuem regiões com taxas de substituição nucleotídicas diferentes, sendo formado por trechos altamente similares entre diferentes grupos de insetos, denominadas regiões conservadas (C1-C5) e para as regiões em que não se é possível o alinhamento entre algumas espécies de *Drosophila*, denominadas regiões não conservadas (n1-n5). (COLOT; HALL; ROSBASH, 1988). Em níveis taxonômicos mais profundos, mesmo a região PAS mostra grande heterogeneidade, por exemplo, entre *Drosophila* e *Lepidoptera* (REPPERT et al., 1994).

As diferentes regiões do gene *period* têm sido amplamente utilizadas para diferentes estudos. Na ordem Díptera (Famílias Siomyidae, Syrphidae, Psilidae, Drosophilidae, Tephritidae e Muscidae) foram comparadas as regiões repetitivas Thr-Gly, e encontrou-se que fora de Drosophilidae estas regiões são extremamente estáveis, mas dentro da família são submetidas à

expansão (NIELSEN et al., 1994). Em populações australianas, européias e africanas de *D. melanogaster* foram avaliadas regiões repetitivas Thr-Gly, encontrando-se uma repetição que apresenta uma clina latitudinal na Austrália, que é paralela à encontrada na Europa, reforçando a visão de que a seleção térmica pode ter contribuído para moldar os padrões continentais da variabilidade da repetição (SAWYER et al., 2006).

Na ordem Lepidoptera o gene *period* foi usado para resolver relações filogenéticas entre tribos, subtribos, família e superfamílias (REGIER et al., 1998; REGIER; COOK; et al., 2008; REGIER; GRANT; et al., 2008). Em esfingídeos (Lepidoptera: Sphingidae) tem sido útil para o entendimento da biogeografia da família Sphingidae (KAWAHARA et al., 2009). No Gênero *Anastrepha* (Tephritidae) foi usado o domínio CCID para a reconstrução filogenética e indicou que o mesmo pode ser considerado potencialmente informativo em estudos filogenéticos (BARR; CUI, 2005). No gênero *Ceratitis* (Tephritidae) foram estudadas as relações filogenéticas usando diferentes marcadores moleculares entre eles o gene *period* (BARR; WIEGMANN, 2009). No cluster *D. buzzatii* foi usado para resolver relações filogenéticas entre suas sete espécies (*D. buzzatii*, *D. koepferae*, *D. antonietae*, *D. serido*, *D. gouveai*, *D. seriema* e *D. borborema*), encontrando-se que a seleção purificadora é a principal força que governa a evolução do gene *period* (FRANCO et al., 2010). Também foi utilizado para avaliar a diferenciação genética entre as populações de *Lutzomyia intermedia* e *Lutzomyia whitmani* no leste do Brasil (MAZZONI et al., 2006) e em populações da espécie *Lutzomyia longipalpis* (BAUZER et al., 2002).

---

## II. OBJETIVOS

- Determinar a estrutura populacional de *D. meridionalis* por meio das análises de sequências parciais do gene nuclear *period* e o gene mitocondrial *COI*.
- Estabelecer uma hipótese filogeográfica para *D. meridionalis* a partir de sequências parciais do gene mitocondrial *COI* e o gene nuclear *period* e, inferindo os eventos históricos e demográficos envolvidos na determinação do padrão atual de diversificação das suas populações.
- Realizar uma comparação entre a filogeografia de *D. meridionalis* e das espécies do cluster *D. buzzatii* obtidas em estudos filogeográficos anteriores, em uma tentativa de inferir e/ou corroborar padrões biogeográficos existentes.

### III. MATERIAL E MÉTODOS

#### 1.0 Material Biológico.

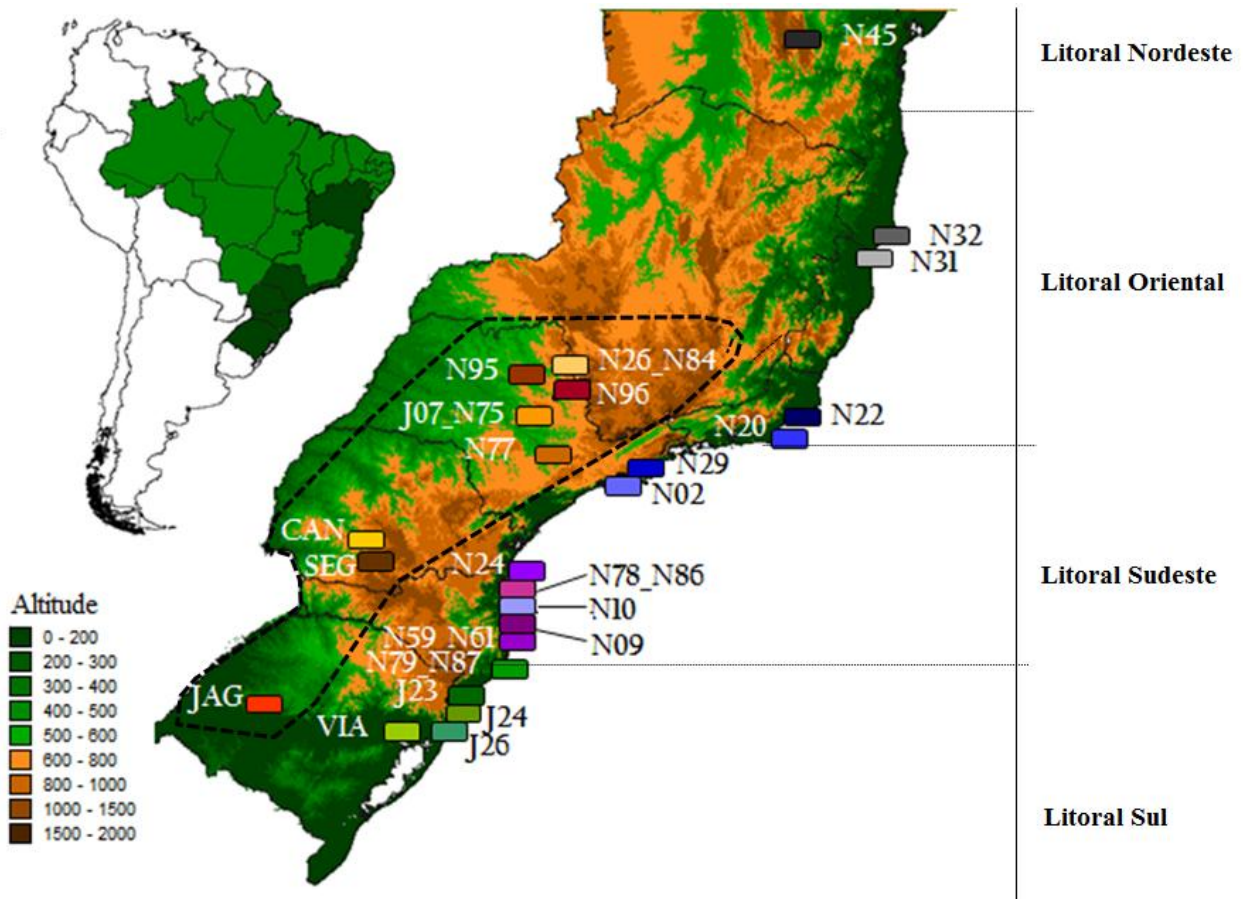
Para realização deste trabalho, foram analisadas sequências gênicas de indivíduos provenientes de populações de *D. meridionalis* (Tabela 1, Figura 3).

O DNA da maior parte das amostras foi extraído previamente e estava armazenado a -20°C no Laboratório de Genética Evolutiva da FMRP-USP. Uma coleta adicional de cladódios em decomposição foi realizada na localidade de Osório-RS. Os cladódios em decomposição foram retirados do campo, envoltos em jornal e transportados para o laboratório, onde foram acomodados em terrários. Destes cladódios, emergiram seis adultos de *D. meridionalis*, que foram identificados a partir de análises comparativas da genitália usando a referência de Vilela (1983).

A extração do DNA das amostras foi feita com auxílio do conjunto de reagentes *Wizard® Genomic DNA Purification Kit* (Promega), de acordo com as instruções do fabricante.

**Tabela 1.** Localidades das populações de *Drosophila meridionalis* amostradas. Para cada localidade estão relacionados: CG = Coordenadas Geográficas, CL = Código de Localidade, N1 = Amostras do gene mitocondrial COI, N2 = Amostras do gene nuclear *period*.

Localidade	CG	CL	N1	N2
Mucuri-BA	18°00'00"S, 39°30'00"W	N32	10	02
Mucugê-BA	13°00'07"S, 41°23'00"W	N45	01	01
Itaúnas-ES	18°24'26"S, 39°41'53"W	N31	07	02
Macaé-RJ	22°24'27"S, 41°48'14"W	N22	02	--
Arraial do Cabo-RJ	22°59'15"S, 42°00'49"W	N20	01	04
Altinópolis-SP	21°02'03"S, 47°20'20"W	N26_N84	10	05
Itirapina-SP	22°16'33"S, 47°48'27"W	J07_N75	11	05
Itu-SP	23°16'48"S, 47°13'52"W	N77	10	06
Sertãozinho-SP	21°10'39"S, 48°05'31"W	N95	11	05
Porto Ferreira-SP	21°51'29"S, 47°28'26"W	N96	06	05
São Sebastião-SP	23°16'48"S, 47°13'52"W	N29	07	06
Bertioga-SP	23°45'58"S, 45°54'42"W	N02	11	--
Foz de Jordão-PR	25°46'26"S, 52°06'59"W	SEG	03	03
Cantagalo-PR	25°25'15"S, 52°04'10"W	CAN	02	02
São Francisco do Sul-SC	26°16'03"S, 48°31'28"W	N24	10	05
Penha-SC	26°46'31"S, 48°35'53"W	N78_N86	10	04
Camboriú-SC	27°02'44"S, 48°35'15"W	N10	02	02
Governador Celso Ramos-SC	27°22'05"S, 48°31'60"W	N09	10	05
Florianópolis-SC	27°35'52"S, 48°25'46"W	N59_N61	13	01
Laguna-SC	28°29'18"S, 48°46'44"W	N79_N87	20	05
Arroio Teixeira-RS	27°22'05"S, 48°31'60"W	J23	10	04
Cornélio-RS	29°38'41"S, 50°03'24"W	J24	10	06
Osório-RS	29°56'15"S, 50°18'45"W	J26	07	05
Viamão-RS	30°03'45"S, 51°03'45"W	VIA	04	03
Jaguari-RS	29°26'55"S, 54°43'50"W	JAG	04	--
<b>Total</b>			<b>192</b>	<b>86</b>



**Figura 3.** Mapa topográfico gerado com auxílio do programa DIVA-GIS v. 7.5.0, disponível em [www.diva-gis.org](http://www.diva-gis.org), com as localizações geográficas das populações de *Drosophila meridionalis* analisadas neste trabalho. A localidade de cada código está descrita na Tabela 1. As divisões das populações nos grupos do interior e litoral que constam no mapa foram consideradas durante a apresentação dos resultados e discussão. \* VIA e N45 são populações incluídas no grupo do litoral.

### 1.1 Obtenção de sequências.

Neste trabalho, foram analisadas sequências de duas regiões gênicas: o gene nuclear *period* localizado no cromossoma X e gene mitocondrial *COI*. O tamanho amostral para cada região gênica analisada e em relação à procedência geográfica está mostrada na Tabela 1. Silva-Bernandi (2010) amplificou e editou as sequências do gene mitocondrial *COI* para 24 populações de *D. meridionalis*.

## 1.2 Isolamento, amplificação e sequenciamento do gene mitocondrial *COI*.

O isolamento e a amplificação do gene mitocondrial *COI* foram realizadas por meio da reação em cadeia da polimerase (PCR). A PCR continha entre 30 a 50 ng de DNA, 200 uM de dNTP, 0,4uM de cada um dos *primers* (oligonucleotídeos iniciadores), 1X de tampão flexi (200 mM Tris-HCl pH 8,4 ; 500 mM KCl), 0,3 ul (1U) de *Taq polimerase* (Promega®) e água mili-Q para completar 25 ul de volume final. A PCR foi feita em um Termociclador AB Applied Biosystems (Veriti 96 well Thermal Cycler) e as condições de amplificação foram: 94°C por 1'30" e 35 ciclos de 40" a 94°C, 40" a 48°C, e 2' a 72°C. Os oligonucleotídeos iniciadores utilizados para amplificar o gene mitocondrial *COI*, foram TY-J-1460 5'-CAA TTT ATC GCC TAA ACT TCA GCC-3' (*forward*) e C1-N-2191 5'-CCC GGT AAA ATT AAA ATA TAA ACT TC-3' (*reverse*) (SIMON et al., 1994). Os produtos amplificados foram verificados por eletroforese, usando um gel de agarose 1% corado com SYBR® (safe DNA gel stain). Os produtos foram visualizados por meio de exposição à luz UV em um transiluminador e seus tamanhos foram estimados a partir do padrão de bandas do marcador de peso molecular *Low DNA Mass Ladder* (©Invitrogen Corporation).

Para a purificação do produto de PCR foi usado o kit de purificação *Illustra™* (Ge Healthcare) seguindo as instruções do fabricante, seguidamente cada amostra foi dosada no nanodrop™ e a concentração ajustada para o sequenciamento. A reação de sequenciamento foi realizada com auxílio do conjunto de reagentes *ABI PRISM BigDye™ Terminator Cycle Sequencing Ready Reaction Kit*. A reação de sequenciamento constitui-se de 1µl de *Big Dye*, 1.5µl de tampão de *Big Dye 5X*, 10umol de oligonucleotídeos iniciadores, 50 ng/ul de DNA e água mili-Q para completar 10 ul de volume final. Para o sequenciamento foram usados os mesmos oligonucleotídeos iniciadores e as condições de amplificação foram: 94°C por 1' e 39 ciclos de 15" a 94°C, 15" a 53°C, e 4' a 60°C.

Finalmente, para precipitar o produto de sequenciamento foi adicionado 80µl de isopropanol 65% à placa de PCR (PCR® microplate) contendo os produtos das reações de sequenciamento e seguidamente mantidos em repouso por 15 minutos à temperatura ambiente. Em seguida, a placa foi centrifugada a 4.000 rpm por 45 minutos a 4°C e o sobrenadante descartado por inversão da placa. Posteriormente, foi adicionado 200ul de etanol 70% gelado a

cada produto e centrifugada a placa a 14.000 rpm por 10 minutos a 4°C e o sobrenadante foi descartado por inversão novamente da placa. O procedimento anterior da adição do etanol gelado foi repetido novamente. Por último a placa é centrifugada com um *spin* invertido a 1000 rpm por 1 minuto e secada a 37°C na placa aquecedora por 1 hora para a evaporação do etanol. As amostras foram sequenciadas em sequenciador automático de DNA modelo *ABI Prism 377*.

### 1.3 Isolamento, amplificação e sequenciamento do gene nuclear *period*.

O isolamento e a amplificação da região CCID do gene *period* foram realizadas por meio da reação em cadeia da polimerase (PCR). A PCR continha entre 30 a 50 ng de DNA, 200 µM de dNTP, 5µM de cada um dos *primers* (oligonucleotídeos iniciadores), 1X de tampão flexi (200 mM Tris-HCl pH 8,4 ; 500 mM KCl), 0,3 µl (1U) de *Taq polimerase* (Promega®) e água mili-Q para completar 25 µl de volume final. A PCR foi feita em um Termociclador AB Applied Biosystems (Veriti 96 well Thermal Cycler) e as condições de amplificação foram: 94°C por 1'30" e 35 ciclos de 30" a 94°C, 30" a 53°C, e 1' a 72°C. Os oligonucleotídeos iniciadores utilizados para amplificar o gene *period* foram per-CBf (5' TGGGAG GGC GAG GCGAAC AA '3) e per-CBr (5'GGC ATG GGT TGG TAC ATC AT'3) (FRANCO et al., 2010). Os produtos amplificados foram verificados por eletroforese, usando um gel de agarose 1% corado com SYBR® (safe DNA gel stain). Os produtos foram visualizados por meio de exposição à luz UV em um transiluminador e seus tamanhos foram estimados a partir do padrão de bandas do marcador de peso molecular *Low DNA Mass Ladder* (©Invitrogen Corporation).

A purificação do produto de PCR foi usado o kit de purificação *Illustra™* (Ge Healthcare) seguindo as instruções do fabricante, seguidamente cada amostra foi dosada no nanodrop™ e a concentração ajustada para o sequenciamento. A reação de sequenciamento foi realizada com auxílio do conjunto de reagentes *ABI PRISM BigDye™ Terminator Cycle Sequencing Ready Reaction Kit*. A reação de sequenciamento constitui-se de 1µl de *Big Dye*, 1.5µl de tampão de *Big Dye 5X*, 10µmol de oligonucleotídeos iniciadores, 50 ng/ul de DNA e água mili-Q para completar 10 µl de volume final. Para o sequenciamento foram usados os mesmos oligonucleotídeos iniciadores e as condições de amplificação foram: 94°C por 1' e 39 ciclos de 15" a 94°C, 15" a 53°C, e 4' a 60°C.



Para a precipitação do produto de sequenciamento e o sequenciamento foi utilizado o mesmo procedimento aplicado para o gene mitocondrial *COI*.

## 2.0 Análises das sequências.

Sequências *forward* e *reverse* foram analisadas, comparadas, corrigidas e editadas usando o programa Chromas Lite 2.0. O alinhamento múltiplo das sequências foi gerado no programa ClustalW Multiple Alignment 1.8 (THOMPSON; HIGGINS; GIBSON, 1994) implementado no programa BioEdit versão 7.0.9.0 (HALL, 1999) e inspecionado visualmente para avaliação de possíveis erros do sequenciamento.

Os índices de estatística descritiva estimados foram: ( $\pi$ ), que corresponde ao número médio de diferenças nucleotídicas por sítio entre duas sequências (NEI, 1987), o número de haplótipos ( $h$ ), o índice de diversidade haplotípica ( $H_d$ ) (NEI, 1987), o número médio de diferenças nucleotídicas ( $k$ ) (TAJIMA, 1983). Todos esses parâmetros foram estimados por meio do programa DnaSP 5.10 (ROZAS et al., 2003). Com o auxílio do programa MEGA 5 (TAMURA et al., 2011), foi selecionado o modelo de substituição nucleotídica que melhor explica a evolução das sequências analisadas.

## 2.1 Inferências filogeográficas e estrutura populacional.

Foi utilizada a Análise Filogeográfica de Agrupamento de Clados Aninhados ou NCPA (*Nested Clade Phylogeographic Analysis*), proposta por Templeton, Routman, Phillips (1995) e Templeton (1998) como um dos métodos estáticos para estimar eventos filogeográficos em *D. meridionalis*. Essa metodologia tem basicamente quatro etapas:

1. *Construção de uma rede de haplótipos*: a rede de haplótipos é gerada considerando o algoritmo de parcimônia estatística (TEMPLETON; CRANDALL; SING, 1992) com auxílio do programa TCS 1.21 (CLEMENT; POSADA; CRANDALL, 2000). O algoritmo de parcimônia estatística usa o modelo de sítios finitos para sequências de DNA e começa por estimar o número máximo das diferenças entre os haplótipos, como resultado de substituições simples com um nível de significância de 95% (ou limite de conexão de parcimônia). Os haplótipos são conectados uns aos outros por um passo mutacional, sendo

que os haplótipos não amostrados são representados por pequenas circunferências (POSADA; CRANDALL, 2001). As reticulações, quando presentes, podem ser eliminadas utilizando três critérios descritos em Crandall e Templeton (1993), sendo eles: a) o critério topológico: sugere que os haplótipos têm maior probabilidade de estarem conectados a haplótipos de interior, mais antigos, do que a haplótipos localizados nas extremidades da genealogia, mais recentes; b) o critério de frequência: sugere que os haplótipos têm maior probabilidade de estarem conectados a haplótipos mais frequentes; c) o critério geográfico: sugere que haplótipos devam ser mais relacionados aos haplótipos de sua população de origem do que com outra população (CRANDALL; TEMPLETON, 1993).

2. *Aninhamento de clados*: O algoritmo de aninhamento consiste em agrupar haplótipos (clados de nível 0) separados por uma única mutação para formar um clado de nível 1, desde os haplótipos das extremidades até o interior da rede. Posteriormente, os clados de nível 1 que estão separados por uma mutação, são agrupados em clados de nível 2 e assim sucessivamente, até o seguinte nível de aninhamento que constitua a rede total ou que o limite de conexão de parcimônia é atingido (TEMPLETON; BOERWINKLE; SING, 1987).

3. *Testes de significância para verificação de associação entre clados e distribuição geográfica*: Os dados geográficos são quantificados de duas formas principais. A primeira é a distância de clados ( $D_c$ ), que mede a amplitude da distribuição geográfica de um determinado clado e a segunda é a distância do clado aninhado ( $D_n$ ), que mede a distância média de um haplótipo ou clado em relação aos haplótipos ou clados aos quais ele está evolutivamente mais relacionado, ou seja, os clados com os quais está aninhado em um clado superior. A hipótese nula de não associação da rede de haplótipos com a distribuição geográfica da espécie é testada por um mínimo de 1000 permutações aleatórias, as quais são necessárias para fazer inferência estatística no nível de 5% (TEMPLETON; ROUTMAN; PHILLIPS, 1995; TEMPLETON, 2006). Os cálculos de  $D_c$  e  $D_n$  e os testes de significância foram feitos pelo programa GeoDis 2.6 (POSADA; CRANDALL; TEMPLETON, 2000), que inclui a correção de Dunn-Sidak para múltiplos testes.

4. *Interpretação dos resultados:* A significância estatística não é a mesma que a significância biológica (TEMPLETON, 2006). Por tanto é preciso interpretar biologicamente a significância estatística com uma chave de inferências (TEMPLETON; ROUTMAN; PHILLIPS, 1995; TEMPLETON, 2004b). Com uma nova versão da chave de inferência de Templeton, 2011 disponível em <http://darwin.uvigo.es/software/geodis.html>, foram interpretados os Dc e Dn gerados e estatisticamente significativos e foi possível a detecção dos prováveis eventos históricos (isolamento por distância, fragmentação alopátrica e expansão de distribuição) e/ou recorrentes responsáveis pela associação significativa entre a variação genética e a distribuição geográfica da amostra analisada de *D. meridionalis*.

Para detectar estrutura populacional, foi realizada uma Análise de Variância Molecular (AMOVA)(EXCOFFIER; SMOUSE; QUATTRO, 1992). A AMOVA estima índices da estrutura genética usando informações sobre as diferenças entre haplótipos, as quais são inseridas em uma matriz de distâncias euclidianas ao quadrado. Esta análise determina as estimativas de componentes de variância e os análogos da estatística F, designados na AMOVA como estatística Phi ( $\Phi$ ), que reflete a correlação da diversidade haplotípica em diferentes níveis hierárquicos (dentro dos indivíduos, dentro de populações, dentro de grupos de populações, entre os grupos). A significância dos componentes de variância e das estatísticas  $\Phi$  são testados usando procedimentos não paramétricos de permutação, sendo que as permutações são diferentes para cada um dos componentes de covariância, eliminado o pressuposto de normalidade convencional para análise de variância, mas imprópria para dados moleculares. A variância molecular total é a soma de todos os componentes de variância devido às diferenças entre os haplótipos dentro das populações ( $\sigma^2_c$ ,  $\Phi_{ST}$ ), às diferenças entre haplótipos em diferentes populações dentro dos grupos ( $\sigma^2_b$ ,  $\Phi_{SC}$ ) e às diferenças entre grupos ( $\sigma^2_a$ ,  $\Phi_{CT}$ ) (EXCOFFIER; SMOUSE; QUATTRO, 1992). Foram realizadas AMOVAs considerando diferentes agrupamentos populacionais e foram realizadas pelo programa Arlequin 3.5 © (EXCOFFIER; LISCHER, 2010).

O Teste de Mantel (MANTEL, 1967) tem sido amplamente usado na genética de populações para comparar duas ou mais matrizes de distância, testando hipóteses sobre estruturas espaciais, temporais ou ambientais em distâncias genéticas (TELLES; DINIZ-FILHO, 2005). Esse teste avalia a significância estatística de duas matrizes correlacionadas, por meio de um

procedimento permutacional (EPPERSON, 2003). Para testar a correlação entre distância genética e a geográfica, foram calculadas as matrizes de distância genética ( $F_{st}$ ) entre as populações no programa Arlequin 3.5 © (EXCOFFIER; LISCHER, 2010) e de distâncias geográficas dessas populações com o auxílio do programa *Geographic Distance Matrix Generator* -Version 1.2.3. O teste de Mantel foi realizado com auxílio do programa Arlequin 3.5 © (EXCOFFIER; LISCHER, 2010) e a significância da mesma foi calculada com 10000 permutações.

## 2.2 Testes de Neutralidade e a *Mismatch Distribution*.

Os testes de neutralidade realizados foram:

-D de Tajima (Tajima, 1989): é baseado em um modelo de coalescência (alelos neutros em uma população de tamanho constante e modelo de sítios infinitos sem recombinação) e comumente aplicado a dados de polimorfismo de DNA de uma única espécie. A quantidade de variação acumulada em uma sequência pode ser medida pelo parâmetro theta ( $\theta$ ), cujo valor é dado pela taxa de mutação de uma sequência e pelo tamanho efetivo populacional ( $\theta=4N_e\mu$  e  $\theta=3N_e\mu$  para genes ligados ao X). Esse parâmetro pode ser estimado através de  $\Pi$ , que corresponde ao número de diferenças entre duas sequências, escolhidas aleatoriamente na amostra, ou por meio de  $S$ , que é o número de sítios polimórficos das sequências amostradas (TAJIMA, 1989). Tajima (1989) propôs que a diferença ( $D$ ) entre as estimativas de  $\theta$ , baseadas em  $\Pi$  e  $S$ , serviriam como teste para detecção de desvios da neutralidade. Em condições de neutralidade (todas as mutações são seletivamente neutras e a população mantém um tamanho constante ao longo do tempo) as estimativas de  $\theta$ , baseadas em  $\Pi$  e  $S$ , seriam aproximadamente as mesmas ( $D=0$ ). Com eventos seletivos assim como mudanças no tamanho efetivo da população através do tempo aconteceram, os valores de  $D$  serão diferentes de zero. Valores negativos e significativos ( $D<0$ ) indicam excesso de haplótipos com baixas frequências, o que sugere expansão populacional, gargalos populacionais, eventos fundadores, efeito carona, varredura seletiva (*selective sweep*) (TEMPLETON, 2006) e seleção balanceadora (HAMILTON, 2009), enquanto valores positivos ( $D>0$ ) indicam um excesso de haplótipos com frequências intermediárias, o que sugere seleção direcional forte ou contração populacional (HAMILTON, 2009).

-Fs de Fu (Fu, 1997): Baseia-se em um modelo de sítios infinitos sem recombinação. Esse teste avalia a probabilidade de ter  $k$  alelos em uma amostra de  $n$  sequências dado o valor de  $\theta$ , sendo o valor de Fs dado pela equação abaixo:

$$F_s = \ln(S'/1 - S') \text{ ((Fu, 1997), equação 1)}$$

Em que  $S'$  ( $S' = \Pr(K > k | \theta = \theta\pi)$ ), é a probabilidade de ocorrência de um número de alelos menor ou igual em relação a uma amostra aleatória, dado que  $\theta$  ( $\theta\pi =$  número médio de diferenças de nucleotídeos entre duas sequências). Em uma amostra com excesso de mutações recentes, o parâmetro  $\theta$ , baseado no valor  $\pi$ , tende a ser menor do que aquele baseado no número de alelos. Sendo assim,  $S'$  pode ser um bom indicador sobre quantidade de mutações recentes na amostra, influenciando o valor de Fs (FU, 1997). A significância do valor Fs é testado por meio da geração de amostras aleatórias, sob a hipótese de neutralidade seletiva e equilíbrio da população, usando um algoritmo de simulação coalescente. Uma particularidade do teste Fs é que ele deve ser considerado significativo apenas quando os valores de  $p$  forem menores ou iguais a 0,02 e não 0,05, como é o caso dos outros testes de neutralidade. Assim, como o teste D de Tajima, valores negativos e significativos do teste Fs são um indicativo de expansão populacional (FU, 1997).

Os testes D de Tajima e Fs de Fu foram realizados com o auxílio do programa Arlequin 3.5 © (EXCOFFIER; LISCHER, 2010).

Teste R2 (Ramos-Onsins; Ramos, 2002): O teste R2 é baseado na diferença entre o número de mutações únicas (U) e o número médio de diferenças nucleotídicas entre duas sequências (k). Neste teste é que o número esperado de haplótipos únicos (*singletons*) em uma genealogia, após de um evento de crescimento populacional severo e recente é  $k/2$ , conseqüentemente, valores pequenos de R2 são esperados sob um cenário de expansão populacional (RAMOS-ONSINS; ROZAS, 2002). Esse teste foi realizado com o auxílio do programa DNAsp 5.10 (ROZAS et al., 2003)

Além dos testes de neutralidade, também foi realizada a *Mismatch distribution*, que é uma avaliação da distribuição do número médio de diferenças par a par entre as sequências de DNA em uma amostra e que pode gerar padrões distintos em função da história demográfica da população. Espera-se que a distribuição seja bi ou multimodal em amostras extraídas de

populações em equilíbrio demográfico ou subdivisão populacional, enquanto a distribuição unimodal mostrará que as populações têm passado por uma recente expansão ou contração demográfica (ROGERS; HARPENDING, 1992) ou por uma faixa de expansão, com elevados níveis de migração entre os demes vizinhos (BERTORELLE; EXCOFFIER, 1998). Com o auxílio do programa Arlequin 3.5© (EXCOFFIER; LISCHER, 2010) foram estimados três parâmetros relacionados à expansão demográfica: o tempo desde a expansão em unidades de tempo mutacional ( $\tau$ ) e as estimativas do tamanho populacional antes e depois da expansão ( $\theta_0 = 2\mu N_0$ , e  $\theta_1 = 2\mu N_1$ , sendo  $N_0$  e  $N_1$  os tamanhos populacionais antes e depois da expansão, respectivamente). As simulações por coalescência foram realizadas usando os parâmetros estimados para a expansão demográfica. Esses parâmetros foram re-estimados para cada simulação, a fim de obter os intervalos de confiança empíricos. A SSD (*Sum of Squared Differences*) entre a *Mismatch distribution* simulada e observada foi usada como teste estatístico. O valor de p é avaliado por:

$$P = \frac{SSD_{sim} \geq SSD_{obs}}{B}$$

Onde p é o p-valor e B é o *bootstrapped* para os parâmetros estimados.

Os testes de neutralidade e a *Mismatch Distribution* tem sido classificados em três categorias com base nas informações que esses utilizam. Os testes de classe I são baseados nas frequências de mutações (D de Tajima e R2), a classe II (Fs de Fu) na distribuição dos haplótipos e a classe III sobre a distribuição do número médio de diferenças par a par entre as sequências (*Mismatch Distribution*). Dentro dos testes mais sensíveis para detecção de eventos de expansão populacional são aqueles que pertencem à classe II, dentro desses, o Fs de Fu. Embora para pequenas amostras, seja recomendado o R2. O Fs de Fu é fortemente afetado pela recombinação, de modo que se há dúvida sobre a ocorrência destes eventos nas amostras, recomenda-se o uso dos testes de D de Tajima e R2 (RAMÍREZ-SORIANO et al., 2008).

### 2.3 Inferências Bayesianas.

Para acrescentar informações sobre as relações evolutivas entre as sequências, história demográfica e estimar o tempo do ancestral comum mais recente ( $T_{MRCA}$ ) para as populações de *D. meridionalis* foram realizadas inferências filogeográficas com uma abordagem bayesiana, usando o modelo de coalescência (KINGMAN, 1982).

A *Bayesian Skyline Plot* (BSP) é um método que estima a dinâmica populacional ancestral através do tempo, a partir de uma amostra de sequências, dado um determinado modelo de substituição nucleotídica e uma taxa de mutação (DRUMMOND et al., 2005). A BSP utiliza o método de Monte Carlo via Cadeias de Markov (MCMC- *Markov Chain Monte Carlo*), um procedimento de amostragem que combina *generalized skyline plots* para gerar uma distribuição posterior do tamanho efetivo da população ao longo do tempo. A BSP inclui intervalos de confiança para o tamanho efetivo populacional, estimado em cada ponto no tempo de volta para o ancestral comum mais recente das sequências de genes (DRUMMOND et al., 2005). A BSP foi realizada com o auxílio do programa BEAST 1.6.2 (DRUMMOND; RAMBAUT, 2007).

Para a realização da BSP do gene mitocondrial *COI* foi usada uma taxa de substituição de 1,4% por milhão de anos (NAGARAJA; RANGANATH, 2004; RUSSO; TAKEZAKI; NEI, 1995) e o modelo evolutivo de substituição nucleotídica TN93+G+I. O número de interações na MCMC foi de 200 milhões, amostradas a cada 1000 interações com um período de *burn-in* de 10% do total de interações. As interações obtidas com o programa BEAST foram analisadas no programa TRACER 1.5 (RAMBAUT; DRUMMOND, 2003), o qual gera o gráfico da BSP.

Árvores plausíveis foram geradas para as sequências do gene mitocondrial *COI*, considerando como parâmetros, o tamanho constante e relógio molecular relaxado (*uncorrelated lognorma*), com auxílio do programa BEAST 1.6.2 (DRUMMOND; RAMBAUT, 2007). Com o programa TreeAnnotator 1.6.2 resumisse as árvores usando o padrão de agrupamento *Maximum clade credibility tree-MCC tree*, que encontra a árvore com a maior probabilidade a posteriori de todos os seus nós. A visualização da árvore de coalescência foi feita pelo programa Fig Tree v.1.3.1. A convergência e estabilidade desta análise foram checadas usando o programa Tracer 1.5 (DRUMMOND; RAMBAUT, 2007) que apresentou valores do tamanho efetivo da amostra

(ESS) superiores a duzentos. Para esta análise foi usado o mesmo modelo evolutivo de substituição e a mesma taxa de substituição nucleotídica, usada para a BSP.

### 3.0 Modelagem de distribuição de espécies.

Com o objetivo de compreender a dinâmica da distribuição da espécie, foram geradas e comparadas projeções com a ocorrência potencial da espécie em dois cenários: sua distribuição atual e durante o último período glacial, há cerca de 21.000 anos. Para a geração dessas projeções foram usadas 25 coordenadas geográficas. Para a modelagem da distribuição atual de *D. meridionalis* foi utilizada a base de dados do *WorldClim v.1.4* (HIJMANS et al. 2005). As 19 variáveis bioclimáticas são derivadas da temperatura mensal e valores de precipitação. Estas variáveis bioclimáticas representam tendências anuais (por exemplo, a temperatura média anual, precipitação anual), sazonalidade (por exemplo, variação anual de temperatura e precipitação) e extremos ou fatores ambientais limitantes (por exemplo, a temperatura do mês mais frio e mais quente, e precipitação dos trimestres mais úmidos e secos). A distribuição de *D. meridionalis*, também foi modelada para o cenário climático do último máximo glacial (*Last glacial maximum LGM*) aproximadamente 21.000 anos antes do presente. Os modelos usados foram fornecidos pela “Comunidade de Sistemas de Modelos Climáticos” (*Community Climate System Model – CCSM*) e o modelo gerado pelo grupo “Modelos para Pesquisa Interdisciplinar do Clima” (*Model for Interdisciplinary Research On Climate – MIROC*), ambos disponíveis em: <http://www.worldclim.org/past>, com resolução de 2,5 arco-minutos, em que cada pixel corresponde a um quadrado de cinco quilômetros de lado. As análises foram realizadas por meio do algoritmo de Entropia Máxima implementado no programa *Maxent v.3.3.3a* (PHILLIPS; ANDERSON; SCHAPIRE, 2006) e os mapas foram editados no programa *DIVA-GIS v. 7.5.0* (Disponível em <[www.diva-gis.org](http://www.diva-gis.org)>. Acesso em 09 de março de 2012).



## IV. RESULTADOS

### 1.0 Análises das sequências do gene mitocondrial *COI*.

Foram analisadas 642 pares de bases (pb) de 192 indivíduos machos distribuídos em 25 populações de *D. meridionalis* (Figura 3), gerando 55 haplótipos. Estas sequências foram geradas parcialmente por Silva-Bernardi (2010) e neste trabalho, e compreendem a posição 1510 do genoma mitocondrial de *Drosophila yakuba* (Número de acesso NCBI GenBank NC\_001322), e o primeiro códon a partir dessa base equivale ao décimo terceiro aminoácido do gene em questão.

O modelo de substituição nucleotídica calculado que melhor explica a evolução das sequências é T92+G+I (AICc 4024,674). No entanto, quando esse modelo não foi disponibilizado pelos programas usados, foi substituído pelo modelo TN93+G+I.

Na Tabela 2 estão apresentados os índices calculados de diversidade genética para a amostra obtida. A diversidade nucleotídica total ( $\pi$ ) foi de  $0,03353 \pm 0,0009$ , a população Laguna-SC (N79\_N87) apresentou o maior índice de diversidade ( $0,02851 \pm 0,0054$ ) e as populações Macaé-RJ (N22), Altinópolis-SP (N26\_N84), Camboriú-SC (N10) apresentaram os menores índices de diversidade nucleotídica ( $0,00000 \pm 0,0000$ ).

**Tabela 2.** Índices de diversidade genética e haplótipos de populações de *Drosophila meridionalis*, utilizando sequências do gene mitocondrial *COI*.

	<b>N</b>	<b>h</b>	<b>S</b>	<b>Hd</b>	$\pi_1$	<b>k</b>	<b>Haplótipos*</b>
<b>N32</b>	10	3	2	0,378	0,00062±0,0003	0,4	Hap 23 (8) ; Hap 33 (1) ; Hap 34 (1)
<b>N45</b>	1	1	0	1,000	---	---	Hap 23 (1)
<b>N31</b>	7	2	1	0,571	0,00089±0,0002	0,5	Hap 23 (3) ; Hap 32 (4)
<b>N22</b>	2	1	0	0,000	0,00000±0,0000	0,0	Hap 28 (2)
<b>N20</b>	1	1	0	1,000	---	---	Hap 27 (1)
<b>N26/N84</b>	10	1	0	0,00	0,000±0,0000	0,0	Hap 41 (10);
<b>J07/N75</b>	11	6	6	0,800	0,00363±0,0005	2,3	Hap 37 (1) ; Hap 38 (5) ; Hap 39 (2) ; Hap 40 (1) ; Hap 41 (1) ; Hap 42 (1)
<b>N77</b>	10	4	11	0,778	0,00768±0,0017	4,9	Hap 38 (3) ; Hap 43 (4) ; Hap 44 (1) ; Hap 45 (2)
<b>N95</b>	11	4	5	0,709	0,00312±0,0004	2,0	Hap 38 (1) ; Hap 39 (5) ; Hap 41 (4) ; Hap 41 (1) ; Hap 51 (1)
<b>N96</b>	6	3	4	0,600	0,00301±0,0010	1,9	Hap 38 (1) ; Hap 39 (4) ; Hap 41 (1)
<b>N29</b>	7	3	39	0,524	0,01736±0,0112	11,1	Hap 21 (5) ; Hap 22 (1) ; Hap 31 (1)
<b>N02</b>	11	4	39	0,709	0,01161±0,0082	7,4	Hap 19 (4) ; Hap 20 (1) ; Hap 2 (5) ; Hap 22 (1)
<b>SEG</b>	3	3	11	1,000	0,01142±0,0049	7,3	Hap 3 (1) ; Hap 52 (1) ; Hap 53 (1)
<b>CAN</b>	2	2	2	1,000	0,00312±0,0016	2,0	Hap 1 (1) ; Hap 2 (1)
<b>N24</b>	10	3	2	0,511	0,00087±0,0003	0,5	Hap 23 (7) ; Hap 29 (1) ; Hap 30 (2)
<b>N78-N86</b>	10	2	1	0,467	0,00073±0,0002	0,4	Hap 23 (7) ; Hap 46 (3)
<b>N10</b>	2	1	0	0,000	0,00000±0,0000	0,0	Hap 23 (2)
<b>N09</b>	10	4	52	0,644	0,01862±0,0097	11,9	Hap 23 (6) ; Hap 24 (1) ; Hap 25 (2) ; Hap 26 (1)
<b>N59-N61</b>	13	3	49	0,295	0,01356±0,0079	8,7	Hap 23 (11) ; Hap 35(1) ; Hap 36 (1)
<b>N79-N87</b>	20	5	42	0,763	0,02851±0,0054	18,3	Hap 8 (8) ; Hap 47 (5) ; Hap 48 (4) ; Hap 49 (2) ; Hap 50 (1)
<b>J23</b>	10	5	12	0,667	0,00498±0,0019	3,2	Hap 5 (1) ; Hap 6 (1) ; Hap 7 (1) ; Hap 8 (6) ; Hap 9 (1)
<b>J24</b>	10	7	13	0,911	0,00526±0,0014	3,3	Hap 6 (3) ; Hap 9 (1) ; Hap 10 (1) ; Hap 11 (2) ; Hap 12 (1) ; Hap 13 (1) ; Hap 14 (1)
<b>J26</b>	7	6	8	0,952	0,00386±0,0011	2,4	Hap 6 (2) ; Hap 13 (1) ; Hap 15 (1) ; Hap 16 (1) ; Hap 17 (1) ; Hap 18 (1)
<b>VIA-1</b>	4	4	6	1,000	0,00467±0,0013	3,0	Hap 6 (1) ; Hap 11 (1) ; Hap 54 (1) ; Hap 55 (1)
<b>JAG</b>	4	2	1	0,667	0,00104±0,0003	0,6	Hap 3 (2) ; Hap 4 (2)
<b>D.</b>	192	5	73	0,923	0,03353±0,0009	21,5	
<b>meridionalis</b>		5					

\*Haplótipos segundo a NCPA, dentro dos parêntesis esta o número de indivíduos por haplótipos.

N = número de indivíduos; h = número de haplótipos; S = número de sítios polimórficos; Hd = diversidade haplotípica;  $\pi_1$  = diversidade nucleotídica (desvio-padrão); k = número médio de diferenças nucleotídicas. Os códigos das populações são os mesmos apresentados na Tabela 1.

### 1.1 Nested Clade Phylogeographic Analysis (NCPA).

Três redes de haplótipos, denominadas A, B e C, foram geradas para as populações de *D. meridionalis* utilizando o programa TCS (CLEMENT; POSADA; CRANDALL, 2000). Cada uma das redes foi usada isoladamente para investigar a associação entre a distribuição geográfica e a genealogia obtida.

#### -Rede A

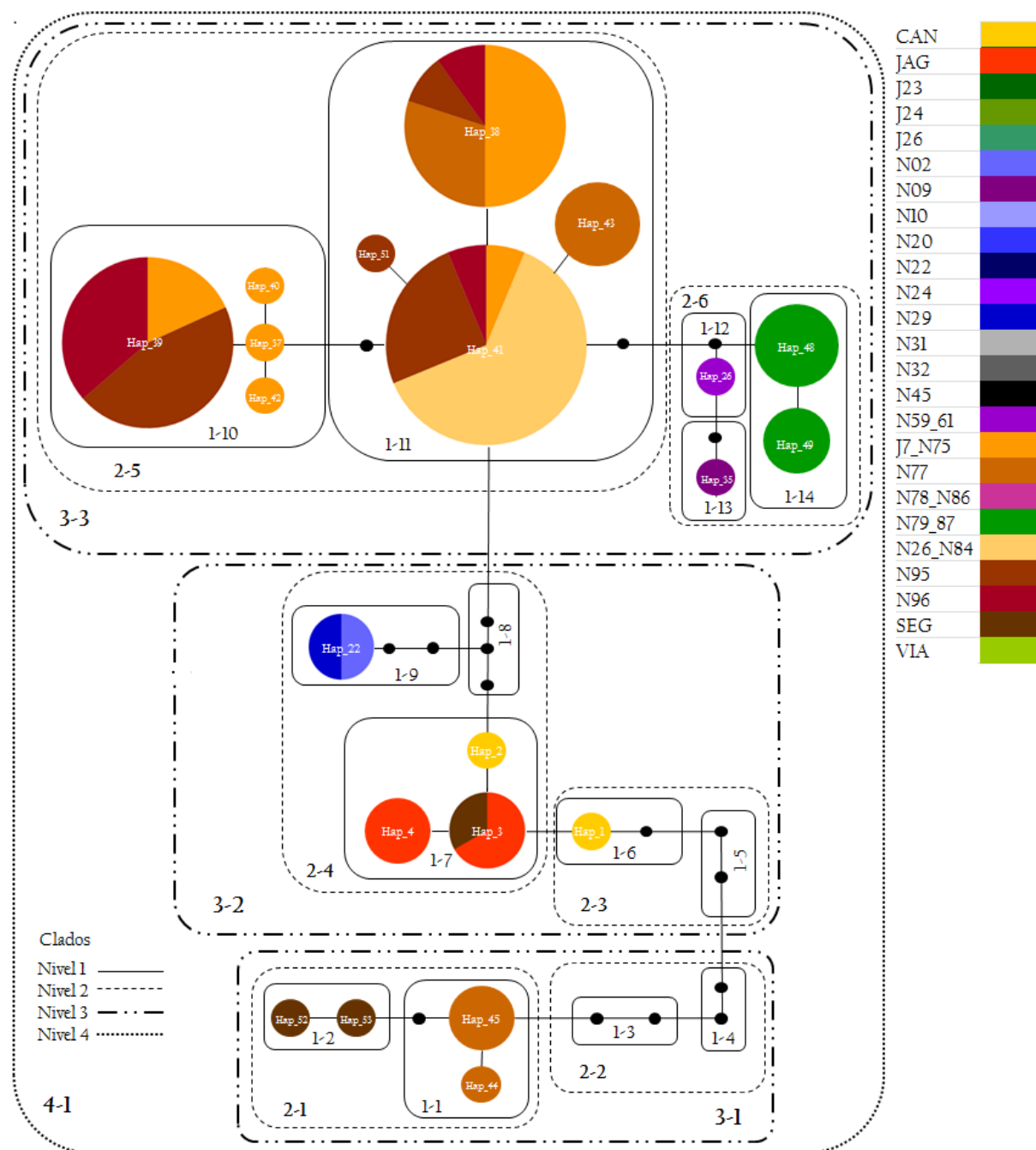
A rede A (Figura 4) é composta por 21 haplótipos provenientes de populações do interior: Altinópolis-SP (N26\_N84), Sertãozinho-SP (N95), Porto Ferreira-SP (N96), Itirapina-SP (J07\_N75), Itu-SP (N77), Foz do Jordão-PR (SEG), Cantagalo-PR (CAN), Jaguari-RS (JAG) e de populações do litoral: São Sebastião-SP (N29), Bertioga-SP (N02), Governador Celso Ramos-SC (N09), Florianópolis-SC (N61) Laguna-SC (N79\_N87).

O Hap\_41 apresentou o maior *outgroup weight* (0,324), indicando que este tem a maior probabilidade de ser o haplótipo mais ancestral da genealogia. Em concordância com tal resultado, esse haplótipo ocupa uma posição interior na rede e tem a maior frequência e ampla distribuição geográfica (CLEMENT; POSADA; CRANDALL, 2000; TEMPLETON, 2006), encontrando-se nas populações do interior de São Paulo.

Na Tabela 3 estão descritas as associações geográficas significativas, encontradas nos diferentes clados da rede A.

**Tabela 3.** Resultados significativos da NCPA para a rede A, obtida a partir de amostras do gene mitocondrial *COI* de populações de *Drosophila meridionalis*, de acordo com a chave de inferências filogeográficas de Templeton (2011), disponível em <http://darwin.uvigo.es/software/geodis.html>

CLADO	Etapas da Chave de Inferências
1-11	1-2-11-12 (Não): Expansão populacional contígua.
2-5	1-2-3-4 (Não): Fluxo gênico restrito com isolamento por distância
3-3	1-19-20-2-11-12 (Não): Expansão populacional contígua.
4-1	1-2-3-4 (Não): Fluxo gênico restrito com isolamento por distância



**Figura 4.** Rede A de haplótipos não enraizada contendo 21 haplótipos do gene mitocondrial *COI* de populações de *Drosophila meridionalis*. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. Haplótipos hipotéticos, não amostrados, são representados por pequenos círculos pretos. As cores dentro de cada haplótipo representam as diferentes populações, sendo essas as mesmas apresentadas na Figura 1.

### **-Rede B**

A rede B (Figura 5) é composta por 20 haplótipos provenientes de populações do Rio Grande do Sul: Arroio Teixeira-RS (J23), Cornélios-RS (J24), Osório-RS (J26), Viamão-RS (VIA) e populações de Santa Catarina: Laguna-SC (N79\_N87), Florianópolis-SC (N61) e Governador Celso Ramos-SC (N09). Essa rede apresenta vários haplótipos com frequências elevadas (Hap\_06 e Hap\_08), sendo o Hap\_08 com o maior *outgroup weight* (0,2781). Esta rede apresentou forma de estrela (*Star Shape*), com um ou poucos haplótipos com frequências elevadas e muitos haplótipos diferindo por uma ou poucas mutações (AVISE, 2000).

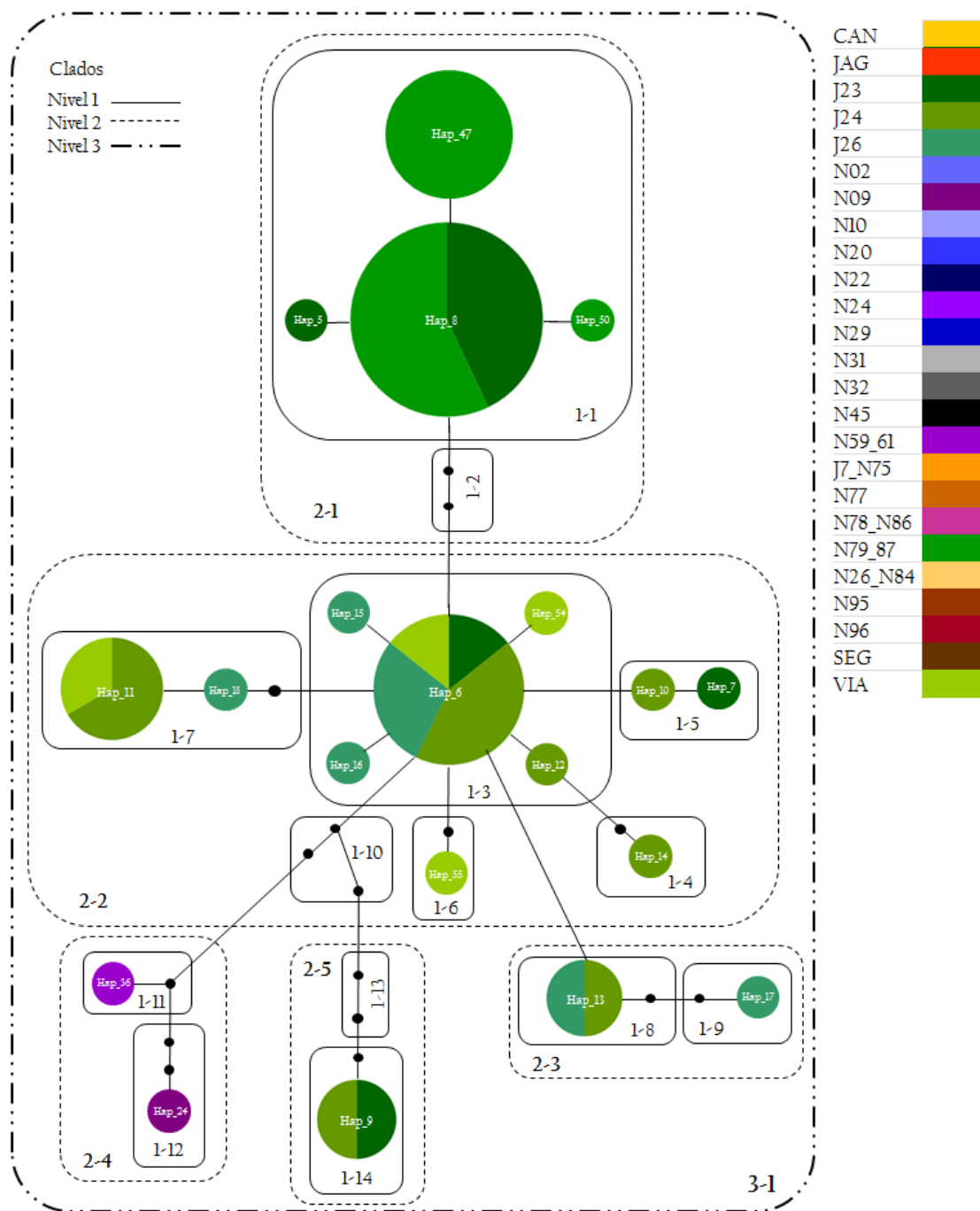
Associação geográfica significativa foi encontrada para o clado 3-1, com eventos de expansão populacional (Etapas da chave de inferência: 1-2-11-12 (Não): *Expansão populacional contígua*).

### **-Rede C**

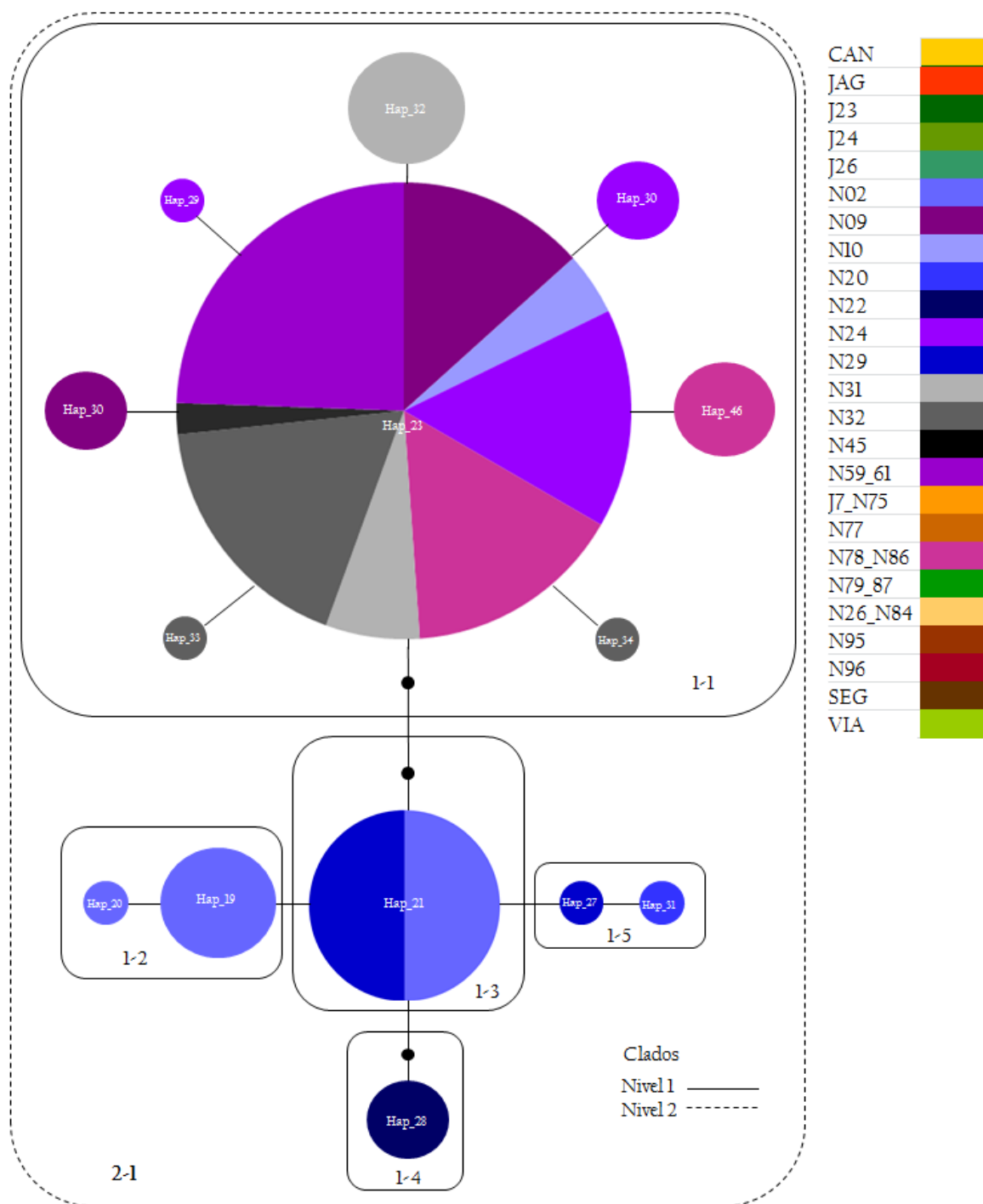
A rede C (Figura 6) é composta por 14 haplótipos de populações do litoral: Mucuri-BA (N32), Itaúnas-ES (N31), Macaé-RJ (N22), Arraial do Cabo-RJ (N20), São Sebastião-SP (N29), Bertiooga-SP (N02), São Francisco do Sul-SC (N24), Penha-SC (N78\_N86), Camboriú-SC (N10), Governador Celso Ramos-SC (N09) e Florianópolis-SC (N59\_N61) e uma população do interior da Bahia: Mucugê-BA (N45).

O Hap\_23 apresenta o maior *outgroup weight* (0,5364) indicando que tem a maior probabilidade de ser o haplótipo mais ancestral da genealogia. Em concordância com esse resultado, ocupa uma posição interior na rede e tem a maior frequência e distribuição geográfica (CLEMENT; POSADA; CRANDALL, 2000; TEMPLETON, 2006). Esta rede também apresentou forma de estrela.

Na Tabela 4 estão descritas as associações geográficas significativas, encontradas nos diferentes clados da rede C.



**Figura 5.** Rede B de haplótipos não enraizada contendo 20 haplótipos do gene mitocondrial *COI* de populações de *Drosophila meridionalis*. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. Haplótipos hipotéticos, não amostrados, são representados por pequenos círculos pretos. As cores dentro de cada haplótipo representam as diferentes populações, sendo essas as mesmas apresentadas na Figura 1.



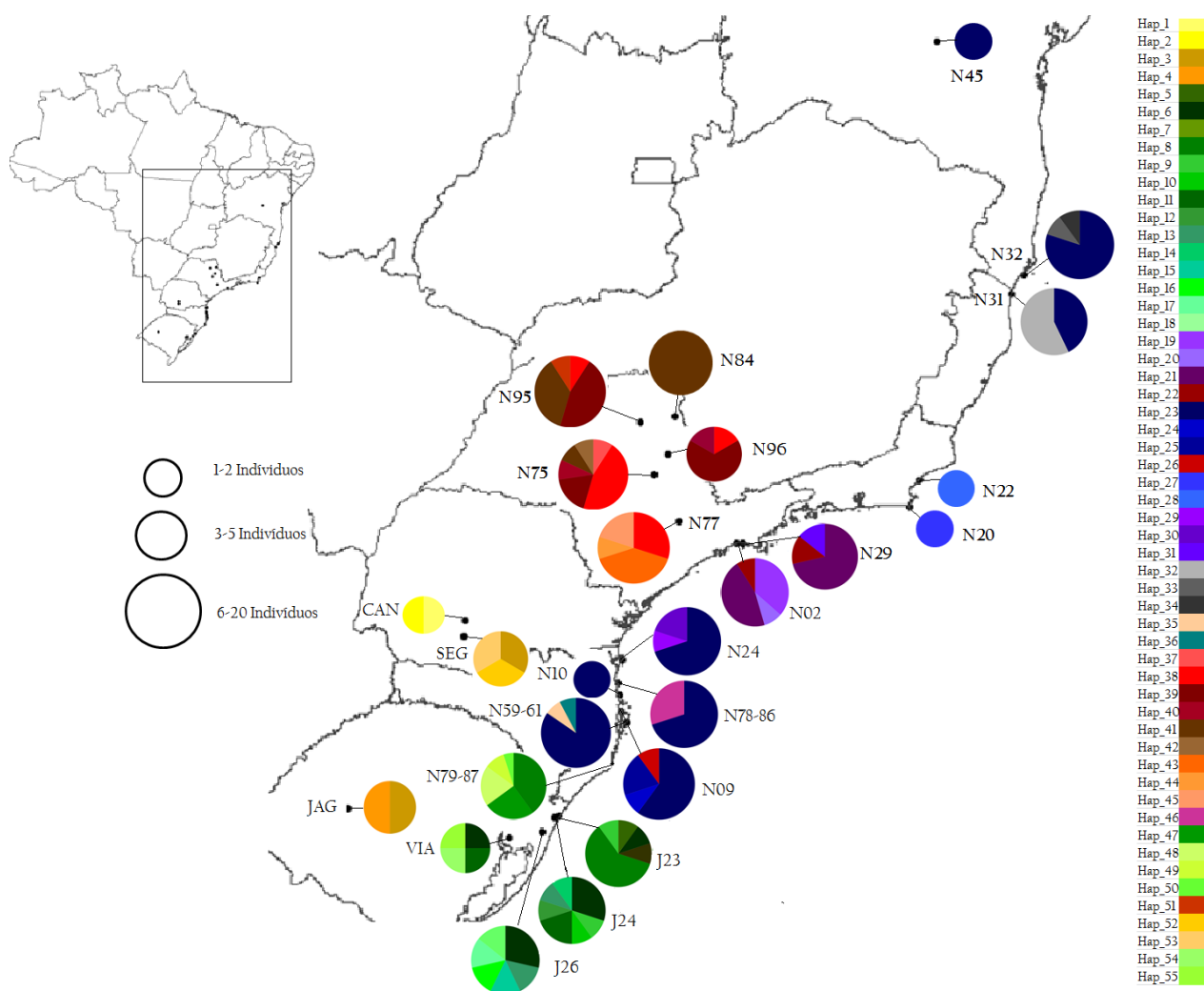
**Figura 6.** Rede C de haplótipos não enraizada contendo 14 haplótipos do gene mitocondrial *COI* de populações de *Drosophila meridionalis*. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. Haplótipos hipotéticos, não amostrados, são representados por pequenos círculos pretos. As cores dentro de cada haplótipo representam as diferentes populações, sendo essas as mesmas apresentadas na Figura 1.

**Tabela 4.** Resultados significativos do agrupamento de clados não enraizados da rede C, obtida a partir de amostras do gene mitocondrial *COI* de populações de *Drosophila meridionalis*, de acordo com a chave de inferências filogeográficas de Templeton (2011), disponível em <http://darwin.uvigo.es/software/geodis.html>

<b>CLADO</b>	<b>Etapas da Chave de Inferências</b>
<b>1-1</b>	1-2-3-4 (Não): Fluxo gênico restrito com isolamento por distância
<b>2-1</b>	1-2-11- 12 (Não): Expansão populacional contígua.

A Figura 7 representa a distribuição dos 55 haplótipos do gene mitocondrial *COI*. As populações de Florianópolis (N59\_N61) e Governador Celso Ramos (N09) do litoral de Santa Catarina estão em simpatria com haplótipos que pertencem às três redes, sendo em maior número de indivíduos e haplótipos na rede C.





**Figura 7.** Distribuição dos 55 haplótipos do gene mitocondrial *COI* das 25 populações de *Drosophila meridionalis*. As populações estão representadas por círculos, cujo tamanho é proporcional ao número de indivíduos que os possuem. As cores na gama azul, roxo e cinza representam as populações da rede C, as cores da gama verde representam as populações da rede B e as cores da gama amarela e vermelha representam as populações da rede A.

### 1.2 AMOVA, testes de Mantel, testes de neutralidade e *Mismatch Distribution*

Os resultados da Análise de Variância Molecular (AMOVA) global e considerando diferentes agrupamentos, foram descritos na Tabela 5. Em todos os agrupamentos testados, foi observada estrutura populacional significativa e principalmente no agrupamento das três redes (A, B e C) definidas pela genealogia de haplótipos.

**Tabela 5.** Análise de Variância Molecular (AMOVA), para as populações de *Drosophila meridionalis*, baseadas nas sequências do gene mitocondrial *COI*.

	<b>Fonte de Variação</b>	<b>G.L</b>	<b>% variação</b>	
<b>AMOVA Global</b>				
$\Phi_{st}=0,75435$ ; $p<0,000$	Entre as populações	24	75,43	
	Dentro das populações	167	24,57	
<b>AMOVA Populações Int-Lit</b>				
$\Phi_{ct}=0,70350$ ; $p<0,000$	Entre os grupos	1	70,35	
	$\Phi_{sc}=0,49728$ ; $p<0,000$	Entre populações dentro dos grupos	23	14,74
	$\Phi_{st}=0,85094$ ; $p<0,000$	Dentro das populações	167	14,91
<b>AMOVA - Redes A B C</b>				
$\Phi_{ct}=0,88761$ ; $p<0,000$	Entre os grupos	2	88,76	
	$\Phi_{sc}=0,57163$ ; $p<0,000$	Entre populações dentro dos grupos	27	6,42
	$\Phi_{st}=0,95186$ ; $p<0,000$	Dentro das populações	162	4,81

G.L. = graus de liberdade. Int= Interior. Lit= Litoral.

As comparações par a par entre os valores de  $F_{ST}$  das populações de *D. meridionalis* estão apresentadas na Tabela 6. Os valores de  $F_{ST}$  mais altos e significativos estão entre as populações que compõem as três redes.

**Tabela 6.** Valores de  $F_{ST}$  calculados entre pares de populações de *Drosophila meridionalis*, para o gene mitocondrial *COI*. Os asteriscos representam os valores de p significativos ( $p \leq 0.05$ ).

	CAN	JAG	J23	J24	J26	N02	N09	N10	N20	N22	N24	N29	N31	N32	N45
<b>CAN</b>	0,00														
<b>JAG</b>	0,27	0,00													
<b>J23</b>	0,92*	0,94*	0,00												
<b>J24</b>	0,91*	0,93*	0,29*	0,00											
<b>J26</b>	0,94*	0,95*	0,33*	-0,01	0,00										
<b>N02</b>	0,81*	0,84*	0,71*	0,68*	0,68*	0,00									
<b>N09</b>	0,71*	0,76*	0,58*	0,55*	0,55*	0,15*	0,00								
<b>N10</b>	0,97	0,99	0,83*	0,81*	0,87*	0,17	-0,29	0,00							
<b>N20</b>	0,95	0,98	0,82	0,78	0,84	-0,52	-0,29	1,00	0,00						
<b>N22</b>	0,97	0,99	0,83*	0,80*	0,86*	0,05	0,09	1,00	1,00	0,00					
<b>N24</b>	0,98*	0,99*	0,88*	0,88*	0,92*	0,42*	0,04	-0,25	0,87	0,91*	0,00				
<b>N29</b>	0,72*	0,78*	0,67*	0,63*	0,62*	-0,08	0,08	0,06	-0,77	-0,06	0,39*	0,00			
<b>N31</b>	0,98*	0,98*	0,87*	0,86*	0,91*	0,40*	0,03	0,26	0,88	0,91*	0,35*	0,35*	0,00		
<b>N32</b>	0,99*	0,99*	0,89*	0,88*	0,92*	0,42*	0,03	-0,32	0,90	0,93*	0,04	0,39*	0,38*	0,00	
<b>N45</b>	0,95	0,98	0,80	0,78	0,84	-0,09	-0,93	0,00	1,00	1,00	-0,85	-0,32	0,00	-1,00	0,00
<b>N59_N61</b>	0,79*	0,82*	0,63*	0,62*	0,62*	0,19*	-0,08	-0,30	-0,14	0,20	0,01	0,14	0,02*	-0,01	-0,95
<b>J7_N75</b>	0,66*	0,74*	0,93*	0,92*	0,93*	0,85*	0,80*	0,94*	0,94	0,94*	0,96*	0,82*	0,96*	0,96*	0,94
<b>N77</b>	0,38	0,52*	0,89*	0,88*	0,89*	0,81*	0,75*	0,88*	0,86	0,88*	0,93*	0,76*	0,92*	0,93*	0,87
<b>N78_N86</b>	0,98*	0,99*	0,89*	0,88*	0,92*	0,43*	0,04*	-0,09	0,89	0,92*	0,15	0,40*	0,41*	0,13	-0,56
<b>N79_N87</b>	0,49*	0,55*	0,15	0,21*	0,18*	0,43*	0,33*	0,36	0,24	0,36	0,52*	0,36*	0,49*	0,51*	0,21
<b>N26_N84</b>	0,96*	0,97*	0,96*	0,95*	0,97*	0,88*	0,83*	1,00*	1,00	1,00*	0,99*	0,86*	0,99*	0,99*	1,00
<b>N95</b>	0,69*	0,77*	0,93*	0,92*	0,94*	0,86*	0,80*	0,95*	0,95	0,95*	0,97*	0,82*	0,96*	0,97*	0,95
<b>N96</b>	0,73*	0,81*	0,93*	0,92*	0,94*	0,84*	0,76*	0,96*	0,95	0,95*	0,97*	0,789*	0,97*	0,97*	0,95
<b>SEG</b>	0,27	0,53	0,89*	0,88*	0,89*	0,77*	0,69*	0,87	0,80	0,86	0,95*	0,68*	0,94*	0,95*	0,80
<b>VIA</b>	0,93	0,95*	0,31*	-0,07	-0,02	0,65*	0,49*	0,86	0,81	0,85	0,93*	0,56*	0,92*	0,93*	0,81

Tabela 6. Continuação.

	<b>N59_N61</b>	<b>J7_N75</b>	<b>N77</b>	<b>N78_N86</b>	<b>N79_N87</b>	<b>N26_N84</b>	<b>N95</b>	<b>N96</b>	<b>SEG</b>	<b>VIA</b>
<b>CAN</b>										
<b>JAG</b>										
<b>J23</b>										
<b>J24</b>										
<b>J26</b>										
<b>N02</b>										
<b>N09</b>										
<b>N10</b>										
<b>N20</b>										
<b>N22</b>										
<b>N24</b>										
<b>N29</b>										
<b>N31</b>										
<b>N32</b>										
<b>N45</b>										
<b>N59_N61</b>	0,00									
<b>J7_N75</b>	0,84*	0,00								
<b>N77</b>	0,80*	0,22*	0,00							
<b>N78_N86</b>	0,01	0,96*	0,93*	0,00						
<b>N79_N87</b>	0,39*	0,59*	0,54*	0,52*	0,00					
<b>N26_N84</b>	0,86*	0,31*	0,23*	0,99*	0,60*	0,00				
<b>N95</b>	0,84*	0,02	0,27*	0,97*	0,59*	0,34*	0,00			
<b>N96</b>	0,82*	0,06	0,31*	0,97*	0,54*	0,65*	-0,06	0,00		
<b>SEG</b>	0,76*	0,63*	0,27	0,95*	0,47*	0,80*	0,66*	0,64*	0,00	
<b>VIA</b>	0,58*	0,93*	0,87*	0,93*	0,13*	0,98*	0,94*	0,93*	0,86*	0,00

O teste de Mantel foi realizado para os clados nos quais foram detectados eventos de fluxo gênico restrito com isolamento por distância pela NCPA, sendo 2-5 e 4-1 da rede A e o clado 1-1 da rede C. Para o Clado 4-1 da rede A foi encontrada uma correlação positiva e significativa ( $rY1=0.457329$ ;  $p=0.0001$ ) entre as distâncias genética e geográfica, indicando que quanto maior a distância geográfica, maior é a diferenciação genética entre as populações. Para o clado 2-5 da rede A e o clado 1-1 da rede C não foram encontradas correlações significativas (Tabela 7).

**Tabela 7.** Teste de Mantel para as redes de haplótipos das populações de *Drosophila meridionalis*, baseadas nas sequências do gene mitocondrial *COI*.

Clado	Rede	Coefficiente de Correlação (rY1)	p valor *
Clado 2-5	A	0,374756	0,2705
Clado 4-1	A	0,457326	0,0001*
Clado 1-1	C	-0,177472	0,7978

\*p valor  $\leq 0.05$

As análises de *Mismatch Distribution* e os testes de neutralidade (Fs de Fu, D de Tajima e R2) (Tabela 8) foram realizados para os clados 1-11 e 3-3 da rede A, o clado 3-1 da rede B e o clado 2-1 da rede C, para os quais foram detectados eventos de expansão populacional pela NCPA.

Os gráficos resultantes mostrados na Figura 8 são irregulares e multimodais, mostrando um padrão diferente do que se esperaria sob o modelo de expansão populacional. No entanto, os valores de p para a rejeição do teste da significância do ajustamento, comparando as SSD simuladas e observadas, são observados valores entre 0,05 e 0,95 o que sugere que não pode ser rejeitada a hipótese de expansão populacional para os clados, (SCHNEIDER; EXCOFFIER, 1999) com exceção do clado 1-11 da rede A.

A *Mismatch Distribution*, o D de Tajima e o R2 detectaram expansão populacional para o clado 3-3 da rede A e para o clado 1-11 da rede A, o R2 apresentou um valor pequeno e significativo, detectando expansão populacional.

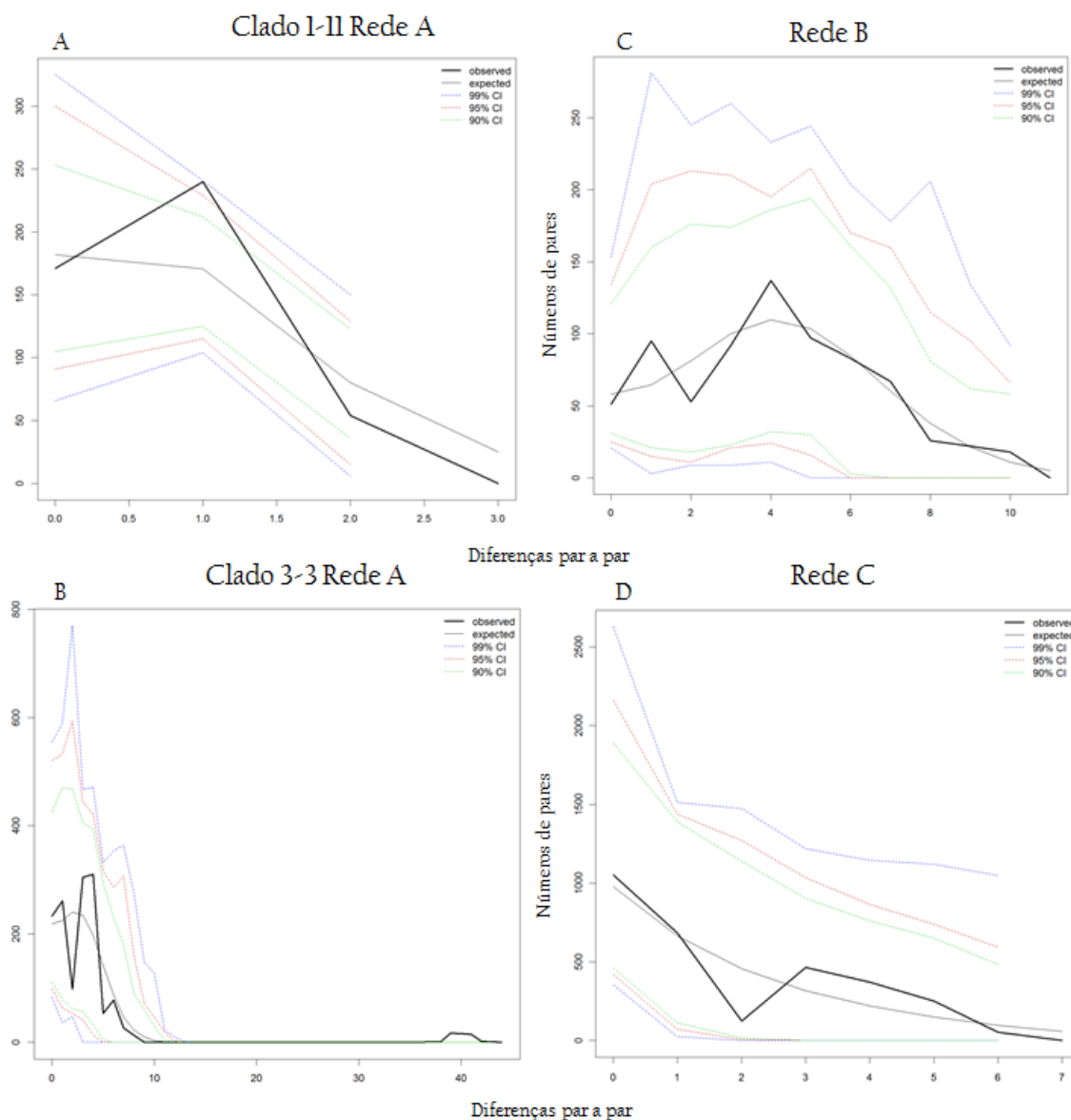
Foram detectados eventos de expansão populacional nas redes B e C, pela *Mismatch Distribution* e os diferentes testes de neutralidade (Tabela 8), em concordância com a NCPA.

**Tabela 8.** Resultados dos testes de neutralidade (Fs de Fu, D de Tajima e R2) e *Mismatch Distribution* para populações de *Drosophila meridionalis*, utilizando sequências do gene mitocondrial *COI*.

Clado	Rede	<i>Mismatch Distribution</i>	D de Tajima	Fs de Fu	R2
<b>Clado 1-11</b>	A	$\Theta_0=0,00$ $\Theta_1=99999$ SSD=0,025 pSSD=0,036	-0,007 p=0,49	-0,26 p=0,39	0,118* p=0,00
<b>Clado 3-3</b>	A	$\Theta_0=0,000$ $\Theta_1=5,950$ SSD=0,023 pSSD=0,240	-2,07* p=0,00	-0,82 p=0,43	0,102* p=0,00
<b>Clado 3-1</b>	B	$\Theta_0=0,007$ $\Theta_1=12,391$ SSD=0,005 pSSD=0,686	-1,44* p=0,05	-8,67** p=0,00	0,106* p=0,00
<b>Clado 2-1</b>	C	$\Theta_0=0,02$ $\Theta_1=2,070$ SSD=0,019 pSSD=0,626	-1,17 p=0,11	-5,13** p=0,019	0,095* p=0,00

\*p<0,05; \*\*p<0,02.

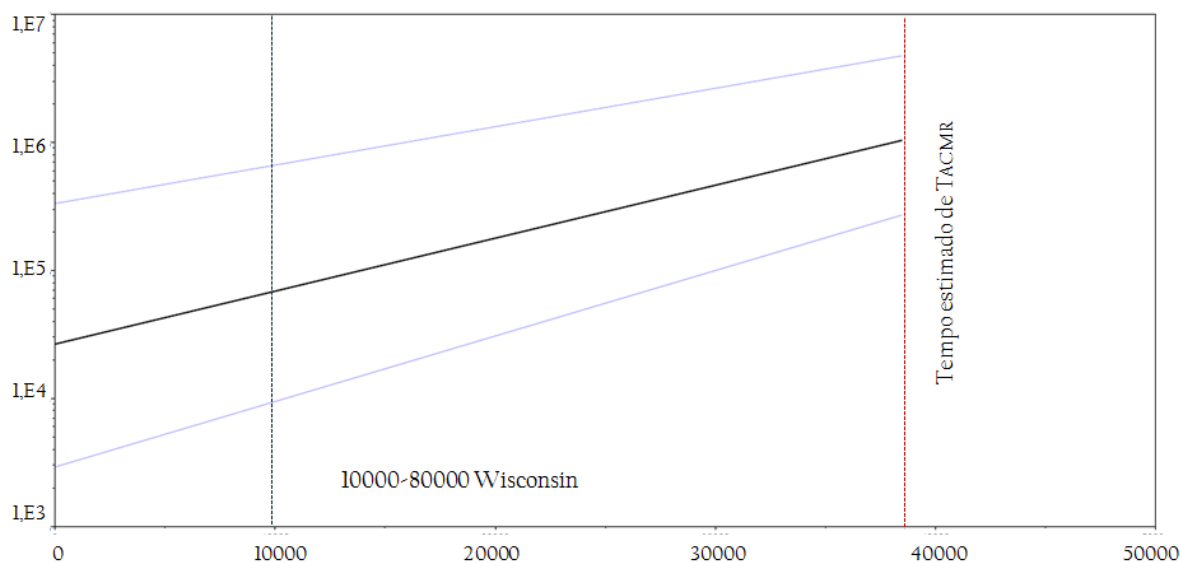
$\theta_1$ , tamanho populacional antes da expansão;  $\theta_2$ , tamanho populacional após a expansão; SSD, teste do desvio da soma dos quadrados (*Sum of squared deviation*).



**Figura 8.** Representação gráfica das diferenças par a par entre as sequências (*Mismatch Distribution*) do gene mitocondrial COI dos indivíduos de *Drosophila meridionalis*. A: Indivíduos alocados no clado 1-11 da rede A. B: Indivíduos alocados no clado 3-3 da rede A. C: Indivíduos agrupados na rede B de haplótipos. D: Indivíduos agrupados na rede C de haplótipos. Linhas mais fortes representam as diferenças par a par observadas, enquanto linha mais tênue representa as diferenças par a par esperadas com eventos de expansão populacional.

### 1.3 Análises Bayesianas

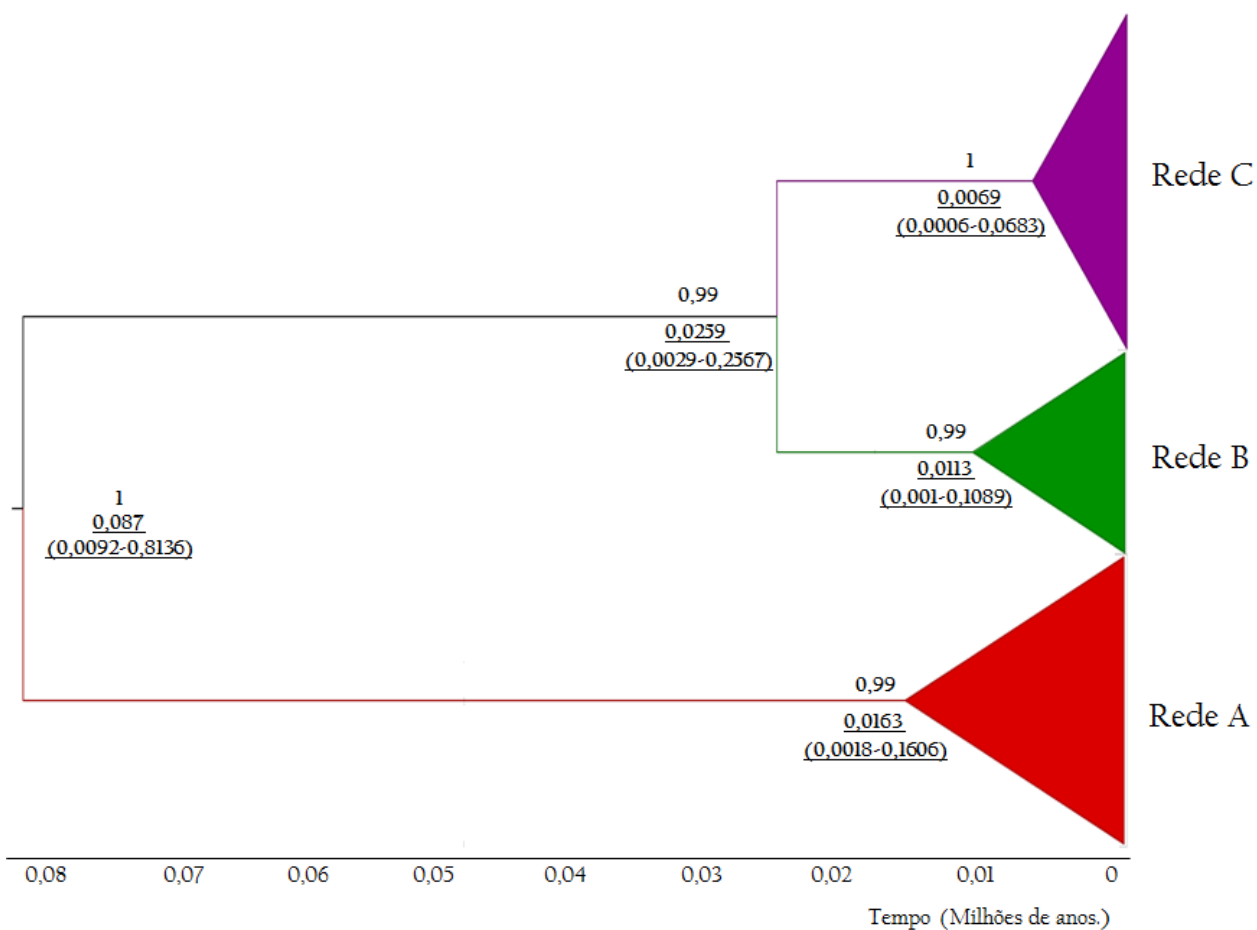
A *Bayesian Skyline Plot* (Figura 9) mostra que o tempo calculado do ancestral comum mais recente da amostra analisada ( $T_{MRCA}$ ), considerando as sequências do gene mitocondrial *COI*, é de 40 mil anos atrás. Uma redução no tamanho populacional próximo ao tempo presente é sugerida no gráfico, fato que é concordante com a atual distribuição geográfica das FTSS.



**Figura 9.** Reconstrução da história demográfica, da espécie *Drosophila meridionalis* a partir das sequências do gene mitocondrial *COI*. A curva negra representa a variação na média do tamanho efetivo populacional, em relação ao tempo em anos atrás. As linhas azuis representam intervalos de confiança (95%). A linha vermelha vertical tracejada indica o tempo para o ancestral comum mais recente da amostra ( $T_{MRCA}$ ) da espécie. O tamanho populacional é medido em unidades de  $\log N_e \cdot \tau$ .

Foram formados três agrupamentos com as sequências do gene mitocondrial *COI*, com alto suporte estatístico, pelas análises de coalescência bayesiana (*Maximum Clade Credibility Tree-MCCT*) (Figura 10). Estes três agrupamentos são coincidentes com as redes A, B e C definidas pelo programa TCS por meio da parcimônia estatística (Figura 4, 5, 6).





**Figura 10.** Relações entre as linhagens e o  $T_{MRCA}$  de 192 indivíduos em 25 populações de *Drosophila meridionalis* baseado em análises de coalescência bayesiana (*Maximum Clade Credibility Tree*) das sequências do gene mitocondrial *COI*. O número superior corresponde ao suporte estatístico do nó (probabilidade a posteriori) e os números sublinhados correspondem ao  $T_{MRCA}$  com seus intervalos de confiança (95%). A escala de tempo está em milhões de anos.

## 2.0 Análises das sequências do gene nuclear *period*.

Foram sequenciados fragmentos de 496 pares de bases (pb) do gene nuclear *period* de 86 indivíduos machos de 22 populações de *D. meridionalis*, gerando 30 haplótipos. Para as sequências do gene nuclear *period* o modelo de substituição nucleotídica que melhor explica a evolução das sequências é HKY+G (AICc 1861.517).

Os índices de diversidade genética calculados para as populações analisadas de *D. meridionalis* estão apresentados na Tabela 9.

**Tabela 9.** Índices de diversidade genética e haplótipos de populações de *Drosophila meridionalis*, utilizando sequências do gene nuclear *period*.

	<b>N</b>	<b>h</b>	<b>S</b>	<b>Hd</b>	$\pi_1$	<b>k</b>	<b>Haplótipos*</b>
<b>N32</b>	2	1	0	0,000	0,00000±0,00000	0,0	Hap 14 (2)
<b>N45</b>	1	1	0	1,000	--	--	Hap 15 (1)
<b>N31</b>	2	2	2	1,000	0,00403±0,00202	2,0	Hap 12 (1) ; Hap 13 (1)
<b>N20</b>	4	4	5	1,000	0,00571±0,00155	2,8	Hap 5 (1) ; Hap 8 (1) ; Hap 9 (1) ; Hap 10 (1)
<b>N26/N84</b>	5	4	5	0,900	0,00484±0,00139	2,4	Hap 17 (2); Hap 23 (1) ; Hap 24 (1) ; Hap 25 (1)
<b>J07/N75</b>	5	5	6	1,000	0,00645±0,00123	3,2	Hap 1 (1) ; Hap 16 (1) ; Hap 17 (1) ; Hap 18 (1) ; Hap 19 (1)
<b>N77</b>	6	3	3	0,600	0,00242±0,00109	1,2	Hap 1 (4) ; Hap 20 (1) ; Hap 21 (1)
<b>N95</b>	5	3	5	0,700	0,00403±0,00188	2,0	Hap 17 (3) ; Hap 25 (1) ; Hap 29 (1)
<b>N96</b>	5	3	5	0,800	0,00365±0,00085	1,8	Hap 1 (2) ; Hap 17 (2) ; Hap 30 (1)
<b>N29</b>	6	3	2	0,733	0,00188±0,00052	0,9	Hap 5 (3) ; Hap 6 (2) ; Hap 11 (1)
<b>SEG</b>	3	2	1	0,667	0,00134±0,00063	0,6	Hap 1 (2) ; Hap 23 (1)
<b>CAN</b>	2	1	0	0,000	0,00000±0,00000	0,0	Hap 1 (2)
<b>N24</b>	5	1	0	0,000	0,00000±0,00000	0,0	Hap 5 (5)
<b>N78-N86</b>	4	1	0	0,000	0,00000±0,00000	0,0	Hap 5 (3) ; Hap 22 (1)
<b>N10</b>	2	2	2	1,000	0,00403±0,00202	2,0	Hap 2 (1) ; Hap 5 (1)
<b>N09</b>	5	2	1	0,400	0,00081±0,00048	0,4	Hap 5 (4) ; Hap 6 (1)
<b>N59-N61</b>	1	1	0	1,000	--	--	Hap 5 (1)
<b>N79-N87</b>	5	3	2	0,700	0,00244±0,00091	1,1	Hap 5 (1) ; Hap 7 (1) ; Hap 26 (1) ; Hap 27 (1) ; Hap 28 (1)
<b>J23</b>	4	4	8	1,000	0,00874±0,00205	4,3	Hap 2 (1) ; Hap 3 (1) ; Hap 4 (1) ; Hap 5 (1)
<b>J24</b>	6	4	5	0,867	0,00390±0,00135	1,9	Hap 2 (1) ; Hap 5 (2) ; Hap 6 (2) ; Hap 7 (1)
<b>J26</b>	5	3	2	0,800	0,00214±0,00059	1,0	Hap 2 (1) ; Hap 5 (2) ; Hap 6 (2)
<b>VIA-1</b>	3	3	2	1,000	0,00269±0,00090	1,3	Hap 5 (1) ; Hap 6 (1) ; Hap 14 (1)
<b><i>D. meridionalis</i></b>	<b>86</b>	<b>30</b>	<b>27</b>	<b>0,850</b>	<b>0,01017±0,00065</b>	<b>4,0</b>	

\*Haplótipos segundo a NCPA, dentro dos parêntesis esta o número de indivíduos por haplótipos. N = número de indivíduos; h = número de haplótipos; S = número de sítios polimórficos; Hd = diversidade haplotípica;  $\pi_1$  = diversidade nucleotídica (desvio-padrão); k = número médio de diferenças nucleotídicas. Os códigos das coletas são os mesmos apresentados na Tabela 1.

A diversidade nucleotídica total ( $\pi$ ) para as populações foi de  $0,01017 \pm 0,00065$ . O maior índice de diversidade nucleotídica foi calculado para a população Arroio Teixeira-SC (J23) e os menores para as populações Mucuri-BA (N32), Cantagalo-PR (CAN), São Francisco do Sul-SC (N24), Penha-SC (N78\_N86). Para as populações Mucugê-BA (N45) e Florianópolis-SC (N59\_N61) não foi possível calcular o índice de diversidade devido ao tamanho amostral (apenas um indivíduo).

### 2.1 Nested Clade Phylogeographic Analysis (NCPA)

Foi gerada apenas uma rede com o programa TCS (CLEMENT; POSADA; CRANDALL, 2000), conectando todos os haplótipos obtidos das sequências do gene *period* (Figura 11).

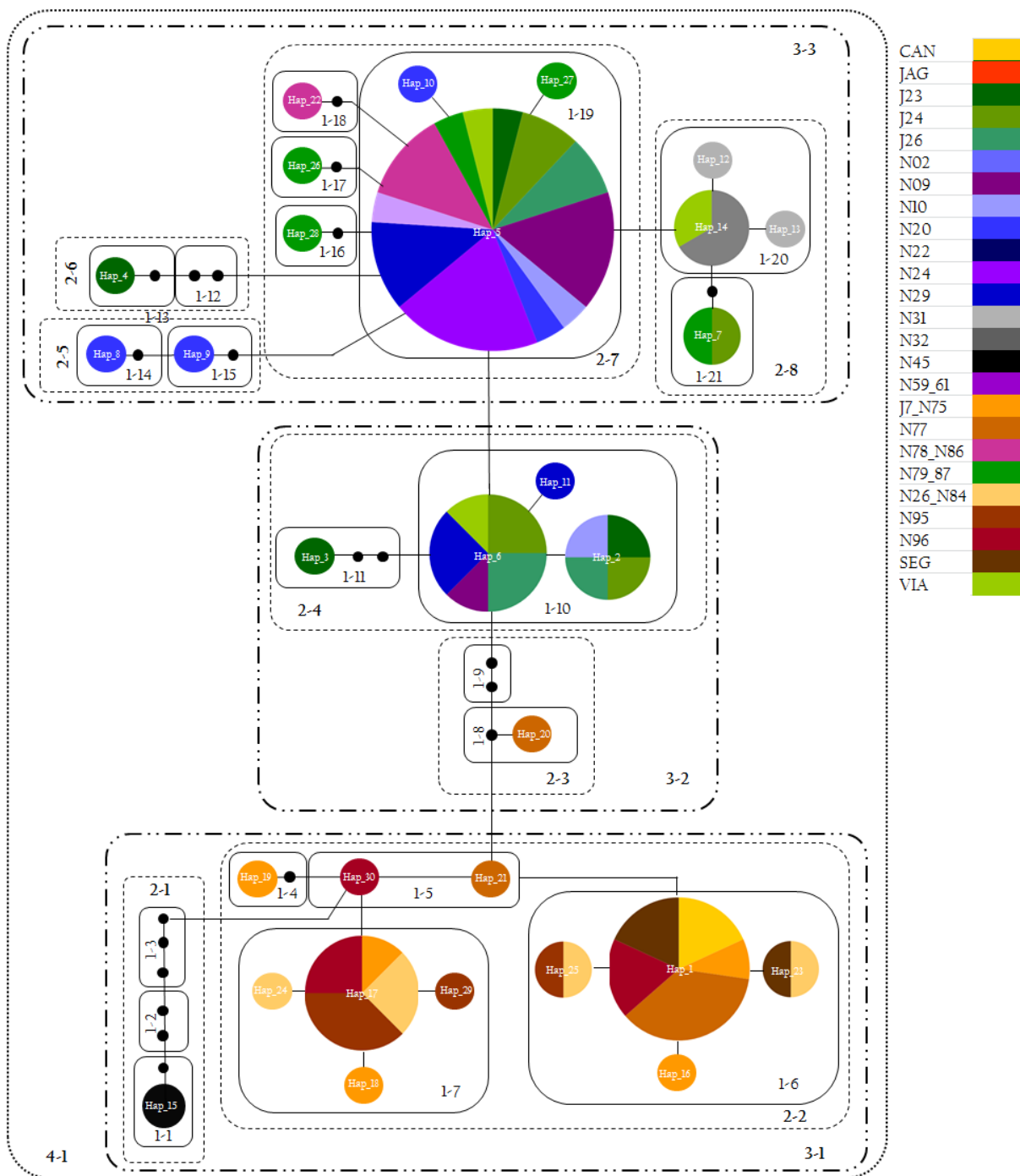
O Hap\_5 apresentou o maior *outgroup weight* (0,1535), indicando que este tem a maior probabilidade de ser o haplótipo mais ancestral da genealogia. Em concordância com tal resultado, esse haplótipo ocupa uma posição interior na rede e tem a maior frequência e ampla distribuição geográfica (CLEMENT; POSADA; CRANDALL, 2000; TEMPLETON, 2006), encontrando-se nas populações do litoral.

Associações geográficas significativas foram observadas para os clados 2-2 que inclui todas as populações do interior e o clado 3-3 que inclui as populações do litoral (Tabela 10).

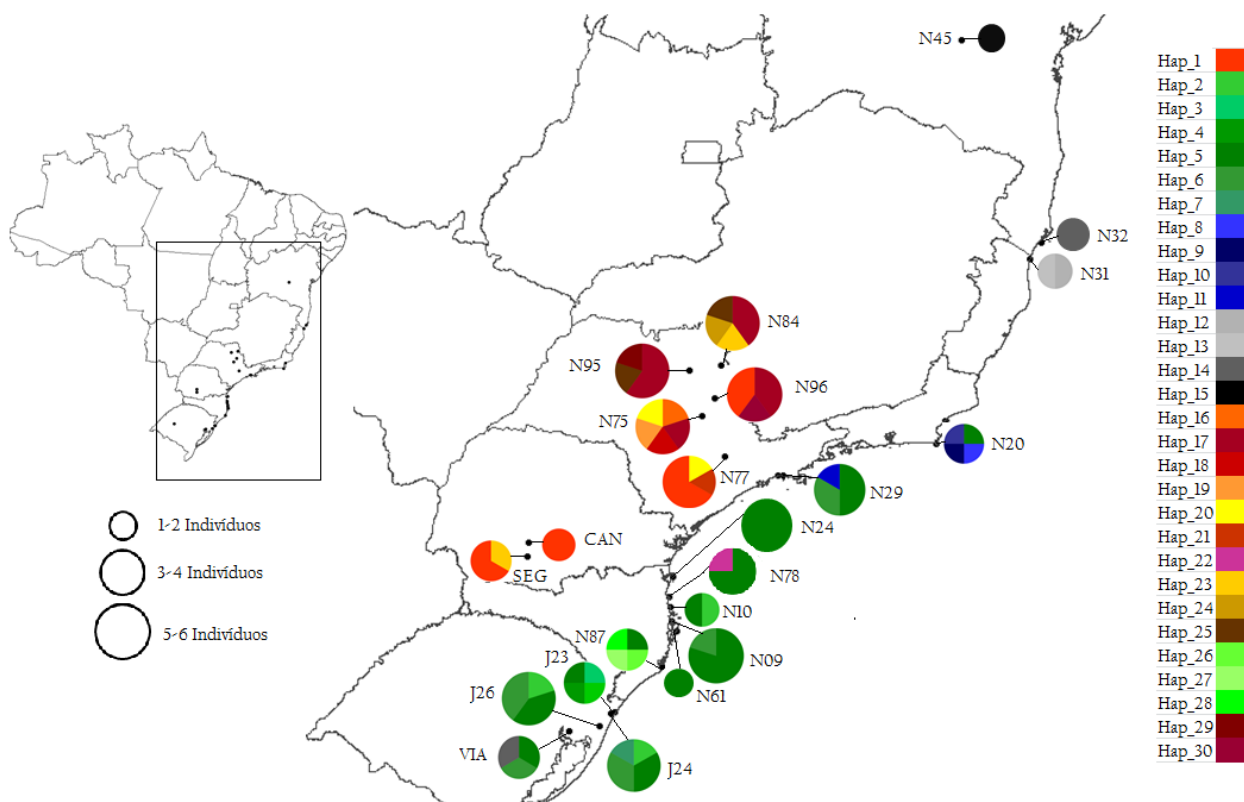
**Tabela 10.** Resultados significativos do agrupamento de clados não enraizados da rede de haplótipos, obtida a partir de amostras do gene nuclear *period* de populações de *Drosophila meridionalis*, de acordo com a chave de inferências filogeográficas de Templeton (2011), disponível em <http://darwin.uvigo.es/software/geodis.html>

CLADO	Etapas da Chave de Inferências
2-2	1-2-3-4 (Não): Fluxo gênico restrito com isolamento por distância
3-3	1-2-11-12 (Não): Expansão populacional contígua

A distribuição dos 30 haplótipos do gene nuclear *period* (Figura 12), sugere a estruturação das populações de *D. meridionalis* em dois grupos: populações do interior e do litoral. Além disso, é possível visualizar a ampla distribuição do Hap\_5 nas populações do litoral.



**Figura 11.** Rede de haplótipos não enraizada contendo 30 haplótipos do gene nuclear *period* de populações de *Drosophila meridionalis*. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. Haplótipos hipotéticos, não amostrados, são representados por pequenos círculos pretos. As cores dentro de cada haplótipo representam as diferentes populações, sendo essas as mesmas apresentadas na Figura 1.



**Figura 12.** Distribuição dos 30 haplótipos do gene nuclear *period* nas 22 populações de *Drosophila meridionalis*. As populações estão representadas por círculos, cujos tamanhos são proporcionais aos números de indivíduos que os possuem. As cores na gama azul, roxo, e cinza e verde representam as populações do litoral e as cores da gama amarela e vermelha representam as populações do interior.

## 2.2 AMOVA, testes de Mantel, testes de neutralidade e *Mismatch Distribution*

Os resultados da Análise de Variância Molecular (AMOVA) global e considerando diferentes agrupamentos, estão descritos na Tabela 11. Em todos os agrupamentos testados, foi observada estrutura populacional significativa e principalmente no agrupamento das populações do interior e o litoral.

Para o gene nuclear *period* a AMOVA mostra valores  $\Phi_{ct}$  maiores entre os grupos formados pelas populações do interior e o litoral, do que nos grupos considerando as três redes obtidas com as sequências do gene *COI*.

**Tabela 11.** Análise de Variância Molecular (AMOVA) para as populações de *Drosophila meridionalis*, baseadas nas sequências do gene nuclear *period*.

	Fonte de Variação	G,L	% variação
<b>AMOVA Global</b>			
$\Phi_{st}=0,667$ ; $p<0,000$	Entre as populações	24	66,70
	Dentro das populações	167	33,30
<b>AMOVA Populações Int-Lit</b>			
$\Phi_{ct}=0,71898$ ; $p<0,000$	Entre os grupos	1	71,90
	Entre populações dentro dos grupos	23	7,09
$\Phi_{st}=0,78986$ ; $p<0,000$	Dentro das populações	167	21,01
<b>AMOVA - Redes A B C</b>			
$\Phi_{ct}=0,64954$ ; $p<0,000$	Entre os grupos	2	64,95
	Entre populações dentro dos grupos	27	9,62
$\Phi_{st}=0,74581$ $p<0,000$	Dentro das populações	162	25,41

G.L. = graus de liberdade. Int= Interior. Lit= Litoral.

O teste de Mantel realizado para o clado 2-2, em que foi detectado fluxo gênico restrito com isolamento por distância pela NCPA, não foi significativo (Tabela 12). O teste de Mantel para o Clado 4-1 (rede total) mostrou uma correlação positiva e significativa ( $r_{Y1}=0.2973$ ;  $p=0.003$ ) entre as distâncias genéticas e as geográficas das populações, indicando que quanto maior a distância geográfica, maior é a diferenciação genética das populações, o que não foi detectado pela NCPA.

**Tabela 12.** Teste de mantel para a rede de haplótipos das populações de *Drosophila meridionalis*, baseadas nas sequências do gene nuclear *period*.

Clado	Coefficiente de Correlação	p valor
Clado 2-2	0,3853	0,079
Clado 4-1	0,2973	0,003*

\* $p\leq 0.05$

As comparações par a par entre os valores de  $F_{ST}$  das populações de *D. meridionalis* estão apresentadas na Tabela 13. Os valores de  $F_{ST}$  mais altos e significativos estão entre as populações do interior e o litoral.

**Tabela 13.** Valores de  $F_{ST}$  calculados entre pares de populações de *Drosophila meridionalis*, para o gene *period*. Os asteriscos representam os valores de p significativos ( $p \leq 0.05$ ).

	CAN	J23	J24	J26	N09	N10	N20	N24	N29	N31	N32	N45	N59_N61
CAN	0,00												
J23	0,59	0,00											
J24	0,82*	0,04	0,00										
J26	0,86*	0,01	-0,18	0,00									
N09	0,95*	0,11	-0,02	0,08	0,00								
N10	0,83	-0,26	-0,30	-0,36	0,11	0,00							
N20	0,77	0,11	0,26	0,30	0,24	0,11	0,00						
N24	1,00*	0,20*	0,20	0,38	0,00	0,47	0,29	0,00					
N29	0,87*	0,08	-0,10	-0,11	-0,01	-0,09	0,29	0,26	0,00				
N31	0,93	0,19	0,40	0,57*	0,68*	0,40	0,36	0,86*	0,56*	0,00			
N32	1,00	0,17	0,37	0,58*	0,74*	0,50	0,36	1,00*	0,57*	0,00	0,00		
N45	1,00	0,57	0,86	0,88	0,95	0,78	0,76	1,00	0,89	0,90	1,00	0,00	
N59_N61	1,00	-0,73	-0,52	-0,25	-1,00	-1,00	-0,56	0,00	-0,40	0,33	1,00	1,00	0,00
J7_N75	0,08	0,56*	0,69*	0,69*	0,74*	0,60*	0,68*	0,78*	0,71*	0,69*	0,70*	0,64	0,56
N77	-0,20	0,66*	0,78*	0,79*	0,85*	0,76*	0,77*	0,88*	0,80*	0,84*	0,85*	0,86	0,79
N78_N86	1,00	0,28*	0,45*	0,57*	0,68*	0,62	0,42*	1,00*	0,53*	0,87	1,00	1,00	1,00
N79_N87	0,29	0,63*	0,75*	0,75*	0,80*	0,68*	0,73*	0,83*	0,77*	0,76*	0,77*	0,71	0,67
N26_N84	0,84*	0,16*	0,13	0,25*	0,10	0,12	0,20	0,13	0,21	0,24	0,18	0,86	-0,75
N95	0,47	0,65*	0,77*	0,78*	0,83*	0,72*	0,75*	0,86*	0,79*	0,80*	0,81*	0,77	0,72
N96	0,21	0,63*	0,76*	0,77*	0,82*	0,71	0,74*	0,86*	0,78*	0,79*	0,81*	0,78	0,72
SEG	-0,20	0,64*	0,82*	0,85*	0,92*	0,82	0,79*	0,96*	0,86*	0,90	0,94	0,92	0,89
VIA	0,85	-0,05	-0,23	-0,07	-0,09	-0,20	0,13	0,19	-0,10	0,20	0,17	0,85	-1,00

Tabela 13. Continuação

	<b>J7_N75</b>	<b>N77</b>	<b>N78_N86</b>	<b>N79_N87</b>	<b>N26_N84</b>	<b>N95</b>	<b>N96</b>	<b>SEG</b>	<b>VIA</b>
<b>CAN</b>									
<b>J23</b>									
<b>J24</b>									
<b>J26</b>									
<b>N09</b>									
<b>N10</b>									
<b>N20</b>									
<b>N24</b>									
<b>N29</b>									
<b>N31</b>									
<b>N32</b>									
<b>N45</b>									
<b>N59_N61</b>									
<b>J7_N75</b>	0,00								
<b>N77</b>	0,19	0,00							
<b>N78_N86</b>	0,77*	0,88*	0,00						
<b>N79_N87</b>	-0,06	0,34*	0,82*	0,00					
<b>N26_N84</b>	0,71*	0,80*	0,17	0,76*	0,00				
<b>N95</b>	0,02	0,46*	0,85*	-0,15	0,79*	0,00			
<b>N96</b>	-0,12	0,20	0,85*	-0,14	0,78*	-0,03	0,00		
<b>SEG</b>	0,08	0,01	0,96*	0,23	0,84*	0,43	0,13	0,00	
<b>VIA</b>	0,65*	0,78*	0,52*	0,72*	-0,08	0,76*	0,74*	0,84	0,00



Os resultados das análises da *Mismatch Distribution* geraram gráficos para as sequências do clado 3-3 e o clado 4-1 (rede total), irregulares e multimodais (Figura 13). Esses gráficos mostram um padrão que não sugere expansão populacional, embora os valores de p para SSD sugere que não pode ser rejeitada a hipótese de expansão populacional, para o clado 3-3.

Os testes D' de Tajima e o R2 para o clado 3-3 detectaram expansão populacional (Tabela 14) como foi sugerido pela NCPA (Tabela 10) e para a rede total (4-1) a *Mismatch Distribution* e todos os testes de neutralidade detectaram a ocorrência de expansão populacional (Tabela 14), o que não foi detectado pela NCPA.

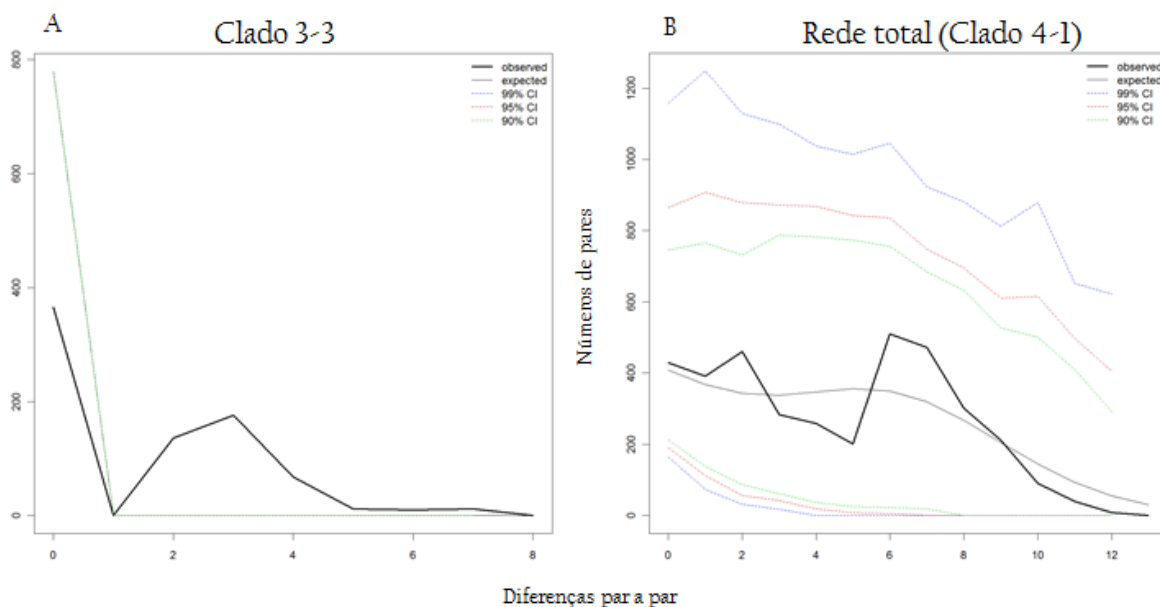
**Tabela 14.** Resultados dos testes de neutralidade (Fs de Fu, D de Tajima e R2) e *Mismatch Distribution* para populações de *Drosophila meridionalis*, aplicados às sequências do gene nuclear *period*.

Clado	<i>Mismatch Distribution</i>	D' de Tajima	Fs de Fu	†R2
<b>Clado 3-3</b>	$\Theta_0=0,000$	-1.81*	-2.36	0,124*
	$\Theta_1=854,266$	0,018	0,097	P=0,00
	SSD=0,370			
	pSSD=0,00			
<b>Clado 4-1</b>	$\Theta_0=0.002$	-1,48*	-11.91**	0,097*
	$\Theta_1=7,985$	0,041	0,00	P=0,00
	SSD=0,008			
	pSSD=0,735			

†Esse teste foi realizado considerando eventos de recombinação.

\*p<0,05; \*\*p<0,02.

$\theta_1$ , tamanho populacional antes da expansão;  $\theta_2$ , tamanho populacional após a expansão; SSD, teste do desvio da soma dos quadrados (*Sum of squared deviation*).



**Figura 13.** Representação gráfica das diferenças par a par entre as sequências (*Mismatch Distribution*) do gene nuclear *period* dos indivíduos de *Drosophila meridionalis*. A: Indivíduos alocados no clado 3-3 B: Indivíduos agrupados na rede total de haplótipos. Linhas mais fortes representam as diferenças par a par observadas, enquanto linha mais tênue representa as diferenças par a par esperadas com eventos de expansão populacional.

### 3.0 Modelagem de Nicho Ecológico

Os modelos de nicho ecológico apresentaram resultados de AUC (*Area Under the Curve*) para o presente de  $0,993 \pm 0,002$ , para o modelo passado CCSM de  $0,990 \pm 0,003$  e para o modelo passado MIROC de  $0,988 \pm 0,004$ . O parâmetro AUC pode apresentar valores entre 0,5 e 1,0, sendo o valor de 0,5 indicativo de que o modelo obtido não é melhor do que um modelo aleatório, valores entre 0,5 a 0,7 correspondem a um modelo de baixa precisão devido à escassa discriminação, e valores maiores que 0,7 correspondem a um modelo de alta precisão e alta discriminação (NEWBOLD et al., 2009).

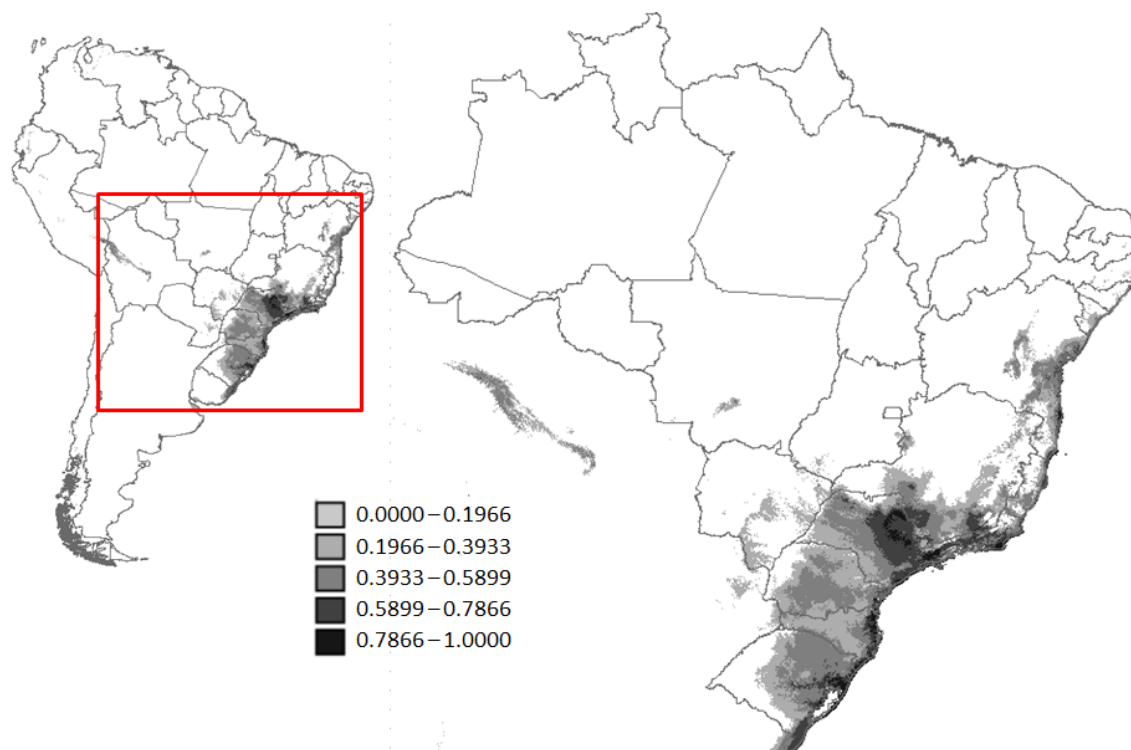
#### - *Projeção atual*

O mapa para o presente (Figura 14) sugere que as variáveis bioclimáticas que mais contribuíram para a elaboração do modelo preditivo são: *Temperature Seasonality* (Bio4) e *Max Temperature of Warmest Mont* (Bio5), com uma porcentagem de contribuição de 19,0 e 18,2 respectivamente. Essa projeção coincide com a distribuição geográfica conhecida, que mostra

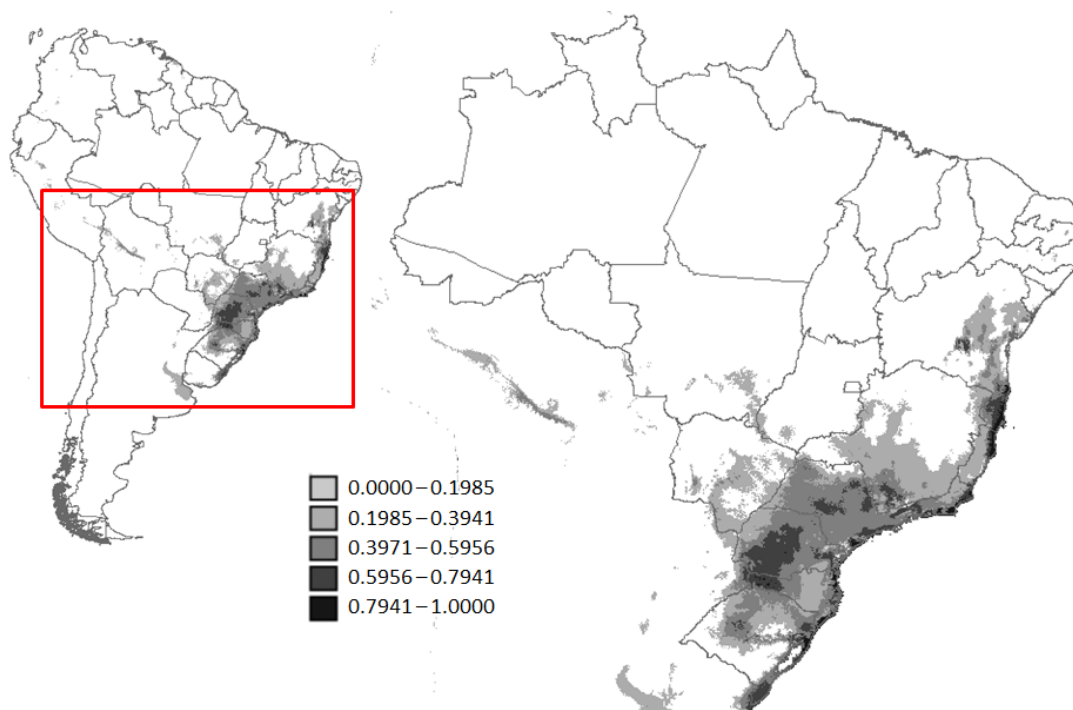
uma predominância geográfica na região sul do Brasil, com pouca probabilidade de ocorrência na região nordeste do Brasil.

- *Projeções CCSM e MIROC*

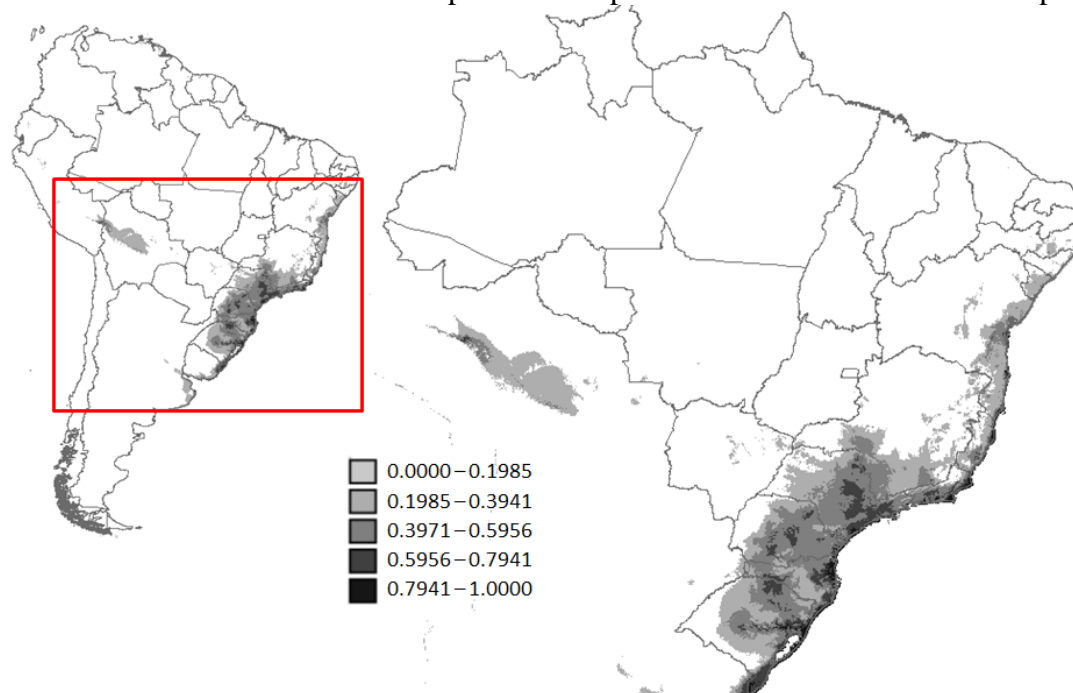
Os mapas de predição para os modelos do passado CCSM e MIROC estão representados nas Figuras 15 e 16, respectivamente. As duas variáveis climáticas que mais contribuíram para a elaboração dos modelos preditivos foram: altitude (porcentagem de contribuição de 29,5 e 24,4 respectivamente) e *Temperature Seasonality* (porcentagem de contribuição de 20,6 e 23,1 respectivamente). Estas projeções sugerem alta probabilidade de ocorrência das populações de *D. meridionalis* no interior dos estados de Paraná, Santa Catarina o Rio Grande do Sul.



**Figura 14.** Modelo de probabilidade de presença atual da espécie *Drosophila meridionalis*, gerada pelo programa MaxEnt. Os diferentes tons de cinza representam a probabilidade da ocorrência da espécie.



**Figura 15.** Modelo de probabilidade de presença da espécie *Drosophila meridionalis* para o último período glacial (21.000 anos atrás), gerada pelo programa MaxEnt, utilizando as projeções do CCSM. Os diferentes tons de cinza representam a probabilidade da ocorrência da espécie.



**Figura 16.** Modelo de probabilidade de presença da espécie *Drosophila meridionalis* para o último período glacial (21.000 anos atrás), gerada pelo programa MaxEnt, utilizando as projeções do MIROC. Os diferentes tons de cinza representam a probabilidade da ocorrência da espécie.

## V. DISCUSSÃO

As análises de estruturação populacional e filogeográficas da espécie *Drosophila meridionalis* a partir das sequências do gene mitocondrial *COI* e o gene *period*, forneceram informações importantes sobre sua história demográfica e a distribuição espacial e temporal de sua variabilidade genética, as quais podem ser comparadas com as espécies cactófilas do cluster *D. buzzatii* e interpretadas em relação às mudanças paleoclimáticas do Quaternário.

### 1.0 Gene mitocondrial *COI*: variabilidade e análises filogeográficas para *Drosophila meridionalis*

A variabilidade genética total (Tabela 2) calculada para o gene mitocondrial *COI* de populações de *D. meridionalis* foi maior que as descritas nas espécies *D. borborema*, *D. seriema*, *D. gouveai*, *D. serido* (FRANCO; MANFRIN, 2013) e *D. buzzatii* (BRITO; MANFRIN; SENE, 2002; SANTOS, 2011). O que é congruente com a elevada variação cariotípica encontrada entre as populações desta espécie (BAIMAI; SENE; PEREIRA, 1983).

#### - *NCPA e variabilidade genética para as redes*

As análises filogeográficas dividem as populações desta espécie nos três agrupamentos distintos redes A, B e C (Figuras 4, 5 e 6) e sugerem eventos de expansão populacional e de fluxo gênico restrito com isolamento por distância, em diferentes clados das redes.

Padrões particulares de distribuição dos haplótipos são esperados sob esses dois eventos demográficos, sendo eles:

- *Fluxo gênico restrito com isolamento por distância*: a dispersão de um haplótipo no espaço ocorre por movimentos geográficos pequenos a cada geração, resultando em uma grande correlação entre quão disperso for um haplótipo (ou clado) e sua posição temporal na árvore de haplótipos. Assim, novos haplótipos podem apenas ser encontrados em sua área de origem, enquanto os haplótipos mais antigos podem ocorrer geograficamente mais dispersos (CASTELLOE; TEMPLETON, 1994; TEMPLETON, 2006). Esse evento demográfico também pode ser evidenciado nas redes, que apresentam haplótipos delimitados espacialmente e que diferem por várias mutações. Em outras palavras esses *gaps* genéticos, podem evidenciar linhagens alopátricas numa árvore de genes. Esta característica supõe também a possibilidade de

uma subestrutura espacial entre as populações e a descontinuidade filogeográfica pode ser explicada por barreiras extrínsecas (história geológica) ao intercâmbio genético (AVISE, 2000).

-*Expansão populacional*: haplótipos encontrados na população ancestral, que sofreu expansão, são geograficamente dispersos. Isto algumas vezes incluirá haplótipos relativamente jovens ou clados que são globalmente raros e frequentemente restritos apenas à área ancestral (TEMPLETON, 2004a, 2006). Esse evento demográfico também pode ser evidenciado nas redes que apresentam forma de estrela (*Star Shape*), apresentando um ou poucos haplótipos com frequências elevadas e muitos haplótipos diferindo por uma ou poucas mutações (AVISE, 2000).

Os critérios anteriormente descritos, junto como os diferentes resultados da NCPA, *Mismatch distribution*, os testes de neutralidade e a variabilidade genética, foram aplicados para realizar as diferentes inferências sobre a história demográfica dos três agrupamentos observados para as populações de *D. meridionalis*.

### 1.1 Rede A

A variabilidade genética intrapopulacional da rede A é caracterizada por valores altos nas populações Itu-SP (N77) e Foz do Jordão-PR (SEG) e valores baixos na população Altinópolis-SP (N26\_N84). Esta última se encontra em um campo rochoso sobre o topo de uma colina isolada, que tem sido considerada como um relicto originado pelas alterações climáticas durante o Quaternário (COSTA; KUHN; SENE, 2000), sendo a causa mais provável para explicar a baixa variabilidade genética calculada.

Foram sugeridos pela NCPA eventos de fluxo gênico restrito com isolamento por distância para os clados 2-5 e 4-1, sendo verificado e confirmado com o teste de Mantel para o último clado (Tabela 7). Eventos de expansão populacional foram sugeridos para os clados 1-11 e 3-3, sendo verificados também por vários testes de neutralidade e/ou a *Mismatch Distribution* (Tabela 8). A discordância entre os diferentes testes para esses clados é devida provavelmente ao tamanho da amostra, que influencia os parâmetros utilizados para testar os modelos demográficos.

Com base nesses resultados e os critérios anteriormente descritos podemos fazer as seguintes inferências para a rede A:

- As populações do interior de São Paulo (clado 2-5) mostraram restrição do fluxo gênico por isolamento por distância, no entanto para esse clado o teste de Mantel não foi significativo. Assim, a restrição do fluxo gênico poderia ser explicada por eventos históricos extrínsecos ao intercâmbio genético, sendo que estas populações apresentam altos e significativos valores de  $F_{ST}$  (Tabela 6), apesar de sua proximidade geográfica, especialmente a população de Altinópolis-SP (N26\_N84) com relação às outras.

- Posteriormente, as populações do interior de São Paulo apresentaram expansão da área de ocorrência (clado 3-3) em direção ao sul, pelo interior. Provavelmente, após a expansão, as populações do interior de Santa Catarina e Rio Grande do Sul sofreram restrição do fluxo gênico por distância. A análise sugere a existência de haplótipos intermediários não amostrados entre as populações do interior de São Paulo e o interior de Paraná e Rio Grande do Sul, de modo que a possibilidade de eventos vicariantes não pode ser descartada.

- Uma característica importante desta rede é a presença de haplótipos exclusivos das populações do litoral de Santa Catarina: Laguna (N79\_N87), Florianópolis (N59\_N61) e Governador Celso-Ramos (N09) e do litoral de São Paulo: Bertioga-SP (N02) e São Sebastião-SP (N29). Dadas as diferentes posições na genealogia destes haplótipos, podemos inferir migrações assíncronas de indivíduos destas populações, em direção ao interior. Sendo provavelmente mais recente o fluxo dos indivíduos das populações do litoral de Santa Catarina.

-Períodos de isolamento geográfico e vários contatos secundários posteriores e assíncronos, podem explicar a presença de haplótipos exclusivos de populações do litoral na rede A.

## 1.2 Rede B

A variabilidade genética intrapopulacional da rede B é caracterizada por valores baixos (Tabela 2) e apresenta a menor distribuição geográfica em relação às outras redes.

Foram sugeridos eventos de expansão populacional para o clado 3-1 (rede total) pela NCPA e confirmado por todos os testes de neutralidade e a *Mismatch Distribution* (Tabela 8).

Dados os resultados encontrados para rede B podemos inferir que:

- Inicialmente a expansão da área de ocorrência de *D. meridionalis* se inicia com as populações do litoral do Rio Grande do Sul, em direção ao norte pelo litoral. A presença de

haplótipos exclusivos e recentes nas populações adjacentes de Florianópolis-SC (N61) e Governador Celso Ramos-SC (N09), evidenciam uma relação histórica entre as redes B e C. Considerando a proximidade destas populações e o pouco compartilhamento de haplótipos também podemos sugerir que possíveis eventos vicariantes nesta região, são em parte responsáveis pela quebra genética das redes B e C.

Esses eventos vicariantes podem estar associados com os afloramentos de rochas cristalinas pré-cambrianas que interrompem a continuidade da planície costeira do litoral sul de Santa Catarina (BAPTISTA DA CUNHA; TEXEIRA GUERRA, 2006).

### 1.3 Rede C

A variabilidade genética intrapopulacional da rede C é caracterizada por valores baixos e altos (Tabela 2) e sua distribuição geográfica é a maior em relação às outras redes.

Foram sugeridos pela NCPA eventos de fluxo gênico restrito com isolamento por distância para o clado 1-1, não confirmado pelo teste de Mantel, e eventos de expansão populacional para o clado 2-1 (rede total), corroborado por vários testes de neutralidade e a *Mismatch Distribution* (Tabela 8).

Dados os resultados encontrados para rede C podemos inferir que:

- As populações do litoral de Santa Catarina apresentaram expansão da área de ocorrência em direção ao norte pelo litoral até o interior da Bahia.

- As populações litorâneas de Santa Catarina e as populações Itaúnas-ES (N31), Mucuri-BA (N32) e Mucugê-BA (N45), que compõem o clado 1-1, se encontram separadas por dois passos mutacionais das populações de Bertioxa-SP (N02), São Sebastião-SP (N29), Arraial do Cabo-RJ (N20) e Macaé-RJ (N22), a pesar de que estas populações estejam na linha de distribuição. Esta característica pode ser um indicativo de fragmentação geográfica, embora esse evento não tenha sido detectado pela NCPA.

- Em termos gerais, a expansão da área de ocorrência da rede C se iniciou com as populações litorâneas de Santa Catarina, seguido de restrição do fluxo gênico devido a: (a) isolamento por distância entre as populações litorâneas de Santa Catarina e as populações litorâneas do norte do Rio de Janeiro e da Bahia; e (b) barreiras biogeográficas entre as populações litorâneas de Santa Catarina e as populações litorâneas de São Paulo e sul do Rio de Janeiro.



Esta quebra genética e geográfica foi observada em estudos cariotípicos de *D. meridionalis*, que mostram dois cariótipos diferentes nas populações de Angra dos Reis (Tipo I) e Arraial do Cabo (Tipo II) na costa do Rio de Janeiro (BAIMAI; SENE; PEREIRA, 1983). Diferenças encontradas pelas análises de parâmetros da corte sonora, entre a população de Bertiooga-SP e as outras populações analisadas por Costa *et al.* (2000), são outras evidências da dinâmica complexa desta região.

Em populações de *D. serido*, também foi observado essa quebra (MORALES, 2005), reforçando a hipótese de existência de barreiras biogeográficas que dificultam o fluxo gênico entre populações dessa região e aquelas das regiões adjacentes. Esse padrão evidencia uma conexão histórica entre populações do litoral de Santa Catarina e o litoral de São Paulo e Rio de Janeiro, sem populações intermediárias entre estas duas regiões, tanto para *D. meridionalis* como para *D. serido*.

Para *D. meridionalis* foi corroborada esta conexão histórica pelas análises filogeográficas no gene *period*. Essa distribuição poderia ser um resultado de eventos de expansão de populações de *Drosophila* acompanhando seus cactos hospedeiros cuja distribuição, hipoteticamente, foi mais ampla no passado, formando o Arco Pleistocênico (TAYLOR; ZAPPI, 2004).

Outros estudos em diferentes táxons têm documentado eventos vicariantes nesta região, por exemplo, na ave *Conopophaga lineata* foram detectados eventos vicariantes na região compreendida pelo vale do Rio Paraíba, que abrange parte do leste de São Paulo e sul do Rio de Janeiro, e fragmentação histórica entre as linhagens dos estados de São Paulo e Paraná, e Rio Grande do Sul até o núcleo de Misiones. Os autores afirmam que mudanças paleoclimáticas e neotectonismo associado à região do vale do Rio Paraíba são os possíveis eventos históricos que influenciaram na distribuição da diversidade desta espécie (OLIVEIRA, 2007).

Baseados na riqueza de espécies de *Opiliones* da Floresta Atlântica foram sugeridos eventos vicariantes nas regiões ao sul do Espírito Santo, no centro e sul de São Paulo e o sul do Rio de Janeiro, sendo relacionados principalmente com o surgimento da Serra do Mar o aparecimento de vales nesta região (PINTO-DA-ROCHA; BERNARDINO DA SILVA; BRAGAGNOLO, 2005).

A principal característica desta região é a proximidade da encosta da Serra do Mar que, em muitos pontos, chega diretamente até o oceano. Estudos geomorfológicos mostram que esta

região tem sido afetada pelo surgimento da Serra do Mar e o aparecimento de vales, como vale do Rio Paraíba do Sul. Uma notável expansão dos climas secos na costa e nas depressões interplanálticas no Rio de Janeiro, tem sido documentada, no entanto esta expansão não tem sido capaz de eliminar todos os contínuos florestais herdados das serras e maciços costeiros. Isto significa uma redução descontínua de florestas em redutos, os quais preservaram a biota florestal que mais tarde se ampliaria e se emendaria vindo a constituir a cobertura vegetal da Serra do Mar e dos morros dos maciços costeiros de Carioca e de Niterói (BAPTISTA DA CUNHA; TEXEIRA GUERRA, 2006).

Estas características fazem que esta região tenha uma dinâmica geomorfológica e biogeográfica complexa, evidenciada em diferentes espécies de *Drosophilas* e outros táxons.

## **2.0 Gene nuclear *period*: variabilidade e análises filogeográficas para *Drosophila meridionalis***

A variabilidade genética (Tabela 9) do gene nuclear *period*, calculada para as populações de *D. meridionalis* foram maiores que as descritas nas espécies do cluster *D. buzzatii*, com exceção de *D. koepferae* (FRANCO et al., 2010).

A variabilidade genética intrapopulacional, destaca-se pelos altos índices das populações (J23) no Rio Grande do Sul, Itirapina (N75) no interior de São Paulo e Arraial do Cabo (N20) no litoral de Rio de Janeiro, sendo o mais alto na população Arroio Teixeira (J23).

A baixa variabilidade encontrada nas populações de São Francisco do Sul (N24) e Penha (N78\_N86), em Santa Catarina, pode ser explicada por um pequeno número de fundadores dessas populações durante a expansão da área de ocorrência da espécie. Além disso, considerando que estas duas populações estão no extremo da distribuição em Santa Catarina, que marca uma descontinuidade das populações pelo litoral, a ocorrência de efeito de gargalo devido a alterações ambientais locais com consequente redução no tamanho populacional, não pode ser descartada.

As populações de Mucuri-BA (N32) e Cantagalo-PR (CAN) são localidades em que o número de indivíduos analisados foi menor que três, de modo que a baixa diversidade nucleotídica apresentada poderia ser decorrente do tamanho amostral.

As populações do litoral apresentam um alto compartilhamento de haplótipos, destacando-se a alta dispersão do Hap\_5, que é distribuído em 12 das 15 populações (Figura12). As

populações do interior de São Paulo e Paraná compartilham haplótipos, embora não compartilhem haplótipos com as populações do litoral.

Para o gene *period* as populações foram divididas em dois grupos, sendo o primeiro congruente com a rede A do gene *COI* e o segundo compreende as redes B e C do gene *COI*.

Por meio da NCPA, eventos de expansão populacional e fluxo gênico restrito com isolamento por distância foram inferidos para o gene *period*. As inferências filogeográficas, baseadas nos resultados das diferentes análises foram:

- Para o clado 2-2, que inclui as populações do interior, foi detectado fluxo gênico restrito com isolamento por distância, embora os valores de  $F_{ST}$  não mostrassem uma diferenciação genética significativa das populações e uma correlação significativa com as distâncias geográficas entre a maioria das populações. No entanto a população de Altinópolis-SP (N84\_N26) apresentou valores de  $F_{ST}$  altos e significativos, com relação a todas as populações do clado (Tabela 13). Para o gene *COI* esta população mostrou baixa variabilidade genética e  $F_{ST}$  altos e significativos em relação às outras populações do interior. Provavelmente esta população ficou geneticamente isolada, dado que esta se encontra em um campo rochoso sobre o topo de uma colina, o que levou posteriormente a uma perda de sua variabilidade e estruturação genética pela interrupção do fluxo gênico por barreiras geográficas.

O padrão deste clado é mais consistente com um fluxo gênico no passado conectando as populações de Paraná e São Paulo, seguido da restrição do fluxo gênico por eventos históricos (barreiras geográficas ou descontinuidade de habitat) mais que pela distância geográfica.

De fato, as projeções do passado para o último período glacial (Figura 15 e 16) sugerem uma alta probabilidade de ocorrência nestas regiões, conectando as atuais populações de São Paulo com as do Paraná, explicando o compartilhamento de haplótipos pelo fluxo gênico contínuo no passado e a restrição por eventos históricos, sendo que a projeção atual da distribuição da espécie (Figura 14) evidencia pouca probabilidade de ocorrência no interior dos estados de Paraná, Santa Catarina e Rio Grande do Sul.

- Eventos demográficos de expansão populacional foram detectados para o clado 3-3, que inclui todas as populações do litoral (mas não todos os indivíduos das populações), desde o Rio Grande do Sul até a Bahia. Provavelmente, a expansão da área de ocorrência se inicia nas

populações do Rio Grande do Sul, seguindo em direção ao norte pelo litoral; esta inferência foi feita considerando a relação dos haplótipos ou clados de interior, supostamente mais antigos, e os haplótipos ou clados de extremidade na genealogia, teoricamente mais recentes (TEMPLETON; BOERWINKLE; SING, 1987). Outras evidências desta inferência são: a ocorrência de populações de *D. meridionalis* em diversas localidades da Argentina e no Paraguai (COSTA; KUHN; SENE, 2000; VILELA, 1983), a pouca ou inexistente ocorrência da espécie no nordeste do Brasil (SILVA-BERNARDI, 2009), as projeções de nicho ecológico e a alta diversidade haplotípica das populações do Rio Grande do Sul.

De acordo com a posição dos haplótipos ou clados das populações do interior na genealogia (haplótipos na extremidade) podemos inferir que estes são mais recentes, havendo colonizado o interior pela ocorrência de migrações de indivíduos de populações do litoral de Santa Catarina e/ou São Paulo.

A rede total de haplótipos não apresentou associações significativas pela NCPA, no entanto, resultados de diferentes análises mostraram expansão populacional (Tabela 14) e uma correlação positiva e significativa entre as distâncias geográfica e genética (Tabela 12). Dado esses resultados, poderia inferir-se que as populações desta espécie foram afetadas por diferentes eventos assíncronicos como: expansão populacional das populações do litoral sul, em direção ao norte, com posterior diferenciação populacional e fragmentação da distribuição geográfica gerada por: (a) isolamento por distância nas populações de litoral e (b) processos vicariantes nas populações do interior, possivelmente à dinâmica geomorfológica e biogeográfica complexa da região do interior e litoral de São Paulo e Santa Catarina, associada aos eventos paleoclimáticos ocorridos durante o Pleistoceno.

### ***Discordância entre os genes***

Grande parte das análises filogeográficas nos dois genes é concordante, tanto nos agrupamentos dos haplótipos como nos eventos sugeridos para esses agrupamentos. A principal discordância é em relação às populações do litoral. O gene mitocondrial *COI* gera duas redes diferentes para essas populações, separando as populações do litoral sul (rede B) e o litoral sudeste e oriental (rede C). No entanto, o gene *period* mostra um compartilhamento de haplótipos destas populações.

Geralmente, essa discordância entre os genes pode ser explicada pelo tempo de coalescência, aos padrões de diferenciação populacional e a forma de como estes marcadores reconstróem as relações entre os grupos (TEMPLETON, 2006; TOEWS; BRELSFORD, 2012). Esse tipo de discordância é esperado dado a diferença no nível da ploidia, que tem um grande impacto no tempo de coalescência. Sendo o DNA mitocondrial haploide e herdado maternalmente, o tamanho populacional efetivo e tempo de coalescência serão menores do que um gene ligado ao X (TEMPLETON, 2006; ZINK; BARROWCLOUGH, 2008). Assim os marcadores nucleares serão mais lentos em indicar processos de estruturação genética de que comparado com o DNA mitocondrial, que dadas suas propriedades, reflete com mais precisão divergência recente.

A discordância entre os padrões de variação genética no DNA mitocondrial e nos genes nucleares, também pode indicar que as populações passaram por períodos de isolamento geográfico com contato secundário posterior (TOEWS; BRELSFORD, 2012).

Assim, o gene mitocondrial *COI* reflete a estruturação das populações do litoral, sendo que este é um indicador mais sensível da estrutura populacional do que os genes nucleares. Por outro lado, os genes nucleares mostraram importantes estimativas das relações profundas da árvore de haplótipos, taxas de crescimento populacional e valores do fluxo gênico.

### 3.0 Hipótese filogeográfica para *Drosophila meridionalis*

Considerando os resultados obtidos pelas diferentes análises realizados para os dois genes, podemos inferir que a expansão da área de ocorrência de *D. meridionalis* teve início com as populações do litoral do Rio Grande de Sul, em direção ao litoral de Santa Catarina com posterior colonização a longa distância dos estados de São Paulo, Rio de Janeiro e Bahia (clado 3-3 do gene *period*-rede B do gene *COI*).

O gene *COI* que separa as populações do litoral em dois agrupamentos (rede B e C) e sendo que esse apresenta tempo de coalescência menor, podemos inferir que este evento é mais recente que a estruturação das populações do interior e o litoral. O que é confirmado pela árvore coalescente *MCCT* (Figura 10).

Outros eventos marcantes na história demográfica e filogeográfica desta espécie, são as migrações assíncronas de indivíduos de populações litorâneas de São Paulo e Santa Catarina em direção ao interior, provavelmente no estado de São Paulo.

As populações do interior e do litoral mostram uma estruturação genética alta e sem compartilhamento de haplótipos no gene *period*, embora, o gene mitocondrial *COI* mostre haplótipos exclusivos de populações do litoral na rede A, evidenciando diferentes migrações de indivíduos daquelas populações em direção ao interior. Estabelecidas as populações no interior, estas ficaram isoladas das populações do litoral, provavelmente por barreiras geográficas, como por exemplo, a Serra do Mar.

Uma segunda expansão populacional foi inferida a partir das populações do interior de São Paulo em direção ao sul pelo interior, colonizando Paraná e Rio Grande do Sul, com um posterior fluxo restrito por isolamento por distância (clado 2-2 do gene *period*-rede A do gene *COI*). No entanto, a possibilidade da influência de outros eventos na interrupção deste fluxo gênico entre as populações do interior, não pode ser descartada. As análises de modelagem de nicho ecológico mostram que o interior do Rio Grande do Sul, Santa Catarina, Paraná e São Paulo tivessem uma conexão no passado, levantando hipóteses de que barreiras ecológicas e/ou barreiras históricas deixaram essas populações isoladas.

A seguir, a hipótese filogeográfica proposta neste trabalho para *D. meridionalis* será comparada com as histórias filogeográficas descritas para as espécies do cluster *D. buzzatii*, levando em consideração o compartilhamento de aspectos ecológicos e geográficos entre elas por convergência, pois são espécies cactófilas.

A espécie *D. antonietae* ocorre no interior dos estados de São Paulo, Paraná e Rio Grande do Sul, extremo leste do Chaco Argentino, e litoral do Rio Grande do Sul e Santa Catarina (SILVA; SENE, 1991), sendo, em maior parte a mesma distribuição de *D. meridionalis*. A partir de sequências do gene mitocondrial *COI* de *D. antonietae* encontram evidências de expansão populacional no passado seguido de uma colonização de áreas costeiras do Atlântico a partir de áreas do núcleo de Misiones das FTSS (DE BRITO; MANFRIN; SENE, 2002), tal expansão pode ter sido favorecida pela presença de cacto por toda a depressão ‘riograndense’ central no estado do Rio Grande do Sul, estendendo a linha da costa do Atlântico. No litoral, o evento de expansão detectado está na direção norte, seguindo as populações de cacto por todas as dunas do estado do Rio Grande do Sul e de Santa Catarina. Manfrin e Sene, (2006) sugerem que esta expansão foi recente, devido às evidências geomorfológicas, que indicam que a origem das planícies costeiras foi próximo ao período Cenozóico; afetada por regressões marinhas e transgressões dos últimos 10 mil anos.

No interior, a hipótese filogeográfica para *D. antonietae* sugere a expansão da área de ocorrência a partir do norte do estado de São Paulo, com posterior colonização dos estados do Paraná, Rio Grande do Sul e Santa Catarina pelo interior (DE BRITO; MANFRIN; SENE, 2002).

A história evolutiva da espécie *D. antonietae* coincide em grande parte com a hipótese filogeográfica sugerida para *D. meridionalis*. No entanto, a principal diferença é que para *D. antonietae* foi sugerido que a região dos estados do Rio Grande do Sul e Santa Catarina, foi colonizada a partir da expansão da área de ocorrência de populações do interior (DE BRITO; MANFRIN; SENE, 2002). Neste trabalho, as análises filogeográficas mostram que é mais provável uma expansão da área de ocorrência, inicialmente das populações da região sul em direção ao norte, com posteriores migrações assíncronicas de indivíduos de populações do litoral (Santa Catarina e São Paulo) em direção ao interior e posteriormente uma expansão da área de ocorrência do interior de São Paulo em direção ao sul pelo interior.

Esta inferência é concordante com uma possível rota de dispersão da espécie *D. buzzatii*, que levanta a hipótese de dispersão desde o noroeste da Floresta Atlântica para o Cerrado principalmente em São Paulo e o norte do Paraná, onde a divisão entre os dois biomas é mais sutil. Esta rota pode ter ocorrido em períodos de condições climáticas favoráveis para a expansão das FTSS adaptadas a climas secos, invadindo o bioma da Floresta Atlântica e possibilitando a formação de um corredor de dispersão ao longo do estado de São Paulo (SANTOS, 2011), não só facilitando a dispersão de *D. buzzatii*, mas também de *D. meridionalis*.

No entanto, a inclusão de populações do interior da Argentina e outras do interior Rio Grande do Sul, serão importantes para validar essa hipótese.

### ***Estruturação de Drosophila meridionalis***

Para o gene nuclear *period* a principal informação apresentada pela AMOVA, são os valores  $\Phi_{ct}$  maiores entre os grupos formados pelas populações do interior e o litoral ( $\Phi_{ct}=0,71898$ ), do que nos grupos formados pelas três redes ( $\Phi_{ct}=0,64954$ ). Para o gene mitocondrial *COI*, contrariamente os valores entre os grupos formados pelas populações do interior e do litoral ( $\Phi_{ct}=0,70350$ ) foram menores do que nos grupos formados pelas três redes ( $\Phi_{ct}=0,88761$ ). No entanto, é importante notar que os valores de  $\Phi_{ct}$  do grupo formado pelas populações do interior e o litoral, são muito altos e similares nos dois genes.

Isto indica e confirma a alta estruturação das populações do litoral e interior, sendo provavelmente os processos associados à estruturação destas populações mais antigos que os processos associados à estruturação das populações do litoral.

Silva-Bernardi (2010) gerou um dendrograma utilizando as distâncias de Mahalanobis calculadas pela análise discriminante das asas das mesmas populações avaliadas neste trabalho, encontrando que as populações de *D. meridionalis* estão divididas em dois grupos: interior e litoral.

#### 4.0 Análises Bayesianas

Árvore de tempo pela máxima credibilidade do clado (*Maximum Clade Credibility Tree-MCCT*) indica que o tempo do ancestral comum mais recente ( $T_{MRCA}$ ) para todos os haplótipos de *D. meridionalis* é de 81.700 anos atrás. Essa data marca a separação das populações do interior e do litoral aproximadamente no final do Pleistoceno, o que explicaria a forte estruturação genética encontrada para esses dois agrupamentos populacionais pelas AMOVAs.

O  $T_{MRCA}$  para os haplótipos das redes B e C foi estimado em 25.900 anos atrás, dentro deste nó a árvore *MCCT* mostra que  $T_{MRCA}$  foi de 6.900 para a rede C e 11.300 para a rede B, mostrando que os haplótipos da rede C são mais recentes. Para a rede A foi estimada um  $T_{MRCA}$  de 16.300 anos atrás.

Em árvores de coalescência é possível detectar fatores históricos demográficos que podem influenciar a genealogia dos genes, gerando padrões específicos. Por exemplo, em uma população em crescimento ao longo do tempo, a probabilidade de um evento de coalescência, esta menos próximo ao presente, porque a população se encontra em seu maior tamanho. Assim, genealogias de populações crescentes tendem a ter tempos mais longos entre eventos coalescentes em direção para ao presente, e menor tempo entre os eventos coalescentes no passado. Portanto, estas genealogias tendem a ter ramos mais longos em direção aos ramos do presente e de menor profundidade na árvore. Em uma população que está diminuindo (*shrinking*) de tamanho ao longo do tempo, a probabilidade de um evento coalescente é mais próximo ao presente, porque a população se encontra em seu menor tamanho. Portanto, estas genealogias tendem a ter ramos mais curtos em direção aos ramos do presente e de maior profundidade na árvore (AVISE, 2000; HAMILTON, 2009).



Dada a topologia da árvore para o gene mitocondrial *COI*, podemos inferir que as populações de *D. meridionalis* estão passando por uma diminuição do tamanho populacional, sendo concordante com a BSP (Figura 9).

Além disso, a árvore confirma a hipótese de que a divergência das populações do interior e o litoral é mais antiga, que a divergência entre as populações do litoral.

### ***História demográfica baseada em Análises Bayesianas***

A história demográfica das populações de *D. meridionalis* para o gene mitocondrial *COI*, sugere uma redução no tamanho populacional próximo ao tempo presente, fato concordante com a atual redução da distribuição geográfica das FTSS, predominantes entre 42.000 e 8.500 anos atrás (BEHLING et al., 2000), e com a diminuição da ocorrência da espécie em localidades anteriormente amostradas (SILVA-BERNARDI, 2010).

Algumas espécies do cluster *D. buzzatii* apresentam histórias demográficas diferentes e  $T_{MRCA}$  mais antigas que *D. meridionalis*. As análises da BSP para *D. serido*, *D. gouveai* e *D. borborema* exibiram um aumento exponencial do tamanho das populações entre 45 e 75 mil anos atrás, que coincide com o período glacial de Wisconsin (FRANCO; MANFRIN, 2013) e para a *D. buzzatii* mostrou um crescimento populacional em torno de 550.000 a 650.000 mil anos atrás (SANTOS, 2011).

As análises da BSP para a espécie *D. seriema* mostrou um período de equilíbrio da população e um gargalo que coincide aproximadamente com o último período glacial, em que houve um decréscimo populacional significativo, entre 30 mil anos atrás até o presente, em que o tamanho populacional dessa espécie foi reduzido cerca de 50%. Para *D. antonietae* é mostrado um período de equilíbrio demográfico com uma fase de crescimento populacional entre 100-15 mil anos, seguido de uma tendência para uma diminuição da população nos últimos 15 mil anos, no entanto com baixo suporte estatístico (FRANCO; MANFRIN, 2013). As histórias demográficas destas duas espécies são concordantes com a história demográfica de *D. meridionalis*, sendo que mostram uma diminuição da população em um intervalo de 40 mil anos atrás até o presente.

*D. meridionalis* tem sido encontrada em simpatria com a espécie *D. antonietae* em localidades do litoral de Rio Grande do Sul e Santa Catarina e o interior de São Paulo, Paraná e

Rio Grande do Sul, compartilhando a maior parte da distribuição geográfica (DE BRITO; MANFRIN; SENE, 2002; MANFRIN; SENE, 2006).

Além disso, *D. meridionalis* também é encontrada em simpatria com *D. seriema* nas localidades de Mucugê-BA (N45) e Serra do Cipó-BA, embora só tenha sido amostrado em Mucugê um único indivíduo de *D. meridionalis*, juntamente com 30 indivíduos do “cluster” *D. buzzatii*, sendo cinco indivíduos de *D. seriema* (FRANCO, 2009). Porém a ocorrência de *D. meridionalis* na Serra do Cipó-MG, Seabra-BA e Milagres-BA foi registrada em coletas anteriores (BAIMAI; SENE; PEREIRA, 1983; TIDON-SKLORZ et al., 1994), os indivíduos dessas populações não foram amostrados nas coletas de campo realizadas durante o trabalho de Silva-Bernardi (2010) e não estão incluídas neste trabalho.

Dadas às adaptações, condições ambientais e distribuições geográficas similares das espécies cactófilas *D. seriema*, *D. antonietae* e *D. meridionalis*, espera-se que estas espécies compartilhem uma história demográfica evolutiva.

A história demográfica influencia a capacidade de adaptação local das populações, de modo que as adaptações dependem da história demográfica subjacente à determinada população, pois movimentos históricos influenciam a quantidade e distribuição da variação genética das populações, e esta por sua vez influencia a resposta das populações às pressões ambientais.

Análises filogeográficas do *cluster D. buzzatii*, tem demonstrado que a diversificação, dinâmica espacial e demográfica destas espécies, tem sido influenciada pelas mudanças climáticas do Quaternário (BRITO; MANFRIN; SENE, 2002; FRANCO; MANFRIN, 2013; MANFRIN; SENE, 2006; MORAES et al., 2009; SANTOS, 2011). Durante essas oscilações paleoclimáticas, a vegetação xerofítica teve expansão e contração de sua distribuição (AB´SABER, 1977; PENNINGTON et al., 2000) e dada à associação destas espécies com os cactos, espera-se que os resultados da expansão e retração das distribuições das populações destas *Drosophila* cactófilas estejam relacionados a esses eventos.

## 5.0 Modelagem de Nicho Ecológico

Os resultados da Modelagem de Nicho Ecológico apresentam, em sua maioria, resultados condizentes com a área de ocorrência conhecida de *D. meridionalis*.

A projeção atual mostra uma predominância geográfica na região sul do Brasil, com pouca probabilidade de ocorrência na região nordeste. Este padrão de distribuição foi corroborado pelas

---

diferentes coletas realizadas por Silva-Bernardi (2010) nos estados de Pernambuco, no interior da Bahia e Minas Gerais, sendo que não foram encontrados indivíduos de *D. meridionalis*. Estas informações contribuem a reforçar a hipóteses de que *D. meridionalis* iniciara sua expansão da área de ocorrência a partir das populações do Rio Grande do Sul em direção norte pelo litoral, ate chegar à Bahia, onde sua ocorrência é fortemente diminuída.

As projeções do passado CCSM e MIROC sugerem alta probabilidade de ocorrência das populações de *D. meridionalis* no interior dos estados do Paraná, Santa Catarina o Rio Grande do Sul, conectando as atuais populações de interior do norte de São Paulo com as populações do interior de Paraná e o interior do Rio Grande do Sul, o que explicaria o compartilhamento de haplótipos entre estas populações. Igualmente, estas projeções mostram uma maior probabilidade de ocorrência no estado da Bahia, mostrando que no passado esta espécie poderia ter colonizado essa região e posteriormente sofrido uma retração da distribuição.

## VI CONCLUSÕES

A partir dos resultados obtidos neste trabalho podemos concluir que:

- As análises filogeográficas do gene mitocondrial *COI* e gene *period* foram concordantes nas inferências da história evolutiva para as populações da espécie *D. meridionalis*; com exceção da subdivisão do litoral mostrada pelo gene *COI*, que pode ser uma evidência da divergência recente destas populações;
- O centro de dispersão dos eventos de expansão populacional detectados para *D. meridionalis*, provavelmente iniciaram com as populações do litoral do Rio Grande do Sul;
- As populações de *D. meridionalis* apresentam elevada estrutura populacional, sendo esta associada a eventos de expansão populacional, seguidos pela fragmentação (barreiras geográficas) em dois principais grupos populacionais: interior e litoral;
- As análises bayesianas para o gene *COI* confirmaram que a divergência das populações do interior e o litoral é mais antiga, que a divergência entre as populações do litoral;
- A história demográfica das populações de *D. meridionalis* para o gene mitocondrial *COI*, sugeriu uma redução no tamanho populacional próximo ao tempo presente;
- A espécie *D. antonietae* e *D. meridionalis* compartilham grande parte de sua história evolutiva;
- *D. meridionalis*, como as espécies do cluster *D. buzzatii*, apresentou indicativos de flutuações demográficas, podendo estar associadas à expansão e contração da distribuição da vegetação xerofítica, durante as oscilações paleoclimáticas do Pleistoceno.

## VII. REFERÊNCIAS BIBLIOGRÁFICAS

AB´SABER, A. N. Os domínios morfoclimáticos na América do Sul. Primeira aproximação. **Geomorfologia**, v. 52, p. 1-22, 1977.

AB´SABER, A. N. Os Domínios de Natureza no Brasil: potencialidades paisagísticas. São Paulo. Atelie Editorial. 159 pp. 2003.

AGARWAL, M.; SHRIVASTAVA, Æ. N.; Á, S. Á. R. Á. P. Advances in molecular marker techniques and their applications in plant sciences. **Plant Cell Rep**, v. 27, p. 617-631, 2008.

AULER, A. S. et al. Quaternary ecological and geomorphic changes associated with rainfall events in presently semi-arid northeastern Brazil. **Journal of Quaternary Science**, v. 19, n. 7, p. 693-701, 27 out. 2004.

AVISE, J. C. et al. Intraspecific Phylogeography: The mitochondrial DNA bridge between population Genetics and systematics. **Annual Review of Ecology and Systematics**, v. 18, p. 489-522, 1987.

AVISE, J. C. **The History and Formation of Species**. Cambridge, Massachusetts: Harvard University Press, 2000.

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

BAIMAI, V. Heterochromatin Accumulation and Karyotypic Evolution in Some Dipteran Insects. **Zoological Studies**, v. 37, n. 2, p. 75-88, 1998.

BAIMAI, V.; SENE, F. M.; PEREIRA, M. A. Q. R. Heterochromatin and karyotypic differentiation of some neotropical cactus-breeding species of the *Drosophila repleta* species group. **Genetica**, v. 60, p. 81-92, 1983.

BAPTISTA DA CUNHA, S.; TEXEIRA GUERRA, A. J. **Geomorfologia do Brasil**. Rio de Janeiro: Editoa Bertrand Brasil LTDA, 2006. p. 329

BARR, N. B.; CUI, L. Molecular Systematics of Nuclear Gene period in Genus *Anastrepha* (Tephritidae). **Annals of the Entomological Society of America**, v. 98, n. 2, 2005.

BARR, N. B.; WIEGMANN, B. M. Phylogenetic relationships of *Ceratitis* fruit flies inferred from nuclear CAD and tango/ARNT gene fragments: testing monophyly of the subgenera *Ceratitis* (*Ceratitis*) and *C.* (*Pterandrus*). **Molecular Phylogenetics and Evolution**, v. 53, n. 2, p. 412-24, nov. 2009.

- BAUZER, L. G. S. R. et al. The period gene and genetic differentiation between three Brazilian populations of *Lutzomyia longipalpis*. **Insect Molecular Biology**, v. 11, n. 4, p. 315-23, ago. 2002.
- BEAUMONT, M. A; RANNALA, B. The Bayesian revolution in genetics. **Nature Reviews. Genetics**, v. 5, n. 4, p. 251-61, abr. 2004.
- BEAUMONT, M. A. et al. In defence of model-based inference in phylogeography. **Molecular Ecology**, v. 19, n. 3, p. 436-446, fev. 2010.
- BEAVER, L. M. et al. Loss of circadian clock function decreases reproductive fitness in males of *Drosophila melanogaster*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 99, n. 4, p. 2134-9, 19 fev. 2002.
- BEAVER, L. M. et al. Noncircadian regulation and function of clock genes period and timeless in oogenesis of *Drosophila melanogaster*. **Journal of Biological Rhythms**, v. 18, n. 6, p. 463-72, dez. 2003.
- BEAVER, L. M.; GIEBULTOWICZ, J. M. Regulation of Copulation Duration by period and timeless in *Drosophila melanogaster*. **Current Biology**, v. 14, p. 1492-1497, 2004.
- BEHLING, H. et al. Late Quaternary vegetational and climate dynamics in northeastern Brazil , inferences from marine core GeoB 3104-1. **Quaternary Science Reviews**, v. 19, p. 981-994, 2000.
- BERTORELLE, G.; EXCOFFIER, L. Inferring admixture proportions from molecular data. **Molecular Biology and Evolution**, v. 15, n. 10, p. 1298-311, out. 1998.
- BRITO, R. A. DE; MANFRIN, M. H.; SENE, F. M. Mitochondrial DNA phylogeography of Brazilian populations of *Drosophila buzzatii*. **Genetics and Molecular Biology**, v. 25, n. 2, p. 161-171, 2002.
- CABANNE, G. S. et al. Nuclear and mitochondrial phylogeography of the Atlantic forest endemic *Xiphorhynchus fuscus* (Aves: Dendrocolaptidae): biogeography and systematics implications. **Molecular Phylogenetics and Evolution**, v. 49, n. 3, p. 760-73, dez. 2008.
- CAMARGO, A.; HEYER, W. R.; DE SÁ, R. O. Phylogeography of the frog *Leptodactylus validus* (Amphibia: Anura): patterns and timing of colonization events in the Lesser Antilles. **Molecular Phylogenetics and Evolution**, v. 53, n. 2, p. 571-9, nov. 2009.
- CARNAVAL, A. C.; BATES, J. M. Amphibian DNA shows marked genetic structure and tracks pleistocene climate change in northeastern Brazil. **Evolution; International Journal of Organic Evolution**, v. 61, n. 12, p. 2942-57, dez. 2007.
- CARNAVAL, A. C.; HICKERSON, M. J.; HADDAD, C. F. B. Stability Predicts Genetic Diversity in the Brazilian Atlantic Forest Hotspot. **Science**, v. 323, n. 6, p. 785-789, 2009.

CARNAVAL, A. C.; MORITZ, C. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. **Journal of Biogeography**, v. 35, n. 7, p. 1187-1201, jul. 2008.

CASTELLOE, J.; TEMPLETON, A. R. Root Probabilities for Intraspecific Gene Trees under Neutral Coalescent Theory. **Molecular Phylogenetics and Evolution**, v. 3, n. 2, p. 102-113, 1994.

CHANG, D. C.; REPERT, S. M. A Novel C-Terminal Domain of Drosophila PERIOD Inhibits dCLOCK : CYCLE-Mediated Transcription. v. 13, n. Ccid, p. 758-762, 2003.

CITRI, Y. et al. A family of unusually spliced biologically active transcripts encoded by a Drosophila clock gene. **Nature**, v. 326, n. 6108, p. 42-47, 1987.

CLEMENT, M.; POSADA, D.; CRANDALL, K. A. TCS: a computer program to estimate gene genealogies. **Molecular Ecology**, v. 9, n. 10, p. 1657-9, out. 2000.

COLOT, H. V; HALL, J. C.; ROSBASH, M. Interspecific comparison of the period gene of Drosophila reveals large blocks of non-conserved coding DNA. **EMBO**, v. 7, n. 12, p. 3929-3937, 1988.

COSTA, C. T. A.; KUHN, G. C.; SENE, F. M. LOW COURTSHIP SONG VARIATION IN SOUTH AND Drosophila meridionalis ( DIPTERA , DROSOPHILIDAE ). **Revista Brasileira de Biologia**, v. 60, n. 1, p. 53-61, 2000.

CRANDALL, K. A; TEMPLETON, A R. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. **Genetics**, v. 134, n. 3, p. 959-69, jul. 1993.

DANTAS MENDONÇA, G. P.; CABANNE, G. S.; SANTOS RODRIGUES, F. How Past Vicariant Events Can Explain the Atlantic Forest Biodiversity? In: **Ecosystems Biodiversity**. [S.l: s.n.]. p. 430-445.

DE BRITO, R. A; MANFRIN, M. H.; SENE, F. M. Nested cladistic analysis of Brazilian populations of Drosophila serido. **Molecular Phylogenetics and Evolution**, v. 22, n. 1, p. 131-43, jan. 2002.

DE OLIVEIRA, P. E.; BARRETO, F. A. M.; SUGUIO, K. Late Pleistocene/Holocene climatic and vegetational history of the Brazilian caatinga : the fossil dunes of the middle São Francisco River. **Palaeogeography, Palaeoclimatology, Palaeoecology**, v. 152, p. 319-337, 1999.

DRUMMOND, A J. et al. Bayesian coalescent inference of past population dynamics from molecular sequences. **Molecular Biology and Evolution**, v. 22, n. 5, p. 1185-92, maio. 2005.

DRUMMOND, A. J.; RAMBAUT, A. BEAST: Bayesian evolutionary analysis by sampling trees. **BMC Evolutionary Biology**, v. 7, p. 214, jan. 2007.

DURANDO, C. M. et al. Phylogenetic analysis of the repleta species group of the genus *Drosophila* using multiple sources of characters. **Molecular Phylogenetics and Evolution**, v. 16, n. 2, p. 296-307, ago. 2000.

D'HORTA, F. M. et al. The genetic effects of Late Quaternary climatic changes over a tropical latitudinal gradient: diversification of an Atlantic Forest passerine. **Molecular Ecology**, v. 20, n. 9, p. 1923-35, maio. 2011.

EPPERSON, B. K. **Geographical Genetics**. Princeton, NJ, USA: Princeton University Press, 2003.

EXCOFFIER, L.; HECKEL, G. Computer programs for population genetics data analysis: a survival guide. **Nature Reviews. Genetics**, v. 7, n. 10, p. 745-58, out. 2006.

EXCOFFIER, L.; LISCHER, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. **Molecular Ecology Resources**, v. 10, n. 3, p. 564-567, 2010.

EXCOFFIER, L.; SMOUSE, P. E.; QUATTRO, J. M. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. **Genetics Society of America**, v. 491, p. 479-491, 1992.

FRANCO, F. F. **História evolutiva do “cluster” *Drosophila buzzatii* (grupo *D. repleta*): eventos históricos e diversificação de espécies no Brasil**. [S.l.] Universidade de São Paulo - Ribeirão Preto, 2009.

FRANCO, F. F. et al. Intra- and interspecific divergence in the nuclear sequences of the clock gene period in species of the *Drosophila buzzatii* cluster. **Journal of Zoological Systematics and Evolutionary Research**, v. 48, n. 4, p. 322-331, 10 nov. 2010.

FRANCO, F. F.; MANFRIN, M. H. Recent demographic history of cactophilic *Drosophila* species can be related to Quaternary palaeoclimatic changes in South America. **Journal of Biogeography**, p. 1-13, 2013.

FU, Y.-X. Statistical Tests of Neutrality of Mutations Against Population Growth, Hitchhiking and Background Selection. **Genetics Society of America**, 1997.

GARRICK, R. C. et al. Babies and bathwater: a comment on the premature obituary for nested clade phylogeographical analysis. **Molecular Ecology**, v. 17, n. 6, p. 1401-1403, mar. 2008.

GLEASON, J. M. Mutations and natural genetic variation in the courtship song of *Drosophila*. **Behavior genetics**, v. 35, n. 3, p. 265-77, maio. 2005.

GLEASON, J. M.; GRIFFITH, E. C.; POWELL, J. R. A Molecular Phylogeny of the *Drosophila willistoni* Group: Conflicts Between Species Concepts? **Evolution**, v. 52, n. 4, p. 1093-1103, 1 ago. 1998.



- GOTO, S. G.; KIMURA, M. T. Phylogenetic utility of mitochondrial COI and nuclear Gpdh genes in *Drosophila*. **Molecular Phylogenetics and Evolution**, v. 18, n. 3, p. 404-22, mar. 2001.
- GRAZZIOTIN, F. G. et al. Phylogeography of the *Bothrops jararaca* complex (Serpentes: Viperidae): past fragmentation and island colonization in the Brazilian Atlantic Forest. **Molecular Ecology**, v. 15, n. 13, p. 3969-82, nov. 2006.
- HAFFER, J. Speciation in Amazonian Forest Birds. **Science**, v. 165, n. 3889, p. 131-136, 1969.
- HAFFER, J.; PRANCE, T. Climatic forcing of evolution in Amazonia during the Cenozoic: on the refuge theory of biotic differentiation. **Amazoniana**, v. 16, p. 579-605, 2001.
- HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41, p. 95-98, 1999.
- HAMILTON, M. B. **Population Genetics**. Hoboken: WILEY-BLACKWELL, 2009. p. 422
- HEBERT, P. D. N.; RATNASINGHAM, S.; DEWAARD, J. R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. **Proceedings. Biological sciences / The Royal Society**, v. 270 Suppl, p. S96-S99, 7 ago. 2003.
- HEWITT, G. M. Quaternary phylogeography: the roots of hybrid zones. **Genetica**, v. 139, n. 5, p. 617-38, maio. 2011.
- HEY, J.; MACHADO, C. A. The study of structured populations--new hope for a difficult and divided science. **Nature Reviews. Genetics**, v. 4, n. 7, p. 535-43, jul. 2003.
- HICKERSON, M. J. et al. Phylogeography's past, present, and future: 10 years after Avise, 2000. **Molecular Phylogenetics and Evolution**, v. 54, n. 1, p. 291-301, jan. 2010.
- HUECK, K. **As Florestas da América do Sul**. São Paulo. Brazil: [s.n.].
- JANZEN, D. H. et al. Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. **Molecular Ecology Resources**, v. 9, p. 1-26, maio. 2009.
- KAWAHARA, A. Y. et al. Phylogeny and biogeography of hawkmoths (Lepidoptera: Sphingidae): evidence from five nuclear genes. **PloS one**, v. 4, n. 5, p. e5719, jan. 2009.
- KINGMAN, J. F. C. The coalescent. **Stochastic Processes and their Applications**, v. 13, n. 3, p. 235-248, set. 1982.
- KNOWLES, L. L. Why does a method that fails continue to be used? **Evolution; international Journal of Organic Evolution**, v. 62, n. 11, p. 2713-7, nov. 2008.
- KNOWLES, L. L. Statistical Phylogeography. **Annual Review of Ecology , Evolution and Systematics**, v. 40, p. 593-612, 2009.

- KNOWLES, L.; MADDISON, W. Statistical phylogeography. **Molecular Ecology**, p. 2623-2635, 2002.
- KONOPKA, R. J.; KYRIACOU, C. P.; HALL, J. C. Mosaic Analysis in the *Drosophila* Cns of Circadian and Courtship-Song Rhythms Affected by a Period Clock Mutation: Short Communication. **Journal of Neurogenetics**, v. 11, n. 1-2, p. 117-139, 1 jan. 1996.
- KOSTÁL, V.; ZÁVODSKÁ, R.; DENLINGER, D. Clock genes period and timeless are rhythmically expressed in brains of newly hatched, photosensitive larvae of the fly, *Sarcophaga crassipalpis*. **Journal of Insect Physiology**, v. 55, n. 5, p. 408-14, maio. 2009.
- KYRIACOU, C. P.; HALL, J. C. Circadian rhythm. **Proceedings of the National Academy of Sciences of the United States of America**, v. 77, n. 11, p. 6729-6733, 1980.
- LAGISZ, M. et al. Two distinct genomic regions, harbouring the period and fruitless genes, affect male courtship song in *Drosophila montana*. **Heredity**, v. 108, n. 6, p. 602-8, jun. 2012.
- LEDRU, M.-P. Late Quaternary Environmental and Climatic Changes in Central Brazil. **Quaternary Research**, v. 39, p. 90-98, 1993.
- LEDRU, M.-P. et al. Regional assessment of the impact of climatic change on the distribution of a tropical conifer in the lowlands of South America. **Diversity and Distributions**, v. 13, n. 6, p. 761-771, 26 jun. 2007.
- LEWIS, R. L.; BECKENBACH, A. T.; MOOERS, A. Ø. The phylogeny of the subgroups within the melanogaster species group: likelihood tests on COI and COII sequences and a Bayesian estimate of phylogeny. **Molecular Phylogenetics and Evolution**, v. 37, n. 1, p. 15-24, out. 2005.
- LINARES, M. C. et al. High mitochondrial diversity in geographically widespread butterflies of Madagascar: a test of the DNA barcoding approach. **Molecular Phylogenetics and Evolution**, v. 50, n. 3, p. 485-95, mar. 2009.
- LINS, R. M. M. A et al. Fixed differences in the paralytic gene define two lineages within the *Lutzomyia longipalpis* complex producing different types of courtship songs. **PloS one**, v. 7, n. 9, p. e44323, jan. 2012.
- MANFRIN, A. M. H.; BRITO, R. O. A. DE; SENE, F. M. Systematics and Evolution of the *Drosophila buzzatii* (Diptera : Drosophilidae) Cluster Using mtDNA Systematics and Evolution of the *Drosophila buzzatii* (Diptera : Drosophilidae) Cluster Using mtDNA. **Annals of the Entomological Society of America**, v. 94, n. 3, p. 333-346, 2001.
- MANFRIN, M. H.; SENE, F. M. Cactophilic *Drosophila* in South America: a model for evolutionary studies. **Genetica**, v. 126, n. 1-2, p. 57-75, jan. 2006.
- MANTEL, N. The detection of disease clustering and a generalized regression approach. **Cancer Research**, v. 27, n. 2, p. 209-20, fev. 1967.

- MATZKIN, L. M. The molecular basis of host adaptation in cactophilic *Drosophila*: molecular evolution of a glutathione S-transferase gene (*GstD1*) in *Drosophila mojavensis*. **Genetics**, v. 178, n. 2, p. 1073-83, fev. 2008.
- MAZZONI, C. J. et al. Molecular polymorphism, differentiation and introgression in the period gene between *Lutzomyia intermedia* and *Lutzomyia whitmani*. **BMC Evolutionary Biology**, v. 6, p. 85, jan. 2006.
- MAZZOTTA, G. M. et al. The clock gene period in the medfly *Ceratitis capitata*. **Genetical Research**, v. 86, n. 1, p. 13-30, ago. 2005.
- MIYATAKE, T. et al. The period gene and allochronic reproductive isolation in *Bactrocera cucurbitae*. **Proceedings. Biological sciences / The Royal Society**, v. 269, n. 1508, p. 2467-72, 7 dez. 2002.
- MORAES, E. M. et al. Phylogeography of the cactophilic species *Drosophila gouveai*: demographic events and divergence timing in dry vegetation enclaves in eastern Brazil. **Journal of Biogeography**, v. 36, n. 11, p. 2136-2147, 2009.
- MORALES, A. C. **Análise histórico-evolutiva de populações das espécies cactofílicas *Drosophila serido* e *Drosophila antonietae* (Diptera, Drosophilidae)**. [S.l.] Universidade de São Paulo - Ribeirão Preto, 2005.
- NAGARAJA, N. J.; RANGANATH, H. A. Molecular phylogeny of the nasuta subgroup of *Drosophila* based on 12S rRNA, 16S rRNA and CoI mitochondrial genes, RAPD and ISSR polymorphisms. **Genes & Genetic Systems**, v. 79, n. 5, p. 293-9, out. 2004.
- NEI, M. **Molecular Evolutionary Genetics**. Columbia U ed. New York: [s.n.].
- NEWBOLD, T. et al. Climate-based models of spatial patterns of species richness in Egypt's butterfly and mammal fauna. **Journal of Biogeography**, v. 36, n. 11, p. 2085-2095, nov. 2009.
- NIELSEN, J. et al. Big flies, small repeats: the "Thr-Gly" region of the period gene in Diptera. **Molecular Biology and Evolution**, v. 11, n. 6, p. 839-53, nov. 1994.
- NIELSEN, R.; BEAUMONT, M. A. Statistical inferences in phylogeography. **Molecular Ecology**, v. 18, n. 6, p. 1034-1047, mar. 2009.
- OLIVEIRA, D. C. S. G. et al. Molecular systematics and geographical distribution of the *Drosophila longicornis* species complex (Diptera: Drosophilidae). **Zootaxa**, v. 1069, p. 1-32, 2005.
- OLIVEIRA, P. R. **Sistemática e Biogeografia Histórica da Família Conophagidae (Aves: Passeriformes): Especiação nas Florestas da América do Sul**. [S.l.] Universidade de São Paulo, 2007.

OLIVEIRA-FILHO, A. T.; FONTES, M. A. Patterns of Floristic Differentiation among Atlantic Forests in Southeastern Brazil and the Influence of Climate. **Biotropica**, v. 32, n. 4b, p. 793-810, 2000.

PANCHAL, M.; BEAUMONT, M. A. The automation and evaluation of nested clade phylogeographic analysis. **Evolution; International Journal of Organic Evolution**, v. 61, n. 6, p. 1466-80, jun. 2007.

PAVAN, A. N. A. C. et al. Patterns of diversification in two species of short-tailed bats ( *Carollia* Gray , 1838 ): the effects of historical fragmentation of Brazilian rainforests. **Biological Journal of the Linnean Society**, v. 102, p. 527-539, 2011.

PENNINGTON, R. T. et al. Neotropical seasonally dry forests and Quaternary vegetation changes. **Journal of Biogeography**, v. 27, p. 261-273, 2000.

PEREIRA, M. A. Q. R.; VILELA, C. R.; SENE, F. M. Notes on breeding and feeding sites of some species of the repleta group of the genus *Drosophila* (Diptera, Drosophilidae). **Ciencia e Cultura**, v. 35, p. 1313-1319, 1983.

PFEILER, E.; MARKOW, T. A. Phylogeography of the Cactophilic *Drosophila* and Other Arthropods Associated with Cactus Necroses in the Sonoran Desert. **Insects**, v. 2, n. 2, p. 218-231, 5 maio. 2011.

PHILLIPS, C. D. et al. Phylogeography of Steller sea lions: relationships among climate change, effective population size, and genetic diversity. **Journal of Mammalogy**, v. 92, n. 5, p. 1091-1104, 14 out. 2011.

PHILLIPS, S. J.; ANDERSON, R. P.; SCHAPIRE, R. E. Maximum entropy modeling of species geographic distributions. **Ecological Modelling**, v. 190, n. 3-4, p. 231-259, jan. 2006.

PICCIN, A. et al. The Clock Gene period of the Housefly, *usca domestica*, Rescues Behavioral Rhythmicity in *Drosophila melanogaster* : Evidence for Intermolecular Coevolution ? **Genetics**, v. 154, p. 747-758, 2000.

PINTO-DA-ROCHA, R.; BERNARDINO DA SILVA, M.; BRAGAGNOLO, C. Faunistic similarity and historic biogeography of the harvestmen of southern and southeastern atlantic rain forest of brazil faunistic similarity and historic biogeography of the harvestmen of southern and southeastern. **Journal of Arachnology**, v. 33, n. 2, p. 290-299, 2005.

POSADA, D.; CRANDALL, K. Intraspecific gene genealogies: trees grafting into networks. **Trends in Ecology & Evolution**, v. 16, n. 1, p. 37-45, 1 jan. 2001.

POSADA, D.; CRANDALL, K. A; TEMPLETON, A R. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. **Molecular Ecology**, v. 9, n. 4, p. 487-8, abr. 2000.

PRADO, D. E.; GIBBS, P. E. Patterns of species distributions in the dry seasonal forests of south America. **Annals of the Missouri Botanical Garden**, v. 80, n. 4, p. 902-927, 1993.

QUIJADA-MASCAREÑAS, A. J. et al. Phylogeographic patterns of trans-Amazonian vicariants and Amazonian biogeography: the Neotropical rattlesnake (*Crotalus durissus* complex) as an example. **Journal of Biogeography**, v. 34, n. 8, p. 1296-1312, 18 ago. 2007.

RAFAEL, V.; ARCOS, G. Subgrupo inca, un nuevo subgrupo del grupo repleta, con descripción de *Drosophila huancavilcae* nov. sp. (Diptera, Drosophilidae). **Evolución Biológica**, v. 3, n. 3, p. 233-243, 1989.

RAMBAUT, A.; DRUMMOND, A. J. **Tracer [Computer Program]**, 2003. Disponível em: <<http://beastbioedacuk/tracer>>

RAMOS-ONSINS, S. E.; ROZAS, J. Statistical properties of new neutrality tests against population growth. **Molecular Biology and Evolution**, v. 19, n. 12, p. 2092-100, dez. 2002.

RAMÍREZ-SORIANO, A. et al. Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. **Genetics**, v. 179, n. 1, p. 555-67, maio. 2008.

REECE, J. S. et al. Phylogeography of two moray eels indicates high dispersal throughout the indo-pacific. **The Journal of Heredity**, v. 101, n. 4, p. 391-402, 2010.

REGIER, J. C. et al. Evolution and phylogenetic utility of the period gene in Lepidoptera. **Molecular Biology and Evolution**, v. 15, n. 9, p. 1172-82, set. 1998.

REGIER, J. C.; COOK, C. P. et al. A phylogenetic study of the “bombycoid complex” (Lepidoptera) using five protein-coding nuclear genes, with comments on the problem of macrolepidopteran phylogeny. **Systematic Entomology**, v. 33, n. 1, p. 175-189, 14 jan. 2008.

REGIER, J. C.; GRANT, M. C. et al. Phylogenetic relationships of wild silkmoths (Lepidoptera: Saturniidae) inferred from four protein-coding nuclear genes. **Systematic Entomology**, v. 33, n. 2, p. 219-228, 4 abr. 2008.

REPPERT, S. M. et al. Cloning of a structural and functional homolog of the circadian clock gene period from the giant silkmoth *Antheraea pernyi*. **Neuron**, v. 13, n. 5, p. 1167-76, nov. 1994.

RIBEIRO, A. C. Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergent continental margin. **Neotropical Ichthyology**, v. 4, n. 2, p. 225-246, 2006.

RICCOMINI, C.; ASSUMPÇÃO, M. Quaternary tectonics in Brazil. **Episodes**, v. 22, n. 1997, p. 221-225, 1999.

- ROGERS, A R.; HARPENDING, H. Population growth makes waves in the distribution of pairwise genetic differences. **Molecular Biology and Evolution**, v. 9, n. 3, p. 552-69, maio. 1992.
- ROZAS, J. et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods. **Bioinformatics**, v. 19, n. 18, p. 2496-2497, 10 dez. 2003.
- RUIZ, A et al. The Drosophila serido speciation puzzle: putting new pieces together. **Genetica**, v. 108, n. 3, p. 217-27, jan. 2000.
- RUSSO, C. A; TAKEZAKI, N.; NEI, M. Molecular phylogeny and divergence times of drosophilid species. **Molecular Biology and Evolution**, v. 12, n. 3, p. 391-404, maio. 1995.
- SAKAI, T.; ISHIDA, N. Circadian rhythms of female mating activity governed by clock genes in Drosophila. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 16, p. 9221-5, 31 jul. 2001.
- SANTOS, H. M. **As rotas de dispersão de Drosophila buzzatii na América do Sul**. [S.l.] Universidade de São Paulo - Ribeirão Preto, 2011.
- SAWYER, L. A et al. The period gene Thr-Gly polymorphism in Australian and African Drosophila melanogaster populations: implications for selection. **Genetics**, v. 174, n. 1, p. 465-80, set. 2006.
- SCHNEIDER, S.; EXCOFFIER, L. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. **Genetics**, v. 152, n. 3, p. 1079-89, jul. 1999.
- SENE, F. M.; PEREIRA, M. A. Q. R.; VILELA, C. R. Evolutionary aspects of cactus breeding Drosophila in South America. In: BARKER, B. J. S. F.; STARMER, W. T. (Eds.). **Ecological Genetics and Evolution. The Cactus—Yeast—Drosophila Model System**. Sydney: Academic Press, 1982. p. 97-106.
- SENE, F. M.; PEREIRA, M. A. Q. R.; VILELA, C. R. Contrasting patterns of differentiation inferred from traditional genetic markers in the process of speciation. **Pacific Science**, v. 42, p. 81-88, 1988.
- SHEARMAN, L. P. et al. Two period homologs: circadian expression and photic regulation in the suprachiasmatic nuclei. **Neuron**, v. 19, n. 6, p. 1261-9, dez. 1997.
- SIEDCHLAG, A. C. et al. Genetic structure, phylogeny, and biogeography of Brazilian eyelidless lizards of genera Calyptommatus and Nothobachia (Squamata, Gymnophthalmidae) as inferred from mitochondrial DNA sequences. **Molecular Phylogenetics and Evolution**, v. 56, n. 2, p. 622-30, ago. 2010.

- SILVA, A. F. .; SENE, F. M. Morphological geographic variability in *Drosophila serido* (Diptera, Drosophilidae). **Revista Brasileira de Entomologia**, v. 35, n. 2, p. 455-468, 1991.
- SILVA-BERNARDI, É. C. C. **Estrutura populacional e filogeografia de *Drosophila meridionalis* (Diptera, Drosophilidae)**. [S.l.] Universidade São Paulo - RP, 2009.
- SIMON, C. et al. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. **Annals of the Entomological Society of America**, v. 87, p. 651-701, 1994.
- SMITH, M. A. et al. Wolbachia and DNA barcoding insects: patterns, potential, and problems. **PloS One**, v. 7, n. 5, p. e36514, jan. 2012.
- STARMER, W. T. et al. Adaptations of *Drosophila* and Yeasts: their Interactions with the Volatile 2-propanol in the Cactus-Micro organism-*Drosophila* Model System. **Australian Journal of Biological Sciences**, v. 39, p. 69-77, 1986.
- SUN, Z. S. et al. RIGUI, a putative mammalian ortholog of the *Drosophila* period gene. **Cell**, v. 90, n. 6, p. 1003-11, set. 1997.
- TAJIMA, F. EVOLUTIONARY RELATIONSHIP OF DNA SEQUENCES IN FINITE POPULATIONS. **Genetics Society of America**, p. 437-460, 1983.
- TAJIMA, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. **Genetics**, v. 123, n. 3, p. 585-95, nov. 1989.
- TAMURA, K. et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Molecular Biology and Evolution**, v. 28, n. 10, p. 2731-9, out. 2011.
- TAUBER, E. et al. Temporal mating isolation driven by a Behavioral Gene in *Drosophila*. **Current Biology**, v. 13, n. 03, p. 140-145, 2003.
- TAYLOR, N. P.; ZAPPI, D. **Cacti of Eastern Brazil**. Reino Unido: Royal Botany Garden, 2004.
- TEI, H. et al. Circadian oscillation of a mammalian homologue of the *Drosophila* period gene. **Nature**, v. 389, n. 6650, p. 512-6, 2 out. 1997.
- TELLES, P. DE C. M.; DINIZ-FILHO, J. A. F. Multiple Mantel tests and isolation-by- distance, taking into account long-term. **Genetics and Molecular Research**, v. 4, n. 4, p. 742-748, 2005.
- TEMPLETON, A. R. Using Haplotype Trees for Phylogeographic and Species Inference in Fish Populations. **Environmental Biology of Fishes**, v. 69, n. 1-4, p. 7-20, mar. 2004a.
- TEMPLETON, A. R. Statistical phylogeography: methods of evaluating and minimizing inference errors. **Molecular Ecology**, v. 13, n. 4, p. 789-809, abr. 2004b.

TEMPLETON, A. R. **Population Genetics and Microevolutionary Theory**. St. Louis Missouri: John Wiley & Sons., Inc., Wiley-Liss, 2006. p. 716

TEMPLETON, A. R. Nested clade analysis: an extensively validated method for strong phylogeographic inference. **Molecular Ecology**, v. 17, p. 1877-1880, 2008.

TEMPLETON, A. R. Statistical hypothesis testing in intraspecific phylogeography: nested clade phylogeographical analysis vs. approximate Bayesian computation. **Molecular Ecology**, v. 18, n. 2, p. 319-31, jan. 2009.

TEMPLETON, A. R.; BOERWINKLE, E.; SING, C. F. A Cladistic Analysis of Phenotypic Associations With Haplotypes Inferred From Restriction Endonuclease Mapping. I. Basic Theory and an Analysis of Alcohol Dehydrogenase Activity in *Drosophila*. **Genetics Society of America**, 1987.

TEMPLETON, A. R.; CRANDALL, K. A.; SING, F. Cladistic Analysis of Phenotypic Associations With Haplotypes Inferred From Restriction Endonuclease Mapping and DNA Sequence Data. 111. Cladogram Estimation. **Genetics Society of America**, v. 132, p. 619-633, 1992.

TEMPLETON, A. R.; ROUTMAN, E.; PHILLIPS, C. A. Separating Population Structure from Population History: A cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. **Genetics Society of America**, v. 140, n. 767-782, 1995.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic acids research**, v. 22, n. 22, p. 4673-80, 11 nov. 1994.

THOMÉ, M. T. C. et al. Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. **Molecular phylogenetics and evolution**, v. 55, n. 3, p. 1018-31, jun. 2010.

TIDON-SKLORZ, R. et al. The genus *Drosophila* (Diptera, Drosophilidae) in the Serra do Cipó, state of Minas Gerais, Brazil. **Revista Brasileira de Entomologia**, v. 38, n. 4, p. 627-637, 1994.

TOBBACK, J. et al. The circadian clock genes affect reproductive capacity in the desert locust *Schistocerca gregaria*. **Insect biochemistry and molecular biology**, v. 41, n. 5, p. 313-21, maio. 2011.

TOEWS, D. P. L.; BRELSFORD, A. The biogeography of mitochondrial and nuclear discordance in animals. **Molecular Ecology**, v. 21, n. 16, p. 3907-30, ago. 2012.

VILELA, C. R. **A revision of the *Drosophila repleta* species group (Diptera, Drosophilida)**, 1983.



- WANG, X. et al. Wet periods in northeastern Brazil over the past 210 kyr linked to distant climate anomalies. **Nature**, v. 432, n. December, p. 2767-2769, 2004.
- WARMAN, G. R. et al. Analysis of the circadian clock gene period in the sheep blow fly *Lucilia cuprina*. **Genetical Research Cambridge**, v. 75, p. 257-267, 2000.
- WASSERMAN, M. Cytological studies of the repleta group of the genus *Drosophila*. V. The mulleri subgroup. **The University of Texas Publications**, v. 6205, p. 85-117, 1962.
- WASSERMAN, M. Evolution of the repleta group. In: ASHBURNER, A.; CARSON, H. L.; THOMPSON, J. N. (Eds.). **The Genetics and Biology of Drosophila**. [S.l: s.n.]. p. 61-139.
- WASSERMAN, M. Cytological Evolution of the *Drosophila* repleta species group. In: KRIMBAS, J. R.; R, P. J. (Eds.). **Drosophila Inversion Polymorphism**. CRC Press ed. Boca Raton: [s.n.]. p. 455-555.
- YANG, Y. et al. Increasing the data size to accurately reconstruct the phylogenetic relationships between nine subgroups of the *Drosophila melanogaster* species group (Drosophilidae, Diptera). **Molecular Phylogenetics and Evolution**, v. 62, n. 1, p. 214-23, jan. 2012.
- YU, Q. et al. Behaviour modification by in vitro mutagenesis of a variable region with the period gene of *Drosophila*. **Nature**, v. 326, n. 23, 1987.
- ZINK, R. M.; BARROWCLOUGH, G. F. Mitochondrial DNA under siege in avian phylogeography. **Molecular ecology**, v. 17, n. 9, p. 2107-21, maio. 2008.