Ana Beatriz Zichinelli

Linhagens evolutivas na espécie taxonômica *Drosophila serido* (grupo *D. repleta*)

Ribeirão Preto – SP

2019

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

Linhagens evolutivas na espécie taxonômica *Drosophila serido* (grupo *D. repleta*)

Aluna: Ana Beatriz Zichinelli

Orientadora: Prof^a. Dra. Maura Helena Manfrin

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto da USP, como parte das exigências para a obtenção do título de Mestre em Ciências. Área: Genética.

Versão corrigida. A versão original encontra-se disponível tanto na Biblioteca da Unidade que aloja o Programa, quanto na Biblioteca Digital de Teses e Dissertações da USP (BDTD).

Ribeirão Preto – SP

2019

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

FICHA CATALOGRÁFICA

Zichinelli, Ana Beatriz

Linhagens evolutivas na espécie taxonômica *Drosophila serido* (grupo *D. repleta*)

Orientadora: Profª Dra. Maura Helena Manfrin

78 p.; 19 il.

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto/USP.
Departamento de Genética. Área: Genética e Evolução.

1. Delimitação de espécies. 2. Marcadores moleculares. 3. Novas linhagens.

Ana Beatriz Zichinelli

Título da dissertação: Linhagens evolutivas na espécie taxonômica *Drosophila serido*
(grupo *D. repleta*)

Dissertação apresentada à Faculdade de
Medicina de Ribeirão Preto da USP,
como parte das exigências para a
obtenção do título de Mestre em
Ciências. Área: Genética.

Aprovado em:

Banca examinadora:

Prof.(a) Dr. (a) _____

Instituição: _____ Assinatura _____

Prof.(a) Dr. (a) _____

Instituição: _____ Assinatura _____

Prof.(a) Dr. (a) _____

Instituição: _____ Assinatura _____

Prof.(a) Dr. (a) _____

Instituição: _____ Assinatura _____

AGRADECIMENTOS

- Á Profa. Dra. Maura Helena Manfrin, primeiramente por ter aceitado a me orientar, e por toda dedicação, apoio, paciência, incentivo, confiança e amizade durante esses anos.
- Aos colegas de laboratório: Dora, Adriano e Marcelo, pela amizade e ajuda nos experimentos e nas análises, e por valiosas discussões que aprimoraram o meu conhecimento.
- Ao técnico do laboratório, Paulo Ricardo Epifânio, pela amizade e por toda ajuda em cuidar das moscas e do laboratório, para tudo funcionar perfeitamente.
- Ao analista de sistemas do Departamento de Genética, Pedro Roberto Prado, pelo auxílio e instalação do sistema operacional em nosso servidor.
- Aos meus pais, Paulo e Eliana, pelo dom da vida e por estar comigo em todos os momentos, me apoiando, incentivando, cuidando. A vocês, minha eterna gratidão.
- Á minha irmã, Izadora, pelo companheirismo e amizade de sempre.
- Às secretárias do Programa de Pós Graduação em Genética, por toda a paciência e ajuda em questões burocráticas.
- Aos colegas do Laboratório de Biologia e Desenvolvimento de Abelhas, por inúmeras conversas e explicações sobre análises e/ou algum tema específico, e por todo apoio e amizade.
- Ao Programa de Pós Graduação em Genética, à Faculdade de Medicina de Ribeirão Preto e à Universidade de São Paulo, pela oportunidade da realização deste trabalho.
- Á CAPES pela concessão da bolsa de mestrado.
- Á CAPES, FAPESP, CNPq, FMRP e USP cujas verbas mantém o Laboratório de Genética Evolutiva.

SUMÁRIO

1 INTRODUÇÃO	10
1.1 Conceito e delimitação de espécie	10
1.2 Espécies cactofílicas de <i>Drosophila</i>	11
1.3 A espécie <i>Drosophila serido</i>	12
1.3.1 Populações do litoral	12
1.3.2 Populações do nordeste	13
1.4 Marcadores moleculares	13
1.4.1 Gene mitocondrial COI	14
1.4.2 Genes nucleares autossômicos GstD1 e α esterase-5	15
1.4.3 Gene nuclear kl-5 ligado ao cromossomo Y	15
2 OBJETIVO	17
3 MATERIAL E MÉTODOS	18
3.1 Material biológico	18
3.2 Amplificação, purificação e sequenciamento dos genes nucleares $\alpha E5$, <i>GstD1</i> , <i>kl-5</i> e mitocondrial COI	24
3.3 Análise dos dados	29
3.3.1 Leitura e alinhamento das sequências	29
3.3.2 Estatística descritiva	29
3.3.3 Modelos de substituição nucleotídica	29
3.3.4 Análise de Variância Molecular e Teste de Mantel	30
3.3.5 Reconstrução filogenética	30
3.3.6 Análise da estrutura populacional	31
3.3.7 Análise de componentes principais	31
3.3.8 Redes haplotípicas	31
4 RESULTADOS E DISCUSSÃO	33
4.1 Estatística descritiva e modelo de substituição nucleotídica	33
4.1.1 Gene COI	33

4.1.2 Gene α esterase-5.....	34
4.1.3 Gene GstD1.....	36
4.1.4 Gene kl-5.....	38
4.2 Análise de variância molecular e teste de Mantel.....	39
4.2.1 Gene COI.....	39
4.2.2 Gene α esterase-5.....	40
4.2.3 Gene GstD1.....	41
4.2.4 Gene kl-5.....	42
4.3 Reconstrução filogenética.....	43
4.3.1 Gene COI.....	43
4.3.2 Gene α esterase-5.....	45
4.3.3 Gene GstD1.....	48
4.3.4 Gene kl-5.....	51
4.3.5 Genes analisados em conjunto.....	53
4.4 Análise de estrutura populacional.....	56
4.4.1 Gene COI.....	56
4.4.2 Gene α esterase-5.....	57
4.4.3 Gene GstD1.....	58
4.4.4 Gene kl-5.....	58
4.5 Análise de componentes principais.....	59
4.6 Redes haplotípicas.....	60
4.6.1 Gene COI.....	60
4.6.2 Gene α esterase-5.....	61
4.6.3 Gene GstD1.....	62
4.6.4 Gene kl-5.....	63
4.7 Discussão.....	64
5 CONCLUSÃO.....	68
6 REFERÊNCIAS.....	69

RESUMO

Zichinelli, A. B. **Linhagens evolutivas na espécie taxonômica *Drosophila serido* (grupo *D. repleta*)**. 2019. 78p. Dissertação (Mestrado). Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 2019.

O “cluster” *Drosophila buzzatii* é formado por sete espécies endêmicas da América do Sul e que apresentam relação ecológica obrigatória com cactos. Entre elas, *Drosophila serido* possui uma distribuição geográfica muito ampla. Análises de inversões cromossômicas polimórficas, hidrocarbonetos da cutícula, diversidade haplotípica mitocondrial e análises cariotípicas apontam a existência de dois grupos populacionais de *D. serido*: populações do litoral e populações do nordeste. Neste trabalho, utilizando sequências dos genes nucleares GstD1, kl5 e α esterase-5 e mitocondrial COI, testamos a hipótese de que a espécie taxonômica *D. serido* é composta por mais de uma linhagem evolutiva, por meio de análises de estrutura populacional, agrupamento populacional, rede haplotípica e filogenia. Os resultados para a análise de estrutura populacional apontam dois (ou mais) agrupamentos para a espécie. Para a análise de agrupamento populacional com base na distância genética, analisamos que algumas populações do litoral estão se distanciando do restante das populações do litoral e nordeste, porém ainda temos as duas populações dos dois agrupamentos hipotéticos próximas. Para a análise das redes haplotípicas, os genes que separam os dois agrupamentos são α esterase-5 e GstD1. Para o restante dos marcadores, ainda apontam o compartilhamento de alguns haplótipos entre as populações. Nas análises filogenéticas, os genes GstD1, α esterase-5 e os quatro analisados em conjunto separam as populações em dois agrupamentos, enquanto os marcadores COI e Kl-5 ainda têm as duas populações hipotéticas misturadas, corroborando com as outras análises. Contudo, nossos dados apoiam que existam duas linhagens evolutivas para *D. serido*.

Palavras-chave: delimitação de espécies, marcadores moleculares, novas linhagens.

ABSTRACT

Zichinelli, A. B. **Evolutionary lineages of the taxonomic species *Drosophila serido* (*D. repleta* group)**. 2019. 78p. Dissertação (Mestrado). Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 2019.

The "cluster" *Drosophila buzzatii* is formed by seven endemic species from South America and have a mandatory ecological relationship with cacti. Among them, *Drosophila serido* has a very broad geographic distribution. Analyses of polymorphic chromosomal inversions, cuticle hydrocarbons, haplotypal mitochondrial diversity and karyotypic analyses indicate the existence of two population groups of *D. serido*: coastal populations and northeastern populations. In this study, using sequences of the nuclear genes GstD1, KL5 and α Esterase-5 and mitochondrial COI, we tested the hypothesis that the taxonomic species *D. serido* is composed of more than one evolutionary lineage, through analysis of population structure, population grouping, haplotypal network and phylogeny. The results for the population structure analysis indicate two (or more) groupings for the species. For population cluster analysis based on genetic distance, we analyzed that some coastal populations are moving away from the rest of the coastal and northeast populations, but we still have the two populations of the two hypothetical clusters nearby. For haplotypic network analysis, the genes that separate the two clusters are α esterase-5 and GstD1. For the remaining markers, they still point to the sharing of some haplotypes among the populations. In phylogenetic analyzes, the GstD1, α esterase-5 and the four genes analyzed together separate the populations into two clusters, while the markers COI and KI-5 still have the two hypothetical populations mixed, corroborating the other analyzes. However, our data support that there are two evolutionary lineages for *D. serido*.

Keywords: delimitation of species, molecular markers, new lineages.

1 Introdução

1.1 Conceito e delimitação de espécie

Uma das unidades comparativas básicas em estudos científicos é a espécie. Apesar de sua importância, não há um consenso em como defini-la como unidade ou limitá-la em campo, considerando suas populações. Entretanto, um conceito muito utilizado e discutido, como um paradigma, é o proposto por Mayr (1977) que define espécie como “agrupamentos de populações naturais inter cruzantes, reprodutivamente isolados de outros grupos com as mesmas características”. Considerando esta definição, o início do processo de especiação é o estabelecimento de algum isolamento reprodutivo pré ou pós-zigótico entre populações (MAYR, 1942; RIDLEY, 2004). Com um isolamento reprodutivo incipiente, as populações estarão evoluindo de forma independente, podendo ocorrer a fixação de variantes genéticas diferentes em cada uma delas por seleção ou deriva. O fluxo gênico entre essas populações será interrompido, e cada uma desenvolverá uma propriedade nova que servirá como barreira para o isolamento reprodutivo. (MAYR, 1942; RITCHIE, 2007).

Neste processo, de acordo com Mayr (1977), os indivíduos de uma espécie devem ser considerados em três níveis, sendo:

- comunidade reprodutiva: os indivíduos assumem entre si o papel de parceiros potenciais e reconhecem um ao outro com o intuito de reproduzir;
- unidade ecológica: os indivíduos interagem com outras espécies, pois compartilham o meio ambiente;
- unidade genética: intercomunicação de um grande patrimônio genético, enquanto o indivíduo é uma pequena parte do todo.

Entretanto, vários trabalhos mostram zonas de simpatria entre populações de diferentes espécies com formação de híbridos, sugerindo que os mecanismos de isolamento reprodutivo são incipientes (MAYR, 1977). Porém, nestas áreas as espécies se mantêm como unidades reconhecíveis, sugerindo que, apesar da potencial troca de material genético, mecanismos genômicos e ecológicos podem interferir impedindo a fusão das populações (MAYR, 1977).

O estudo de como fatores históricos, como processos históricos, barreiras ambientais e fluxo gênico podem atingir a dinâmica das populações ainda não é totalmente entendido, mas informações importantes podem ser obtidas a partir de análises genéticas de diferentes marcadores, com dinâmicas diferentes, entre populações de uma mesma espécie.

1.2 Espécies cactofílicas de *Drosophila*

Alguns grupos dentro do gênero *Drosophila* apresentam relação obrigatória ecológica com cactos (família Cactaceae), dispondo de seus tecidos necrosados para seu desenvolvimento larval (BARKER; STARMER, 1982). Esta associação pode ter induzido a diferenciação das espécies, tanto em termos históricos quanto em ecológico-adaptativos (MARKOW; O'GRADY, 2005). Pelo último, a diferenciação populacional pode ser influenciada pela composição química dos cactos e pela expressão de genes relacionados aos processos de desintoxicação de compostos secundários (MATZKIN et al, 2006; MATZKIN, 2008, 2012). Do ponto de vista histórico, acredita-se que eventos subsequentes e recorrentes, mas não sincrônicos, de retração e expansão das formações vegetais abertas (matas úmidas/semiárido-árido), durante paleoeventos climáticos modificaram a distribuição de cactos (BONATELLI et al, 2014) e, assim, atuando nas populações de *Drosophila* (FRANCO; MANFRIN, 2013; BARRIOS-LEAL et al, 2018).

Uma radiação de espécies cactofílicas de *Drosophila* encontrada na América do Sul forma o “cluster” *Drosophila buzzatti*, composto por sete espécies: *Drosophila buzzatti* (que ocorre por toda a área de distribuição do “cluster”); *D. koepferae* (ocorre no domínio do Chaco); *D. antonietae* (ocorre nas bacias dos rios Paraná-Paraguai); *D. borborema*, *D. seriema*, *D. gouveai* e *D. serido* (ocorrem na caatinga, costa atlântica e cerrado) (MANFRIN; SENE, 2006). Esse “cluster” é considerado monofilético, apontado através de análises de genes nucleares (FRANCO et al, 2010), herança citoplasmática (MANFRIN; DE BRITO; SENE, 2001) e marcadores cromossômicos (RUIZ; WASSERMAN, 1993), entretanto as topologias das árvores filogenéticas apontam incongruências, que podem ser devido ao pouco tempo de divergência (FRANCO; MANFRIN, 2013).

1.3 A espécie *Drosophila serido*

A espécie *Drosophila serido* é uma espécie endêmica do Brasil e tem sua distribuição geográfica muito ampla: presente na Caatinga, em populações isoladas nos estados de Minas Gerais (MG) e Goiás (GO) e em enclaves da vegetação xerofítica na costa Atlântica. Foi descrita em 1977 por meio de coletas na localidade de Milagres – BA (VILELA; SENE, 1977). Essas populações são reconhecidas pela inversão fixa $2x^7$ (WASSERMAN; RICHARDSON, 1987) e edeago morfotipo “A” (SILVA; SENE, 1991), apesar de existir variação morfológica ao longo de sua distribuição geográfica (FRANCO et al, 2008). Esta espécie é cactofílica e está relacionada com várias espécies de cactos, por exemplo: *Cephalocereus piauhyensis*, *Cereus hildmannianus*, *Cereus fernambucensis*, *Opuntia ficus-indica*, *Opuntia monacantha* (MANFRIN; SENE, 2006). Análises de inversões cromossômicas polimórficas (RUIZ et al, 2000), hidrocarbonetos da cutícula (OLIVEIRA et al, 2011), diversidade haplotípica mitocondrial (MANFRIN et al, 2001) e análises cariotípicas (BAIMAI et al, 1983) apontam a existência de dois grupos populacionais de *D. serido*: populações do litoral e populações do nordeste.

1.3.1 Populações do litoral

As populações ao longo da costa atlântica mostram complexidade evolutiva, apresentada através da não congruência e politipia observada em diversos marcadores utilizados. As populações dos estados de Santa Catarina e São Paulo apresentam, além da inversão $2x^7$, a inversão cromossômica fixa $2y^9$ e, além dessas, as inversões $2x^8$ e $2w^8$ (RUIZ et al, 2000). No que se refere às placas metafásicas, foram descritos: cromossomos metafásicos tipo III em Arraial do Cabo – RJ, tipo IV em Peruíbe – SP (BAIMAI et al, 1983) e um outro tipo em Florianópolis – SC. A análise da variação morfológica de edeagos em populações de *D. serido* também mostrou a presença de dois grupos distintos na costa atlântica: um grupo formado pelas populações de Bertioga – SP e Penha – SC e o outro grupo com as populações de São Sebastião – SP e Arraial do Cabo – RJ (FRANCO et al, 2008). Análises filogeográficas, com base em sequências parciais do gene mitocondrial COI, mostraram que as populações do litoral

formam um grupo diferente das populações do nordeste, juntando as populações do estado de São Paulo com as populações do litoral sul do estado da Bahia, e as populações do litoral do estado de Santa Catarina com as populações da caatinga (MORALES, 2005; KOKUDAI et al, 2013).

Segundo Lavagnini-Pizzo (2015), análises filogeográficas utilizando genes nuclear e mitocondrial (*period* e COI, respectivamente), concordam com a homogeneidade populacional observada por outros marcadores utilizados. Porém, quanto à estrutura populacional ao longo do litoral atlântico, há divergência entre os marcadores. Para o gene mitocondrial COI, as populações estariam divididas em dois grupamentos, sendo que as populações de Santa Catarina estariam geneticamente isoladas das demais, formando o grupo “Sul”, enquanto para o gene nuclear *period* as populações desde o sul da Bahia até Santa Catarina formam um único agrupamento, denominado “Sudeste”.

1.3.2 Populações do Nordeste

As populações de *D. serido* do Nordeste apresentam as inversões cromossômicas polimórficas $2a^8$, $2b^8$, $2c^8$ e $2d^8$ (TOSI; SENE, 1989) e são homogêneas para a maioria dos marcadores utilizados até o momento (FRANCO et al, 2008; FRANCO; MANFRIN, 2013). Análises filogeográficas realizadas a partir de sequências parciais do gene mitocondrial COI, sugerem que o centro de dispersão desta espécie é o interior do estado da Bahia, em localidades próximas à Chapada Diamantina e que os eventos de expansão se iniciaram há aproximadamente 40 mil anos. A partir daí, expansão de área em direção ao norte explica a colonização de localidades na Chapada da Borborema – PB e em direção ao litoral explica a colonização de localidades na costa Atlântica brasileira (FRANCO; MANFRIN, 2013).

1.4 Marcadores moleculares

Por meio do estudo da diversidade existente entre indivíduos e sua história evolutiva, é possível inferir como surgem linhagens diferentes. Dentro de uma população, a relação entre as

gerações mantém a unidade entre os indivíduos. A partir do momento que essa população é dividida e isolada, as mudanças gênicas produzidas naturalmente pelas mutações e relações ecológicas poderão determinar duas novas linhagens (JUDD et al, 2008).

Marcadores fisiológicos, morfológicos e fenotípicos eram os únicos disponíveis até a década de 1960 para contribuição dos métodos estatísticos de genética de populações. Marcadores fenotípicos podem ser influenciados por fatores não genéticos, e os caracteres disponíveis para cada organismo é muito baixo. As principais vantagens dos marcadores moleculares são o grande número de caracteres informativos e o fato de serem pouco influenciados pelo ambiente (AVISE, 1994).

Duas fontes possíveis de informações a partir de moléculas de DNA estão disponíveis, sendo: nuclear ou organelar – mitocondrial em animais e mitocondrial e plastidial em plantas.

1.4.1 Gene mitocondrial COI

O gene *Citocromo Oxidase I* – COI está no genoma da mitocôndria, organela herdada predominantemente de forma clonal e matrilinear. Foi o marcador mais utilizado no início dos estudos filogeográficos, devido à elevada taxa evolutiva e ausência de recombinação (AVISE, 2000, 2009; BRITO, EDWARDS, 2009; HICKERSON et al, 2010). O tamanho efetivo populacional dos genes deste genoma é, consequentemente, 25% dos genes nucleares. As taxas de mutações nesse gene também são maiores do que em genes nucleares (AVISE et al, 1987; TEMPLETON, 2004).

A evolução do DNA mitocondrial é principalmente consequência de mutações de ponto. Isso simplifica o alinhamento das sequências e compreensão dos dados nas análises populacionais (AVISE, 1987; TEMPLETON, 2004).

Dentre os genes mitocondriais, o COI é o mais adequado para estudos intra e interespecíficos de metazoários, pois possui divergência suficiente entre as espécies e entre as populações, para que estas possam ser apontadas e terem suas relações evolutivas estabelecidas (AVISE, 2000, 2009).

1.4.2 Genes nucleares autossômicos *GstD1* e α esterase-5

A composição química dos cactos, que varia entre as espécies e pode ser tóxica para alguns grupos, é de grande influência na relação cacto hospedeiro e as espécies de *Drosophila* (FRANK; FOGLEMAN, 1992). Para os processos de desintoxicação nessa relação, diferentes *loci* estão envolvidos. Entre eles, os genes da família citocromo P450 (BONO et al, 2008), glutathione S-transferase (Gst) (MATZKIN et al, 2006) e esterase (ROBIN et al, 2000).

A relação entre os cactos hospedeiros e linhagens evolutivas foi analisada em *D. mojavensis* (MATZKIN, 2008) e *D. mettleri* (BONO et al, 2008) por meio de análises de genes pertencentes às famílias glutathione S-transferase (*GstD1*) e citocromo P450, respectivamente, e o que se observou foi a presença significativa de diferenciação genética entre moscas coletadas de diferentes espécies de plantas hospedeiras. Em análises filogenéticas, o gene α -esterase5 (α E5) mostrou-se instrutivo em nível interespecífico para as espécies do “cluster” *Drosophila buzzatti* (SANTOS et al, 2009).

1.4.3 Gene nuclear *kl-5* ligado ao cromossomo Y

O cromossomo Y no gênero *Drosophila* foi um dos primeiros a ser definido geneticamente e citologicamente, contudo há poucas informações para identificação de genes com cópia única, assim como quais forças evolutivas moldaram este cromossomo (CARVALHO, 2002). Contudo, o cromossomo Y mostra ser pouco conservado devido ao grande número de ganhos e perdas de genes, sendo que o ganho mostra que ele seja mais recente com relação aos demais cromossomos (KOERICH et al, 2008). É possível que a variação genética do cromossomo Y esteja envolvida na competição entre espermatozoides (fator que pode influenciar o *fitness* dos machos) (CARVALHO, 2002; CHIPPINDALE, RICE, 2001) e, com isso, selecionar determinados polimorfismos.

Segundo Gepner e Hays (1993), o gene *kl-5* foi identificado como um gene de cópia única que codifica a proteína *dynein heavy chain*, um fator de fertilidade que está relacionado com o processo de espermatogênese. Os fatores de fertilidade *kl-1*, *kl-2*, *kl-3*, *ks-1* e *ks-2* foram

descritos para o cromossomo Y e considerados de cópia única (CARVALHO, 2002). Estas descobertas contrariaram a hipótese de que o cromossomo Y foi originado a partir do cromossomo X, uma vez que genes parálogos dos fatores de fertilidade são autossômicos e não ligados ao X como o esperado (CARVALHO, 2002).

2 Objetivo

O presente trabalho teve como objetivo testar a existência de linhagens evolutivas em *Drosophila serido* por meio de análises filogenéticas, estrutura e agrupamento populacional e redes haplotípicas utilizando sequências parciais de genes nucleares autossômicos e sexual e mitocondrial. Os objetivos específicos foram:

- 1- Aumentar o número amostral de *Drosophila serido*;
- 2- Sequenciar parcialmente genes mitocondrial e nucleares;
- 3- Realizar análises de estrutura populacional para as populações hipoteticamente estabelecidas (nordeste e litoral);
- 4- Realizar análises comparativas de haplótipos e sequências de DNA para agrupamentos das populações;
- 5- Estabelecer as possíveis linhagens evolutivas para a espécie.

3 Material e métodos

3.1 Material biológico

Neste trabalho, foram analisadas 299 sequências de 42 localidades (Tabela 1, Figura 1), sendo amplificadas 410 pb do gene nuclear autossômico GstD1, 759 pb do gene nuclear autossômico α E5, 912 pb do gene nuclear sexual kl-5 e 629 pb do gene mitocondrial COI. Parte das sequências, assim como amostras de DNA previamente extraídos de indivíduos de *D.serido* fazem parte do acervo do Laboratório de Genética Evolutiva (FFCLRP-USP) e foram utilizadas neste trabalho. Para complementação das amostras, o DNA genômico de indivíduos foi extraído por meio do conjunto de reagentes *DNeasy Blood & Tissue Kit* (Qiagen). Além dessas, foram incluídas nas análises sequências de *Drosophila mojavensis* (grupo *D. repleta*: acesso GenBank COI- Dmoj\GI16346; Esterase- Dmoj\GI15943; GstD1- EU079442.1) que foram utilizadas como *outgroup*.

Tabela 1. Localidades de populações amostradas, com respectivos coletores e datas de coletas, e número de sequências gênicas de indivíduos da espécie *Drosophila serido*.

CL	Localidade	Coletores	Data da coleta	CG	COI	kl5	Gstd1	esterase
N40	10km Morro do Chapéu - BA	Franco FF, Oliveira C, Esguícero A	05 a 14-VIII-2008	11,65°S; 41,29°W	5	4	5	4
R56	Almenara - MG	Lavagnini TC, Leal DYB, Lavagnini AC	19 a 28-IX-2013	16,00°S; 40,95°W	10	9	8	6
N46	Andaraí - BA	Franco FF, Oliveira C, Esguícero A	05 a 14-VIII-2008	12,86°S; 41,31°W	6	6	5	5
N20	Arraial do Cabo - RJ	Morales AC, Franco FF, Silva-Bernardi ECC	14 a 20-III-2004	23,00°S; 42,00°W	6	6	6	6
H49	Bertioga - SP	Sene FM, Monteiro SG, Ruiz A	10 a 18-V-1995	23,90°S; 46,10°W	4	4	4	3
R38	Bom Jesus - RN	Lavagnini TC, Leal DYB, Lavagnini AC, Lavagnini MC	23 a 30-VI-2012	06,02°S; 35,65°W	6	6	6	5
N71	Cabedelo - PB	Franco FF, Santos MHS	16 a 23-02-2008	07,03°S; 34,83°W	7	6	6	5
N35	Cabrália - BA	Sene FM, Franco FF, Almeida JG	04 a 12-VII-2005	16,31°S; 39,01°W	5	5	5	5
N42	Cach. Ferro Doido - BA	Franco FF, Oliveira C, Esguícero A	05 a 14-VIII-2008	11,63°S; 41,00°W	10	6	6	6
N10	Camboriu - SC	Morales AC, Franco FF, Silva-Bernardi ECC	24 a 31-III-2003	27,05°S; 48,59°W	0	3	0	3
R32	Costa Azul - BA	Lavagnini TC, Leal DYB, Manfrin MH, Mateus RP	15 a 22-IX-2012	11,61°S; 37,46°W	6	4	3	6
N67	Exu - PE	Franco FF, Santos MHS	16 a 23-02-2008	07,65°S; 39,79°W	2	2	0	1
N9	Gov. Celso Ramos - SC	Morales AC, Franco FF, Silva-Bernardi ECC	24 a 31-III-2003	27,37°S; 48,53°W	0	3	0	3

Tabela 1. Continuação.

N2	Guaratuba - SP	Manfrin MH, Prado N	VIII-2002	23,85°S; 46,13°W	6	3	4	6
Ilha	Ilha de Alcatrazes - SP			24,10°S; 45,69°W	0	6	6	6
N43	Irecê - BA	Franco FF, Oliveira C, Esguícero A	05 a 14-VIII-2008	11,22°S; 41,95°W	9	6	6	6
J95	Itaberaba - BA	Morales AC, Kuhn GCS, Franco FF, Brisson JÁ	23-III-2002	12,56°S; 40,31°W	4	2	3	4
N31	Itaúnas - ES	Sene FM, Franco FF, Almeida JG	04 a 12-VII-2005	18,41°S; 39,70°W	6	2	4	6
R59	Jequitinhonha - MG	Lavagnini TC, Leal DYB, Lavagnini AC	19 a 28-IX-2013	16,52°S; 41,30°W	1	1	1	0
R34	Jucurutu - RN	Lavagnini TC, Leal DYB, Lavagnini AC, Lavagnini MC	23 a 30-VI-2012	05,95°S; 37,04°W	1	1	1	1
N70	Junco do Serido - PB	Franco FF, Santos MHS	16 a 23-02-2008	07,00°S; 36,71°W	10	6	6	6
R37	Lajes Pintadas - RN	Lavagnini TC, Leal DYB, Lavagnini AC, Lavagnini MC	23 a 30-VI-2012	06,17°S; 36,11°W	8	6	6	6
N22	Macaé - RJ			22,32°S; 41,82°W	5	3	3	5
J93	Manoel Vitorino - BA	Morales AC, Kuhn GCS, Franco FF, Brisson JA	23/03/2002	14,15°S; 40,24°W	0	6	6	5
R39	Maxaranguape - RN	Lavagnini TC, Leal DYB, Lavagnini AC, Lavagnini MC	23 a 30-VI-2012	05,54°S; 35,27°W	5	6	6	6
J92	Milagres - BA	Morales AC, Kuhn GCS, Franco FF, Brisson JA	23/03/2002	11,20°S; 39,90°W	6	6	4	5
R28	Monte Alegre de Sergipe - SE	Lavagnini TC, Leal DYB, Manfrin MH, Mateus RP	15 a 22-IX-2012	10,13°S; 37,54°W	6	5	3	5

Tabela 1. Continuação.

N37	Morro da Barrinha - BA	Franco FF, Oliveira C, Esguícero A	05 a 14-VIII-2008	09,90°S; 40,30°W	10	6	4	5
N45	Mucugê - BA	Franco FF, Oliveira C, Esguícero A	05 a 14-VIII-2008	13,00°S; 41,40°W	5	4	3	4
N32	Mucuri - BA	Sene FM, Franco FF, Almeida JG	04 a 12-VII-2005	18,00°S; 39,42°W	5	6	6	6
R31	Pacatuba - SE	Lavagnini TC, Leal DYB, Manfrin MH, Mateus RP	15 a 22-IX-2012	10,46°S; 36,63°W	6	6	6	5
N19	Paúba - SP	Morales AC, Franco FF	05 a 09-X-2003	23,80°S; 45,55°W	0	6	6	6
R30	Penedo - AL	Lavagnini TC, Leal DYB, Manfrin MH, Mateus RP	15 a 22-IX-2012	10,32°S; 36,52°W	6	6	6	6
J52	Penha - SC			26,75°S; 48,66°W	7	5	1	5
N68	Pocinhos - PB	Franco FF, Santos MHS	16 a 23-02-2008	07,17°S; 36,05°W	4	6	6	6
R29	Porto das Cabras - SE	Lavagnini TC, Leal DYB, Manfrin MH, Mateus RP	15 a 22-IX-2012	10,80°S; 36,93°W	5	4	1	4
R27	Riacho - BA	Lavagnini TC, Leal DYB, Manfrin MH, Mateus RP	15 a 22-IX-2012	09,61°S; 38,22°W	4	5	5	5
R33	Rodolfo Fernandes - RN	Lavagnini TC, Leal DYB, Lavagnini AC, Lavagnini MC	23 a 30-VI-2012	05,82°S; 37,99°W	1	1	1	1
R36	Russas - CE	Lavagnini TC, Leal DYB, Lavagnini AC, Lavagnini MC	23 a 30-VI-2012	04,85°S; 38,04°W	6	6	4	4
N65	Salvador - BA	Franco FF	07 a 09-X-2007	12,93°S; 39,32°W	10	5	6	6
N29	São Sebastião - SP	Morales AC, Franco FF	05 a 09-X-2003	23,82°S; 45,42°W	7	3	2	5

Tabela 1. Continuação.

R35	Tabuleiro do Norte - CE	Lavagnini TC, Leal DYB, Lavagnini AC, Lavagnini MC	23 a 30-VI-2012	05,28°S; 38,16°W	2	2	2	2
-----	-------------------------	---	-----------------	---------------------	---	---	---	---

CL: códigos das localidades; **CG:** coordenadas geográficas (grau-decimal); **genes COI, kl5, Gstd1 e esterase:** número de amostras sequenciadas para cada gene.

Figura 1. Mapa do Brasil indicando as localidades amostradas. Os códigos das localidades estão discriminados na Tabela 1. As localidades agrupadas em azul pertencem ao grupo nordeste e as localidades agrupadas em vermelho pertencem ao grupo litoral.

3.2 Amplificação, purificação e sequenciamento dos genes nucleares α E5, GstD1, kl-5 e mitocondrial COI

A amplificação das regiões de interesse dos genes nucleares α E5 (Figura 2), GstD1 (Figura 3), kl-5 (Figura 4) e do gene mitocondrial COI (Figura 5) foi realizada por meio da reação em cadeia da polimerase (PCR) em um termociclador *Veriti® Thermal Cycler 96-well* (Applied Biosystems).

Figura 2 – Estrutura do gene nuclear α -esterase5 em *Drosophila borborema*. Os íntrons são representados por linhas e os éxons por retângulos. Abaixo da estrutura do gene, a barra negra indica a região analisada no presente trabalho. Modificada a partir de Santos et al (2009).

Figura 3 – Esquema do gene nuclear GstD1 em *Drosophila mojavensis*. O retângulo cinza representa a totalidade do gene, que não possui íntrons. Abaixo desta representação, a barra negra indica a região analisada no presente trabalho.

Figura 4 A – A representação parcial da estrutura do cromossomo Y em *Drosophila melanogaster* com detalhe para a localização dos genes *kl-5*, *kl-3* e *kl-2*. **B** – detalhe para o éxon 12 do gene *kl-5*, sendo que a barra negra corresponde à região amplificada e analisada no presente trabalho.

Figura 5 – Esquema do genoma mitocondrial, obtido em Oliveira; Garesse; Kaguni (2010), em *Drosophila melanogaster* com detalhe para a localização do gene COI. Os retângulos branco, cinza e negro correspondem, respectivamente, aos genes ND2, COI e COII, presentes na

estrutura do genoma mitocondrial. Abaixo da estrutura do gene COI há uma barra negra que corresponde à região amplificada e analisada no presente trabalho.

A concentração de cada reagente e a quantidade de DNA extraído utilizadas para amplificação dos genes de interesse estão descritas na Tabela 2. Em todas as reações foi utilizada água estéril para completar o volume final de 25µl. Os oligonucleotídeos iniciadores *forward* e *reverse* para cada gene e as condições da amplificação estão descritos na Tabela 3.

Tabela 2 – Concentração dos reagentes utilizados para realização da reação em cadeia da polimerase (PCR) para amplificação dos genes de interesse.

	<i>GstD1</i>	<i>kl-5</i>	<i>COI</i>	<i>αE-5</i>
DNA extraído	50ng	50ng	50ng	50ng
Tampão	1x	1x	1x	1x
MgCl ₂	1,5mM	2,0mM	1,5mM	1,5mM
Taq polimerase*	1,25U	1,25U	1,25U	1,25U
dNTP**	200µM	200µM	200µM	200µM
Oligonucleotídeos iniciadores	0,3µM	0,3µM	0,3µM	0,5µM

*GoTaq® Hot Start Polymerase (Promega)

**Fermentas

Tabela 3 – Oligonucleotídeos iniciadores *forward* e *reverse* e condições de PCR para cada gene de interesse.

Gene (referência)	Oligonucleotídeo iniciador <i>forward</i>	Oligonucleotídeo iniciador <i>reverse</i>	Condições PCR
<i>GstD1</i> (PIZZO, T. C. L. 2015)	GstD1 deg-F 5' CCCGAGTTCGTGAAGATCAAYCCNCA 3'	GstD1 deg-R 5' TTCTCGTACCACTTGTTACGTTNKCR 3'	1 ciclo – 2' a 94°C 30 ciclos – 1' a 94°C 1' a 56°C 2' a 76°C 1 ciclo – 5' a 72°C
<i>kl-5</i> (DYER et al, 2011)	KI5-F8 5' KTTYGARYTRCAAGGKCCAGATCC 3'	KI5-R6 5' GCSGGCCAYTCRAATATCCAMAC 3'	1 ciclo – 2' a 94°C 40 ciclos – 1' a 94°C 1' a 56°C 2' a 72°C 1 ciclo – 5' a 72°C
<i>COI</i> (SIMON et al, 1994)	TY-J-1460 5' TACAATTTATCGCCTAAACTTCAGCC 3'	C1-N-2191 5' CCCGGTAAAATTAAAATATAAACTTC 3'	1 ciclo – 1'30'' a 94°C 25 ciclos – 40'' a 94°C 40'' a 48°C 2' a 72°C 1 ciclo – 4' a 72°C
<i>αE-5</i> (SANTOS et al, 2009)	Ester 2F 5' TGGACTGAAGGACCAGGTTT 3'	Ester CH1R 5' AGCCATGCCAGAAGATCCTA 3'	1 ciclo – 4' a 94°C 35 ciclos – 40'' a 94°C 40'' a 51°C 1' a 72°C 1 ciclo – 3'30'' a 72°C

Os produtos de PCR foram verificados por meio da eletroforese em gel de agarose 1% e tampão de cuba TBE 1x (Tris-base 0,89M; Ácido Bórico 0,89M; EDTA 20mM). Para visualização dos produtos de PCR, o gel de agarose 1% foi corado com *SYBR Safe DNA gel stain (Life Technologies)*. A eletroforese foi realizada em 90V por aproximadamente 1 hora. O DNA no gel foi visualizado por meio da exposição à luz UV em um transiluminador. Depois de checados, os produtos de PCR foram purificados com o conjunto de reagentes *GFX PCR and gel band purification kit (GE Healthcare)* seguindo-se as instruções do fabricante. Em seguida, as amostras foram analisadas em um NanoDrop ND-1000-V.3 para quantificação do DNA e preparo da reação de sequenciamento.

A reação de sequenciamento foi realizada com auxílio do conjunto de reagentes *ABI PRISM BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (Life Technologies)*. O volume final da reação de sequenciamento era de 10µL, composta de 1µl de BigDye, 10µM de cada oligonucleotídeo iniciador, tampão 0,75x, até 6,5µl de produto de PCR purificado, para que a concentração fosse de 50ng/µl, e água estéril para completar o volume final, quando necessário. Para todos os genes, no sequenciamento foram usados os mesmos oligonucleotídeos iniciadores utilizados nas reações de amplificação mencionadas anteriormente (Tabela 3). As condições gerais da reação de sequenciamento consistiram de uma etapa inicial de 1' a 96°C e 39 ciclos de 15'' a 96°C, 15'' a T°C e 4' a 60°C, em que "T" corresponde às temperaturas de pareamento dos oligonucleotídeos iniciadores para cada gene (Tabela 3).

Para a precipitação, foram realizadas várias lavagens do produto da reação de sequenciamento. Primeiramente, foram adicionados 80µl de isopropanol 65% às amostras e estas foram deixadas em repouso por 15 minutos à temperatura ambiente. Em seguida, as amostras foram centrifugadas a 4.000 rpm por 45 minutos à 4°C e o sobrenadante descartado. Posteriormente, cada amostra recebeu 200µl de etanol 70% resfriado, sendo centrifugadas a 4.000 rpm por 10 minutos à 4°C. O sobrenadante foi descartado e esta etapa foi repetida. As amostras foram mantidas na placa aquecedora a 37°C por aproximadamente 1 hora, até a completa evaporação do álcool. As amostras foram enviadas para sequenciamento no Centro de Recursos Biológicos e Biologia Genômica – CREBIO, UNESP/FCAV, Jaboticabal – SP.

3.3 Análise dos dados

3.3.1 Leitura e alinhamento das sequências

A leitura e alinhamento das sequências de DNA foram realizados com auxílio dos programas Chromas v.2.01 Copyright© (1998-2005) Technelysium Pty Ltda, ClustalW v.1.8 (THOMPSON; HIGGINS; GIBSON, 1994), implementado no programa BioEdit v.7.1.11 (HALL, 1999) e MUSCLE v.3.8.31 (EDGAR, 2004). Quando necessário, as sequências foram inspecionadas manualmente.

3.3.2 Estatística descritiva

Os valores da estatística descritiva, que inclui o número de sítios polimórficos (S), diversidade haplotípica (Hd), diversidade nucleotídica (π), número médio de diferenças nucleotídicas (k), foram obtidos por meio do programa DnaSP v.5.1 (LIBRADO; ROZAS, 2009), que também foi utilizado para o cálculo do teste Rm, que avalia o número mínimo de eventos de recombinação na amostra. Além do teste Rm, o teste PHI foi realizado para avaliar recombinação (BRUEN; PHILIPPE; BRYANT, 2006), por meio do programa SplitsTree4 v.4.13.1 (HUSON; BRYANT, 2006). Este teste indica evidência de recombinação quando o valor de $p < 0,05$.

3.3.3 Modelos de substituição nucleotídica

Os modelos de substituição nucleotídica foram obtidos com auxílio do programa JModelTest v.2.1.2 (POSADA, 2008) e escolhidos com base no *Akaike Information Criterion* (AIC), em que o melhor modelo é indicado pelo menor valor de AIC (POSADA; BUCKLEY, 2004; POSADA, CRANDALL, 1998). Este critério escolhe o modelo que melhor se aproxima do real processo de evolução molecular dos fragmentos (POSADA; BUCKLEY, 2004).

3.3.4 Análise de Variância Molecular e Teste de Mantel

A estruturação entre as populações foi analisada por meio da análise de variância molecular (AMOVA) (EXCOFFIER et al, 1992), com auxílio do programa Arlequin v.3.5 (EXCOFFIER; LISCHER, 2010). Esta análise estima a variabilidade intra- e interespecífica de cada espécie, fornecendo valor de Φ , que reflete a correlação da diversidade haplotípica: Φ_{st} : covariância entre indivíduos dentro de populações.

Para avaliar se há correlação entre a distância genética e a distância geográfica foi realizado o teste de Mantel por meio do programa Arlequin v.3.5 (EXCOFFIER; LISCHER, 2010). A matriz de distância geográfica entre as populações foi estabelecida por meio do programa Geographic Distance Matriz Generator v.1.2.3 (ERSTS, 2006). A matriz de distância genética, F_{st} par-a-par, foi calculada pelo programa Arlequin v.3.5 (EXCOFFIER; LISCHER, 2010). O coeficiente de correlação (r) foi considerado significativo quando $p < 0,05$.

3.3.5 Reconstrução filogenética

A análise filogenética bayesiana foi realizada por meio do programa BEAST v.1.8.2 (BOUCKAERT et al, 2014). Foi realizada apenas 1 réplica e cada análise realizada com 10 milhões de interações, com registros a cada 1.000 passos, para cada gene de interesse. Entretanto, para a análise conjunta dos genes, foram realizadas 2 réplicas com 100 milhões de interações, com registros a cada 1.000 passos.

As análises foram combinadas no programa Log Combiner e a convergência e estabilidade checadas usando o programa Tracer v.1.6 (RAMBAUT et al, 2014). As árvores geradas foram sumarizadas em uma com *Maximum clade credibility tree* (MCC), no programa TreeAnnotator. O valor de *burn in* foi de 10% do total de árvores amostradas. As árvores foram visualizadas por meio do programa FigTree v.1.4.2 (RAMBAUT, 2012).

3.3.6 Análise da estrutura populacional

A estrutura populacional foi estimada usando o método de agrupamento bayesiano implementado no programa BAPS v.6.0 (CORANDER et al, 2013). Este programa tem duas principais vantagens, sendo uma delas considerar o número de populações como um parâmetro desconhecido e a segunda todos os grupos de combinações são considerados igualmente prováveis *a priori*. Posteriormente, determinará qual a partição mais provável entre os grupos de população, dentro de um número de agrupamentos empiricamente plausível. As análises foram realizadas considerando o modelo *clustering with linked loci* com um número máximo de grupos (k) de 1 - 10 e dez réplicas para cada.

3.3.7 Análise de componentes principais

A análise de componentes principais foi realizada por meio do programa GenALEx 6.51B2 (PEAKALL; SMOUSE, 2012), utilizando uma matriz haplotípica. Os haplótipos foram gerados pelo programa TCS1.21 (CLEMENT et al, 2000) e a matriz foi feita manualmente usando o Microsoft Excel. A análise foi realizada considerando o método de agrupamento por populações baseado na distância genética entre elas. Esse método permite a redução da dimensionalidade dos pontos representativos das amostras, pois, embora a informação estatística presente nas n-variáveis originais seja a mesma dos n-componentes principais, é comum obter em apenas 2 ou 3 dos primeiros componentes principais boa parte desta informação. O primeiro PC explica a maior parte da variação, enquanto o segundo PC explica a segunda maior fração e o terceiro PC, a terceira maior fração.

3.3.8 Redes haplotípicas

As redes haplotípicas para cada gene foram geradas através do programa Haplotype Viewer (<http://www.cibiv.at/~greg/haploviewer>), para visualização e edição de cores concernentes aos dados de sequências nucleares e mitocondrial. Para uma visualização mais

nítida, as populações do litoral foram apresentadas com a cor vermelha e, as populações do nordeste, foram apresentadas com a cor azul.

4 Resultados e discussão

4.1 Estatística descritiva e modelo de substituição nucleotídica

4.1.1 Gene COI

Foram sequenciados 629 pb em 211 indivíduos amostrados em 34 localidades. A estatística descritiva mostrou que há 60 haplótipos (h) com diversidade haplotípica (Hd) de 0,771, 60 sítios polimórficos (S), diversidade nucleotídica (π) de 0,00708 e número médio de diferenças nucleotídicas (k) de 4,154. A estatística descritiva para as populações de *D. serido* analisadas por meio de sequências parciais do gene COI está apresentada na Tabela 4.

Tabela 4. Resultados obtidos da estatística descritiva para cada população de *Drosophila serido* a partir de sequências do gene mitocondrial COI.

Localidade	S	h	Hd	π	K
N71	04	05	0,90	0,00212	1,333
N70	07	08	0,95	0,00272	1,711
N68	02	03	0,83	0,00185	1,167
N67	01	02	1,00	0,00159	1,000
J92	02	03	0,73	0,00148	0,933
N32	09	02	0,60	0,00859	5,400
N37	18	07	0,93	0,01035	6,511
N40	03	03	0,80	0,00223	1,400
N43	19	07	0,92	0,01157	7,278
N42	01	02	0,20	0,00032	0,200
N45	00	01	0,00	0,00000	0,000
N65	11	06	0,87	0,00392	2,467
N35	09	03	0,70	0,00572	3,600
N46	03	03	0,60	0,00191	1,200
J95	00	01	0,00	0,00000	0,000
N31	11	04	0,80	0,00880	5,533
N20	00	01	0,00	0,00000	0,000
N22	02	03	0,70	0,00127	0,800
N29	01	02	0,29	0,00045	0,286
H49	01	02	0,67	0,00106	0,667
N2	01	02	0,53	0,00085	0,533
J52	00	01	0,00	0,00000	0,000
R27	01	02	0,67	0,00106	0,667

Tabela 4. Continuação.

R28	02	03	0,73	0,00138	0,867
R29	05	05	1,00	0,00318	2,000
R30	03	04	0,80	0,00159	1,000
R31	03	04	0,80	0,00159	1,000
R32	05	04	0,80	0,00265	1,667
R35	01	02	1,00	0,00159	1,000
R36	03	04	0,87	0,00191	1,200
R37	03	04	0,64	0,00119	0,750
R38	01	02	0,33	0,00053	0,333
R39	02	03	0,70	0,00159	1,000
R56	02	03	0,51	0,00088	0,556

S = número de sítios polimórficos; h = número de haplótipos; Hd = diversidade haplotípica; π = diversidade nucleotídica e k = número médio de diferenças nucleotídicas. Os códigos das localidades são os mesmos apresentados na Tabela 1.

O modelo de substituição nucleotídica que melhor explica o processo de evolução dos fragmentos obtidos para o gene COI é TPM3uf+ I + G (manual programa JModelTest), contudo, na ausência deste modelo, dentre as opções dos programas utilizados, foi escolhido o modelo HKY+ I + G (HASEGAWA; KISHINO; YANO, 1985). A taxa de substituição nucleotídica para o gene COI, estimada a partir de calibração secundária, é de 1,084 substituições/sítio/Myr.

4.1.2 Gene α esterase-5

Foram sequenciados 759 pb em 136 indivíduos amostrados em 41 localidades. A estatística descritiva mostrou que há 117 haplótipos (h) com diversidade haplotípica (Hd) de 0,958, 145 sítios polimórficos (S), diversidade nucleotídica (π) de 0,10452 e número médio de diferenças nucleotídicas (k) de 15,155. A estatística descritiva para as populações de *D. serido* analisadas por meio de sequências do gene α E5 está apresentada na Tabela 5. O teste Rm ($p=1,0$) não encontrou evidência significativa para um número mínimo de eventos de recombinação na amostra, entretanto, o teste PHI ($p=0,0001$) encontrou evidência estatisticamente significativa a favor da presença de eventos de recombinação.

Tabela 5. Resultados obtidos da estatística descritiva para cada população de *Drosophila serido* a partir de sequências do gene α esterase-5.

Localidade	S	h	Hd	π	K
N71	06	03	0,64	0,01050	1,533
N70	32	07	0,83	0,06040	8,818
N68	20	07	0,88	0,03829	5,591
N67	00	01	0,00	0,00000	0,000
J92	17	08	0,93	0,04384	6,400
J93	22	05	0,82	0,06408	9,356
N32	28	09	0,91	0,05334	7,788
N37	14	04	0,73	0,03044	4,444
N40	18	05	8,6	0,05626	8,214
N43	27	07	0,92	0,05677	8,288
N42	28	10	0,97	0,07078	10,333
N45	22	04	0,82	0,07290	10,643
N65	17	07	0,86	0,03856	5,258
N35	11	05	0,82	0,02953	4,311
N46	25	07	0,93	0,06956	10,156
J95	36	07	0,96	0,07877	11,500
N31	16	07	0,77	0,03279	4,788
N20	15	07	0,83	0,03809	5,561
N22	14	07	0,91	0,04186	6,111
N29	11	05	0,80	0,03105	4,533
H49	01	02	0,53	0,00365	0,533
N2	02	03	0,68	0,00560	0,818
ILHA	23	06	0,68	0,06382	9,318
N19	09	06	0,85	0,02646	3,864
N9	01	02	0,33	0,00228	0,333
N10	00	01	0,00	0,00000	0,000
J52	00	01	0,00	0,00000	0,000
R27	15	05	0,84	0,04612	6,733
R28	22	08	0,95	0,04673	6,822
R29	13	05	0,89	0,03596	5,250
R30	12	08	0,89	0,02252	3,288
R31	16	05	0,80	0,04323	6,311
R32	14	05	0,67	0,02127	3,106
R33	12	02	1,00	0,08219	12,000
R34	11	02	1,00	0,07534	11,000
R35	04	02	0,50	0,01370	2,000
R36	06	03	0,61	0,01663	2,429
R37	22	09	0,95	0,06155	8,924
R38	14	06	0,89	0,02968	4,333

Tabela 5. Continuação.

R39	14	06	0,81	0,02698	3,939
R56	27	08	0,89	0,05853	8,545

S = número de sítios polimórficos; h = número de haplótipos; Hd = diversidade haplotípica; π = diversidade nucleotídica e k = número médio de diferenças nucleotídicas. Os códigos das localidades são os mesmos apresentados na Tabela 1.

O modelo de substituição nucleotídica que melhor explica o processo de evolução dos fragmentos obtidos para o gene $\alpha E5$ é TPM2+ I + G (manual programa JModelTest), contudo, na ausência deste modelo dentre as opções dos programas utilizados, foi escolhido o modelo HKY+ I + G (HASEGAWA; KISHINO; YANO, 1985). A taxa de substituição nucleotídica para o gene $\alpha E5$, estimada a partir de calibração secundária, é de 1,035 substituições/sítio/Myr.

4.1.3 Gene GstD1

Foram sequenciados 410 pb em 147 indivíduos amostrados em 39 localidades. A estatística descritiva mostrou que há 32 haplótipos (h) com diversidade haplotípica (Hd) de 0,849, 41 sítios polimórficos (S), diversidade nucleotídica (π) de 0,11421 e número médio de diferenças nucleotídicas (k) de 4,683. A estatística descritiva para as populações de *D. serido* analisadas por meio de sequências do gene GstD1 está apresentada na Tabela 6. O teste Rm ($p=0,1$) não encontrou evidência significativa para um número mínimo de eventos de recombinação na amostra, no entanto, o teste PHI ($p=0,002$) encontrou evidência estatisticamente significativa a favor da presença de eventos de recombinação.

Tabela 6. Resultados obtidos da estatística descritiva para cada população de *Drosophila serido* a partir de sequências do gene GstD1.

Localidade	S	h	Hd	π	K
N71	04	02	0,54	0,05322	2,182
N70	06	04	0,70	0,06135	2,515
N68	05	03	0,62	0,05728	2,348
J92	08	05	0,86	0,08014	3,286
J93	07	04	0,71	0,07206	2,955
N32	08	06	0,85	0,08167	3,348

Tabela 6. Continuação.

N37	07	04	0,75	0,07404	3,036
N40	06	04	0,73	0,06396	2,622
N43	10	05	0,76	0,07761	3,182
N42	08	05	0,79	0,07280	2,985
N45	09	05	0,93	0,09919	4,067
N65	07	05	0,76	0,06541	2,682
N35	11	07	0,93	0,11707	4,800
N46	09	05	0,80	0,07859	3,222
J95	08	05	0,93	0,09106	3,733
N31	11	05	0,86	0,12631	5,179
N20	13	06	0,80	0,09978	4,091
N22	09	06	1,00	0,10894	4,467
N29	07	03	0,83	0,10569	4,333
H49	12	05	0,86	0,13850	5,679
N2	12	05	0,86	0,13415	5,500
ILHA	09	05	0,79	0,10015	4,106
N19	13	05	0,76	0,12454	5,106
J52	07	02	1,00	0,17073	7,000
R27	05	03	0,64	0,05908	2,422
R28	05	04	0,87	0,07154	2,933
R29	04	02	1,00	0,09756	4,000
R30	04	02	0,54	0,05322	2,182
R31	05	03	0,62	0,05728	2,348
R32	05	03	0,73	0,06667	2,733
R33	04	02	1,00	0,09756	4,00
R34	05	02	1,00	0,12195	5,00
R35	04	02	0,67	0,06504	2,667
R36	04	02	0,57	0,05575	2,286
R37	08	06	0,81	0,06948	2,848
R38	05	03	0,62	0,05728	2,348
R39	06	03	0,62	0,06135	2,515
R56	10	06	0,82	0,08638	3,542
R59	06	02	1,00	0,14634	6,000

S = número de sítios polimórficos; h = número de haplótipos; Hd = diversidade haplotípica; π = diversidade nucleotídica e k = número médio de diferenças nucleotídicas. Os códigos das localidades são os mesmos apresentados na Tabela 1.

O modelo de substituição nucleotídica que melhor explica o processo de evolução dos fragmentos obtidos para o gene GstD1 é TIM2ef +I (manual programa JModelTest), contudo, na ausência deste modelo dentre as opções dos programas utilizados, foi escolhido o modelo

TN93+ I (TAMURA; NEI, 1993). A taxa de substituição nucleotídica para o gene GstD1, estimada a partir da calibração secundária, é de 1,071 substituições/sítio/Myr.

4.1.4 Gene Kl-5

Foram sequenciados 912 pb em 194 indivíduos amostrados em 39 localidades. A estatística descritiva mostrou que há 5 haplótipos (h) com diversidade haplotípica (Hd) de 0,337, 6 sítios polimórficos (S), diversidade nucleotídica (π) de 0,00107 e número médio de diferenças nucleotídicas (k) de 1,221. A estatística descritiva para as populações de *D. serido* analisadas por meio de sequências do gene kl-5 está apresentada na Tabela 7.

Tabela 7. Resultados obtidos da estatística descritiva para cada população de *Drosophila serido* a partir de sequências do gene Kl-5.

Localidade	S	h	Hd	π	K
N71	00	01	0,00	0,00000	0,000
N70	00	01	0,00	0,00000	0,000
N68	00	01	0,00	0,00000	0,000
N67	01	02	1,00	0,00083	1,000
J92	01	02	0,53	0,00044	0,533
J93	00	01	0,00	0,00000	0,000
N32	00	01	0,00	0,00000	0,000
N37	00	01	0,00	0,00000	0,000
N40	00	01	0,00	0,00000	0,000
N43	00	01	0,00	0,00000	0,000
N42	00	01	0,00	0,00000	0,000
N45	00	01	0,00	0,00000	0,000
N65	00	01	0,00	0,00000	0,000
N35	00	01	0,00	0,00000	0,000
N46	00	01	0,00	0,00000	0,000
J95	00	01	0,00	0,00000	0,000
N31	00	01	0,00	0,00000	0,000
N20	00	01	0,00	0,00000	0,000
N22	00	01	0,00	0,00000	0,000
N17	00	01	0,00	0,00000	0,000
H49	01	02	0,50	0,00041	0,500
N2	00	01	0,00	0,00000	0,000

Tabela 7. Continuação.

ILHA	00	01	0,00	0,00000	0,000
N19	00	01	0,00	0,00000	0,000
N9	01	02	0,67	0,00055	0,667
N10	00	01	0,00	0,00000	0,000
J52	00	01	0,00	0,00000	0,000
R27	00	01	0,00	0,00000	0,000
R28	01	02	0,40	0,00033	0,400
R29	00	01	0,00	0,00000	0,000
R30	00	01	0,00	0,00000	0,000
R31	00	01	0,00	0,00000	0,000
R32	00	01	0,00	0,00000	0,000
R35	00	01	0,00	0,00000	0,000
R36	00	01	0,00	0,00000	0,000
R37	01	02	0,33	0,00028	0,333
R38	00	01	0,00	0,00000	0,000
R39	00	01	0,00	0,00000	0,000
R56	00	01	0,00	0,00000	0,000

S = número de sítios polimórficos; h = número de haplótipos; Hd = diversidade haplotípica; π = diversidade nucleotídica e k = número médio de diferenças nucleotídicas. Os códigos das localidades são os mesmos apresentados na Tabela 1.

O modelo de substituição nucleotídica que melhor explica o processo de evolução dos fragmentos obtidos para o gene KI-5 é TPM1uf+ I (KIMURA, 1981), contudo, na ausência deste modelo dentre as opções disponíveis nos programas utilizados, foi escolhido o modelo HKY +I (HASEGAWA; KISHINO; YANO, 1985). A taxa de substituição nucleotídica para o gene kl-5, estimada a partir de calibração secundária, é de 1,004 substituições/sítio/Myr.

4.2 Análise de variância molecular e teste de Mantel

4.2.1 Gene COI

O resultado da AMOVA global mostrou estruturação populacional entre as amostras de *D. serido* (Tabela 8).

Tabela 8. Análise da variância molecular (AMOVA) utilizando sequências do gene mitocondrial COI obtidas a partir de populações de *D. serido*.

	Fonte de variação	Componentes de variância	% de variação	<i>Fixation index</i>
AMOVA global	Entre populações	1,24 (Va)	54,05	$\Phi_{st} = 0,54$ ($p=0,00$)
	Dentro das populações	1,05 (Vb)	45,95	
AMOVA hierárquica NOR-LIT	Entre grupos	2,42 (Va)	66,12	$\Phi_{ct} = 0,66$ ($p=0,00$)
	Entre populações dentro dos grupos	0,18 (Vb)	5,04	$\Phi_{sc} = 0,15$ ($p=0,00$)
	Dentro das populações	1,05 (Vc)	28,84	$\Phi_{st} = 0,71$ ($p=0,00$)

*Valores estatisticamente significativos ($p<0,05$)

O teste de Mantel não apresentou correlação significativa entre a distância genética e a distância geográfica ($p=1,00$) para as populações analisadas, ou seja, a distância geográfica não explica a distância genética observada.

4.2.2 Gene α esterase-5

A AMOVA global, realizada a partir de sequências do gene nuclear $\alpha E5$, apresentou evidência de estruturação populacional (Tabela 9).

Tabela 9. Análise da variância molecular (AMOVA) utilizando sequências do gene α esterase-5 obtidas a partir de populações de *D. serido*.

	Fonte de variação	Componentes de variância	% de variação	<i>Fixation index</i>
AMOVA global	Entre populações	4,98 (Va)	66,75	$\Phi_{st} = 0,67$ ($p=0,00$)
	Dentro das populações	2,48 (Vb)	33,25	
AMOVA hierárquica NOR-LIT	Entre grupos	7,94 (Va)	68,7	$\Phi_{ct} = 0,69$ ($p=0,00$)
	Entre populações dentro dos grupos	1,13 (Vb)	9,82	$\Phi_{sc} = 0,31$ ($p=0,00$)
	Dentro das populações	2,48 (Vc)	21,48	$\Phi_{st} = 0,79$ ($p=0,00$)

*Valores estatisticamente significativos ($p<0,05$).

O teste de Mantel não apresentou correlação significativa entre a distância genética e a distância geográfica ($p=1,00$) para as populações analisadas, ou seja, a distância geográfica não explica a distância genética observada.

4.2.3 Gene GstD1

A AMOVA global utilizando sequências do gene nuclear GstD1 apresentou evidência para presença de estruturação populacional (Tabela 10).

Tabela 10. Análise da variância molecular (AMOVA) utilizando sequências do gene GstD1 obtidas a partir de populações de *D. serido*.

	Fonte de variação	Componentes de variância	% de variação	<i>Fixation index</i>
AMOVA global	Entre populações	0,70 (Va)	33,76	$\Phi_{st} = 0,34$ ($p=0,00$)
	Dentro das populações	1,37 (Vb)	66,24	
AMOVA hierárquica NOR-LIT	Entre grupos	1,35 (Va)	83,45	$\Phi_{ct} = 0,83$ ($p=0,00$)
	Entre populações dentro dos grupos	0,26 (Vb)	16,55	$\Phi_{sc} = 0,12$ ($p=0,00$)
	Dentro das populações	1,37 (Vc)	44,53	$\Phi_{st} = 0,55$ ($p=0,00$)

*Valores estatisticamente significativos ($p<0,05$).

O teste de Mantel não apresentou correlação significativa entre a distância genética e a distância geográfica ($p=1,00$) para as populações analisadas, ou seja, a distância geográfica não explica a distância genética observada.

4.2.4 Gene KI-5

A AMOVA global realizada com sequências do gene nuclear KI-5 mostrou que há estruturação genética entre as populações de *D. serido* (Tabela 11).

Tabela 11. Análise da variância molecular (AMOVA) utilizando sequências do gene *kl-5* obtidas a partir de populações de *D. serido*.

	Fonte de variação	Componentes de variância	% de variação	<i>Fixation index</i>
AMOVA global	Entre populações	0,59 (Va)	95,16	$\Phi_{st} = 0,95$ ($p=0,00$)
	Dentro das populações	0,03 (Vb)	4,84	
AMOVA hierárquica NOR-LIT	Entre grupos	0,53 (Va)	68,85	$\Phi_{ct} = 0,68$ ($p=0,00$)
	Entre populações dentro dos grupos	0,24 (Vb)	31,15	$\Phi_{sc} = 0,59$ ($p=0,00$)
	Dentro das populações	0,03 (Vc)	2,4	$\Phi_{st} = 0,98$ ($p=0,00$)

*Valores estatisticamente significativos ($p<0,05$).

O teste de Mantel não apresentou correlação significativa entre a distância genética e a distância geográfica ($p=1,00$) para as populações analisadas, ou seja, a distância geográfica não explica a distância genética observada.

4.3 Reconstrução filogenética

4.3.1 Gene COI

A árvore filogenética bayesiana gerada a partir dos dados de sequências parciais do gene mitocondrial COI revela que não há separação total em dois agrupamentos, como sugere nossa hipótese inicial (Figura 6). Alguns indivíduos do litoral, das seguintes populações: N31, N32, N35 e J52 estão presentes junto com os indivíduos do nordeste, enquanto alguns indivíduos do nordeste, das seguintes populações: N37 e N43 estão presentes junto com os indivíduos do litoral.

A

...

...

Figura 6 A. Relações filogenéticas estabelecidas por meio da análise de sequências parciais do gene COI para as populações de *Drosophila serido*. As reticências representam a continuidade da árvore cortada. Os números representam a probabilidade *a posteriori* dos clados. Os códigos das localidades são os mesmos apresentados na Tabela 1. **B.** Apresentação completa para melhor visualização.

4.3.2 Gene α esterase-5

A árvore filogenética bayesiana gerada a partir dos dados de sequências parciais do gene autossômico α esterase-5 revela a separação total de dois agrupamentos para *D.serido*, assim como sugere nossa hipótese inicial: litoral e nordeste, com probabilidade *a posteriori* igual a 1 (Figura 7).

Figura 7 A. Relações filogenéticas estabelecidas por meio da análise de sequências parciais do gene α esterase-5 para as populações de *Drosophila serido*. As reticências representam a continuidade da árvore cortada. Os números representam a probabilidade *a posteriori* dos clados. Os códigos das localidades são os mesmos apresentados na Tabela 1. A segunda figura (7A) nos apresenta a divisão entre nordeste e litoral (azul e vermelho, respectivamente). **B.** Apresentação completa para melhor visualização.

4.3.3 Gene GstD1

A árvore filogenética bayesiana gerada a partir dos dados de sequências parciais do gene autossômico GstD1 revela a separação total de dois agrupamentos para *D. serido*, assim como sugere nossa hipótese inicial: litoral e nordeste, com probabilidade *a posteriori* igual a 1 (Figura 8).

A

Phylogenetic tree showing relationships between various strains of *D. mojavensis*. The tree is rooted on the left and branches to the right. Bootstrap values are indicated at the nodes. A red box highlights a specific clade of strains.

Key bootstrap values and strain labels include:

- 0.72, 0.01, 0.07, 0.08, 0.44, 0.03, 0.01, 0.05, 0.04, 0.01, 0.03, 0.01, 0.03, 0.01, 0.04, 0.05, 0.02, 0.04
- Strain labels: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100

LITORAL

Figura 8 A. Relações filogenéticas estabelecidas por meio da análise de sequências parciais do gene GstD1 para as populações de *Drosophila serido*. As reticências representam a continuidade da árvore cortada. Os números representam a probabilidade *a posteriori* dos clados. Os códigos das localidades são os mesmos apresentados na Tabela 1. A segunda figura (8A) nos apresenta a divisão entre nordeste e litoral (azul e vermelho, respectivamente). **B.** Apresentação completa para melhor visualização.

4.3.4 Gene Kl-5

A árvore filogenética bayesiana gerada a partir dos dados de sequências parciais do gene sexual Kl-5 revela que não há separação total em dois agrupamentos, como sugere nossa hipótese inicial (Figura 9). Os indivíduos do litoral de SP, sendo as seguintes populações: N2, H49, Ilha, N19 e N29, são encontrados em totalidade com os indivíduos do nordeste. Não há indivíduos do nordeste no clado do litoral.

B

Figura 9 A. Relações filogenéticas estabelecidas por meio da análise de sequências parciais do gene KI-5 para as populações de *Drosophila serido*. As reticências representam a continuidade da árvore cortada. Os números representam a probabilidade *a posteriori* dos clados. Os códigos das localidades são os mesmos apresentados na Tabela 1. **B.** Apresentação completa para melhor visualização.

4.3.5 Genes analisados em conjunto

A árvore filogenética bayesiana gerada a partir dos dados de sequências parciais dos genes nucleares e mitocondrial, analisados em conjunto, revela a separação total de dois agrupamentos para *D. serido*, assim como sugere nossa hipótese inicial: litoral e nordeste, com probabilidade *a posteriori* igual a 1 (Figura 10).

Figura 10 A. Análise filogenética das populações estudadas da espécie *Drosophila serido* para todos os marcadores. As reticências representam a continuidade da árvore cortada. Os números representam a probabilidade posteriori dos clados. Os códigos das localidades são os mesmos apresentados na Tabela 1. A segunda figura (10A) nos apresenta a divisão entre nordeste e litoral (azul e vermelho, respectivamente). **B.** Apresentação completa para melhor visualização.

4.4 Análise da estrutura populacional

4.4.1 Gene COI

A análise de estruturação populacional realizada por meio do programa BAPS, para o marcador molecular COI, resultou em duas populações (Figura 11).

Como resultado dessa análise, as populações foram parcialmente separadas em nordeste e litoral, porém alguns haplótipos ainda são compartilhados entre as populações. Todos os indivíduos do nordeste estão marcados pela cor vermelha, enquanto os indivíduos do litoral estão divididos entre vermelho e verde.

Figura 11. Estrutura populacional para *D. serido* realizada por meio do programa BAPS, para um modelo populacional com $k=2$ para o marcador molecular COI. As linhas verticais em preto representam os indivíduos amostrados. Cada cor representa um *cluster*. A ordem dos indivíduos amostrados está de acordo com a região geográfica. O quadrado preto representa os indivíduos do litoral.

4.4.2 Gene α esterase-5

A análise de estruturação populacional realizada por meio do programa BAPS, para o marcador molecular α esterase-5, resultou em três populações (Figura 12).

Como resultado dessa análise, as populações foram separadas de acordo com nossa hipótese inicial, entre litoral e nordeste. Contudo, todos os indivíduos do litoral estão marcados pela cor azul, enquanto os indivíduos do nordeste estão divididos entre verde e vermelho.

Figura 12. Estrutura populacional para *D. serido* realizada por meio do programa BAPS, para um modelo populacional com $k=3$ para o marcador molecular α esterase-5. As linhas verticais em preto representam os indivíduos amostrados. Cada cor representa um *cluster*. A ordem dos

indivíduos amostrados está de acordo com a região geográfica. O quadrado preto representa os indivíduos do litoral.

4.4.3 Gene GstD1

A análise de estruturação populacional realizada por meio do programa BAPS, para o marcador molecular GstD1, resultou em duas populações (Figura 13).

Como resultado dessa análise, as populações foram separadas em duas populações e validando nossa hipótese inicial de separação: litoral e nordeste. As populações do litoral estão marcadas em vermelho e as populações do nordeste estão marcadas em verde. Nessa análise, não temos mistura entre as populações.

Figura 13. Estrutura populacional para *D. serido* realizada por meio do programa BAPS, para um modelo populacional com $k=2$ para o marcador molecular GstD1. As linhas verticais em preto representam os indivíduos amostrados. Cada cor representa um *cluster*. A ordem dos indivíduos amostrados está de acordo com a região geográfica. O quadrado preto representa os indivíduos do litoral.

4.4.4 Gene KI-5

A análise de estruturação populacional realizada por meio do programa BAPS, para o marcador molecular KI-5, resultou em cinco populações (Figura 14).

Como resultado dessa análise, as populações foram parcialmente separadas em nordeste e litoral, porém alguns haplótipos ainda são compartilhados entre as populações. As cores amarelo e azul possuem indivíduos de ambas as populações. A cor vermelha possui indivíduos apenas do litoral. O quarto e quinto cluster está marcado em rosa e verde, respectivamente, contendo apenas um indivíduo (N67-2 e N9-1).

Figura 14. Estrutura populacional para *D. serido* realizada por meio do programa BAPS, para um modelo populacional com $k=5$ para o marcador molecular KI-5. As linhas verticais em preto representam os indivíduos amostrados. Cada cor representa um *cluster*. A ordem dos indivíduos amostrados está de acordo com a região geográfica. O quadrado preto representa os indivíduos do litoral.

4.5 Análise de componentes principais

A análise de componentes principais, que tem como objetivo condensar a informação contida em várias variáveis originais em um conjunto menor de variáveis estatísticas (componentes) com uma perda mínima de informação teve como melhor resultado a somatória dos eixos 1 e 2, sendo a porcentagem cumulativa desses eixos igual a 63,13% (Figura 15).

Nessa análise, observamos que as populações do litoral estão se distanciando das populações do nordeste, apoiando a nossa hipótese inicial.

Figura 15. Agrupamento populacional para *Drosophila serido*, utilizando uma matriz de haplótipos para todos os genes e indivíduos estudados. Os códigos das localidades são os mesmos apresentados na Tabela 1. As cores das localidades são as mesmas cores utilizadas pelo programa Haplotype Viewer, sendo vermelho para as populações do litoral e azul para as populações do nordeste.

4.6 Redes haplotípicas

4.6.1 Gene COI

Foram analisadas 211 sequências para o marcador molecular COI. Por meio do programa Haplotype Viewer, foram gerados 52 haplótipos para estas sequências (Figura 16).

Os haplótipos marcados em azul são os indivíduos do nordeste, enquanto os haplótipos marcados em vermelho são os indivíduos do litoral. Os haplótipos que visualizamos as duas

cores, correspondem aos indivíduos do litoral que compartilham o mesmo haplótipo com os indivíduos do nordeste.

Figura 16. Rede haplotípica gerada por meio do programa Haplotype Viewer para o marcador molecular COI. As populações do nordeste estão marcadas com a cor azul, enquanto as populações do litoral estão marcadas com a cor vermelha. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujo tamanhos são proporcionais aos números de indivíduos que os possuem.

4.6.2 Gene α esterase-5

Foram analisadas 136 sequências para o marcador molecular α esterase-5. Por meio do programa Haplotype Viewer, foram gerados 97 haplótipos para estas sequências (Figura 17).

Nessa análise, os haplótipos dos indivíduos foram separados em duas populações: nordeste e litoral, apoiando nossa hipótese inicial. Não há compartilhamento de haplótipos entre os indivíduos dessas populações.

Figura 17. Rede haplotípica gerada por meio do programa Haplotype Viewer para o marcador molecular α esterase-5. As populações do nordeste estão marcadas com a cor azul, enquanto as populações do litoral estão marcadas com a cor vermelha. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujo tamanhos são proporcionais aos números de indivíduos que os possuem.

4.6.3 Gene GstD1

Foram analisadas 147 sequências para o marcador molecular GstD1. Por meio do programa Haplotype Viewer, foram gerados 45 haplótipos para estas sequências (Figura 18).

Nessa análise, os haplótipos dos indivíduos também foram separados em duas populações: nordeste e litoral, apoiando nossa hipótese inicial. Não há compartilhamento de haplótipos entre os indivíduos dessas populações.

Figura 18. Rede haplotípica gerada por meio do programa Haplotype Viewer para o marcador molecular GstD1. As populações do nordeste estão marcadas com a cor azul, enquanto as populações do litoral estão marcadas com a cor vermelha. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujo tamanhos são proporcionais aos números de indivíduos que os possuem.

4.6.4 Gene KI-5

Foram analisadas 194 sequências para o marcador molecular KI-5. Por meio do programa Haplotype Viewer, foram gerados 5 haplótipos para estas sequências (Figura 19).

Assim como no marcador COI, alguns indivíduos do litoral compartilham haplótipos com os indivíduos do nordeste.

Figura 19. Rede haplotípica gerada por meio do programa Haplotype Viewer para o marcador molecular KI-5. As populações do nordeste estão marcadas com a cor azul, enquanto as

populações do litoral estão marcadas com a cor vermelha. Cada linha da rede representa um passo mutacional. Haplótipos são representados por círculos, cujo tamanhos são proporcionais aos números de indivíduos que os possuem.

4.7 Discussão

Drosophila serido é uma espécie cactofílica, endêmica da América do Sul, cujo suas populações foram hipoteticamente divididas em dois grupos: populações do litoral e populações do nordeste.

Para as populações do Nordeste, embora haja diversidade de biomas nessa região, que inclui: 1- Mata Atlântica, 2- Agreste e 3- Caatinga (PRADO, 2003), os dados sugerem ausência de estruturação populacional. Duas hipóteses são sugeridas para compreensão da homogeneidade dessas populações: 1- a estabilidade climática na Caatinga e 2- a geomorfologia da região (WERNECK et al, 2011).

Mapas com projeções paleoclimáticas, realizadas para táxons de ocorrência na Caatinga, apresentam grandes áreas de estabilidade climática na região Nordeste, ao longo das flutuações climáticas do Quaternário, sendo que apenas as ecorregiões Raso da Catarina, Planalto da Borborema e porção sul da Depressão Sertaneja apresentam-se nos mapas de predição como áreas de instabilidade climática (WERNECK et al, 2011). Com relação à geomorfologia da região Nordeste, ela apresenta complexidade topográfica relativamente reduzida, possuindo poucas formações como as Chapadas do Araripe, Apodi, Borborema e a Serra do Espinhaço. Dessa forma, uma vez que a quantidade de barreiras é reduzida, o fluxo gênico entre as populações é favorecido (WERNECK et al, 2012b).

O sistema hídrico desta região parece constituir barreira ao fluxo gênico entre populações de *D. serido*. Contudo, há evidências de que, dependendo do grupo taxonômico estudado, a barreira dos rios pode ou não atuar ao fluxo gênico. Para a lagartixa *Gymnodactylus darwinii* (PELLEGRINO et al, 2005) análises genéticas mostraram estruturação populacional consoante com esses rios, porém, para o jacarandá-da-Bahia, *Dalbergia nigra*, a análise de estruturação genética apresentou agrupamentos ao norte e ao sul do Rio Jequitinhonha

(RIBEIRO et al, 2011). Para a espécie *D. serido*, estudada no presente trabalho, as análises também apontaram agrupamentos ao norte e ao sul do Rio Jequitinhonha, evidenciando barreira ao fluxo gênico dessa espécie.

Estudos anteriores (LAVAGNINI-PIZZO, 2015) revelam que populações do Nordeste e vale do rio Jequitinhonha possuem haplótipos compartilhados. Estudos analisando filogeograficamente todas essas populações, usando sequências parciais do gene mitocondrial COI, sugerem que o centro de dispersão desta espécie é o interior do estado da Bahia (FRANCO; MANFRIN, 2013).

As populações do litoral, ao longo da costa Atlântica, mostram complexidade evolutiva. Apesar de não estar relacionada diretamente ao bioma Mata Atlântica, sua ocorrência no litoral se dá na área de restinga, que estão sujeitas às influências da seca, salinidade e quantidade reduzida de nutrientes (LACERDA; ARAUJO; MACIEL, 1993; SCARANO, 2002). Desta forma, ela pode ter influenciado na dinâmica evolutiva desta espécie nessa região.

As populações com distribuição na costa Atlântica apresentam, de forma geral, uma divisão norte-sul, sendo que a posição geográfica desta divisão pode variar de acordo com o grupo taxonômico e a barreira considerada (CARNAVAL et al, 2009; FITZPATRICK et al, 2009; RESENDE et al, 2010; TONINI et al, 2013). Neste trabalho, a divisão norte-sul também foi observada em *D. serido*. Para a espécie, a descontinuidade genética ocorre acima da localidade de Cabralia-BA (N35), sendo esta, geograficamente, a primeira localidade agrupada em litoral/sul. Este fato já havia sido sugerido anteriormente por análises de hidrocarbonetos de cutícula (OLIVEIRA et al, 2011), inversões cromossômicas polimórficas (RUIZ et al, 2000; TOSI, SENE, 1989), diversidade haplotípica mitocondrial (FRANCO; MANFRIN, 2013; MANFRIN, BRITO, SENE, 2001) e análises cariotípicas (BAIMAI, SENE, PEREIRA, 1983).

Estudos têm mostrado que as flutuações climáticas do Quaternário, associadas às flutuações no nível do mar, também influenciaram as áreas de ocorrência dos manguezais, favorecendo eventos de expansão e retração de sua vegetação (COHEN et al, 2008, 2009; FRANÇA et al, 2013). Desta forma, considerando a possibilidade de mudanças nas áreas de ocorrência dos manguezais, devido aos paleoclimas do Quaternário, e as incompatibilidades entre as características do solo para estabelecimento de manguezais e populações de

Cactaceae, é possível inferir que o Rio Jequitinhonha pode ter atuado como barreira à dispersão e fluxo gênico dos cactos, assim como das populações de moscas cactófilas.

A análise da variação morfológica de edeagos nessas populações mostrou a presença de dois grupos distintos na costa atlântica: um grupo formado pelas populações de Bertioga – SP e Penha – SC e o outro grupo com as populações de São Sebastião – SP e Arraial do Cabo – RJ (FRANCO et al, 2008). Análises filogeográficas, com base em sequências parciais do gene mitocondrial COI, mostraram que as populações do litoral formam um grupo diferente das populações do nordeste, juntando as populações do estado de São Paulo com as populações do litoral sul do estado da Bahia, e as populações do litoral do estado de Santa Catarina com as populações da caatinga (MORALES, 2005; KOKUDAI et al, 2013). Quanto à estrutura populacional ao longo do litoral atlântico, há divergência entre os marcadores. Para o gene mitocondrial COI, as populações estariam divididas em dois grupamentos, sendo que as populações de Santa Catarina estariam geneticamente isoladas das demais, formando o grupo “Sul”, enquanto para o gene nuclear *period* as populações desde o sul da Bahia até Santa Catarina formam um único agrupamento, denominado “Sudeste” (LAVAGNINI-PIZZO, 2015).

No presente trabalho, a divisão norte-sul é mostrada com nitidez para os marcadores α esterase-5 e GstD1. Esses genes possuem ação antioxidante e estudos anteriores relacionando esses marcadores e linhagens evolutivas em espécies de *Drosophila* mostraram ligação direta com cactos (MATZKIN, 2008; SANTOS et al, 2009). Os dados sugerem que em duas populações de *D. mojavensis*, o gene GstD1 passou por um período de evolução adaptativa de aminoácidos. Outras análises indicaram que das sete fixações de aminoácidos que ocorreram na linhagem de *D. mojavensis*, duas delas ocorrem no sítio ativo, tendo um efeito significativo na especificidade do substrato e na adaptação a hospedeiros alternativos de cactos (MATZKIN, 2008). Outros estudos analisando o gene α esterase-5 mostraram que a família desse gene possui altas taxas evolutivas e baixa identidade de aminoácidos, provavelmente devido à sua diversificação na sequência de DNA e na função bioquímica, em relação às mudanças ecológicas e comportamentais das espécies hospedeiras (SANTOS et al, 2009). Com isso, relacionamos que as variações genéticas que foram encontradas para a espécie *D. serido* para esses dois

marcadores podem ser variações para adaptação local, em relação aos cactos hospedeiros, devido às diferenças na diversidade de espécies de cactos na Caatinga e litoral atlântico.

Para os outros marcadores, a divisão entre norte-sul também é amostrada, porém alguns haplótipos ainda são compartilhados entre indivíduos de ambas as populações. A ausência de recombinação, elevada taxa evolutiva e redução do tamanho efetivo populacional são as condições que moldaram os genes COI (AVISE, 2000, 2009; BRITO, EDWARDS, 2009; HICKERSON et al, 2010) e kl-5 (CARVALHO, 2002; DUPIM et al, 2018). Contudo, são marcadores muito informativos e úteis para estudos intra e interespecíficos quando combinados com outros marcadores (YANG, RANNALA, 2010; DUPUIS et al, 2012).

O fato de esses genes conseguirem resgatar as relações evolutivas entre populações e entre espécies é o que faz com que apareçam as duas linhagens evolutivas para a espécie *D. serido*, ainda que recuperem o compartilhamento de haplótipos entre indivíduos da mesma espécie.

5 Conclusão

Com base nos resultados apresentados neste trabalho, concluimos que:

- O Rio Jequitinhonha parece ser uma barreira ao fluxo gênico para as espécies de *D. serido*;
- Há uma divisão norte-sul para os marcadores utilizados, que foram chamadas de “populações do nordeste” e “populações do litoral”;
- Essa divisão ocorre acima de Cabrália, na Bahia;
- Existem duas linhagens evolutivas para a espécie *D. serido*.

6 Referências

AVISE, J. C. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. **Annual Review of Ecology, Evolution and Systematics**, v. 18, p. 489-522, 1987.

AVISE, J. C. **Molecular markers, natural history and evolution**. New York: Chapman and Hall, 1994.

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

AVISE, J. C. **Phylogeography: The history and formation of species**. Cambridge: Harvard University Press, 2000.

BAIMAI, V.; SENE, F. M.; PEREIRA, M. A. O. R. Heterochromatin and karyotypic differentiation of some neotropical cactus-breeding species of the *Drosophila repleta* species group. **Genetica**, v. 60, n. 2, p. 81-92, 1983.

BARKER, J. S.; STARMER, W. T. **Ecological Genetics and Evolution. The cactus-yeast-Drosophila Model System**. Austrália: Academic Press, 1982.

BARRIOS-LEAL, D. Y.; ROCHA, J. N.; MANFRIN, M. H. Genetics and Distribution Modeling: The Demographic History of the Cactophilic *Drosophila buzzatti* Species Cluster in Open Areas of South America. **Journal of Heredity**, v. 110, n. 1, p. 22-33, 2018.

BONATELLI, I. A. S. et al. Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. **Molecular Ecology**, v. 23, p. 3044-3063, 2014.

BONO, J. M. et al. Molecular Evolution and population genetics of two *Drosophila mettleri* cytochrome P450 genes involved in host plant utilization. **Molecular Ecology**, v. 17, p. 3211-3221, 2008.

BOUCKAERT, R. et al. BEAST2: A software platform for bayesian evolutionary analysis. **PLOS**, 2014.

BRITO, P. H.; EDWARDS, S. V. Multilocus phylogeography and phylogenetics using sequence-based markers. **Genetica**, v. 135, p. 439-455, 2009.

BRUEN, T. C.; PHILIPPE, H.; BRYANT, D. A simple and robust statistical test for detecting the presence of recombination. **Genetics**, v. 172, p. 2665-2681, 2006.

CARNAVAL, A. C. et al. Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. **Science**, v. 323, p. 785-789, 2009.

CARVALHO, A. B. Origin and Evolution of the *Drosophila* Y chromosome. **Current Opinion in Genetics and Development**, v. 12, p. 664-668, 2002.

CHIPPINDALE, A. K.; RICE, W. R. Y chromosome polymorphism is a strong determinant of male fitness in *Drosophila melanogaster*. **PNAS**, v. 98, n. 10, p. 5677-5682, 2001.

CLEMENT, M.; POSADA, D.; CRANDALL, K. A. TCS: a computer program to estimate gene genealogies. **Molecular Ecology**, v. 9, p. 1657-1659, 2000.

COHEN, M. C. L. et al. Impact of sea-level and climatic changes on the Amazon coastal wetlands during the late Holocene. **Vegetation History and Archaeobotany**, v. 18, p. 425-439, 2009.

COHEN, M. C. L. Wetland dynamics of Marajó Island, northern Brazil, during the last 1000 years. **Catena**, v. 76, p. 70-77, 2008.

CORANDER, J. et al. BAPS: Bayesian Analysis of Population Structure. Department of Mathematics and statistics. **University of Helsinki**. 2013.

CUZZUOL, G. R. F.; CAMPOS, A. Aspectos nutricionais na vegetação de manguezal do estuário do Rio Mucuri. **Revista Brasileira de Botânica**, v. 24, n. 2, p. 227-234, 2001.

DUPIM, E. G. et al. An investigation of Y chromosome incorporations in 400 species of *Drosophila* and related genera. **Genetics**, v. 14, 2018.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32, n. 5, p. 1792-1797, 2004.

ERSTS, P. J. Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Disponível em http://biodiversityinformatics.amnh.org/open_source/gdmg.

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**, v. 131, n. 2, p. 479-491, 1992.

FITZPATRICK, S. W. et al. Geographical variation in genetic structure of na Atlantic Coastal Forest frog reveals regional differences in habitat stability. **Molecular Ecology**, v. 18, p. 2877-2896, 2009.

FRANÇA, M. C. et al. Mangrove vegetation changes on Holocene terraces of the Doce River, southeastern Brasil. **Catena**, v. 110, p. 59-69, 2013.

FRANCO, F. F. et al. Phenotypic variation of the aedeagus of *Drosophila serido* Vilela & Sene (Diptera: Drosophilidae). **Neotropical Entomology**, v. 37, n. 5, p. 558-563, 2008.

FRANCO, F. F.; MANFRIN, M. H. Recent demographic history of cactophilic *Drosophila* species can be related to the Quaternary paleoclimatic changes in South America. **Journal of Biogeography**, v. 40, p. 142-154, 2013.

FRANCO, F.F.; BERNARDI, E. C. C. S.; SENE, F. M.; HASSON, E. R.; MANFRIN, M. H. 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, 2010.

FRANK, M. R.; FOGLEMAN, J. C. Involvement of cytochrome P450 in host-plant utilization by Sonoran Desert *Drosophila*. **PNAS**, v. 89, p. 11998-12002, 1992.

GEPNER, J.; HAYS, T. S. A fertility region on the Y chromosome of *Drosophila melanogaster* encodes a dynein microtubule motor. **PNAS**, v. 90, p. 11132-11136, 1993.

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.

HASEGAWA, M.; KISHINO, H.; YANO, T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. **Journal of Molecular Evolution**, v. 22, p. 160-174, 1985.

HICKERSON, M. J. et al. Phylogeography's past, present, and future: 10 years after Avise, 2000. **Molecular Phylogenetics and Evolution**, v. 54, p. 291-301, 2010.

HUSON, D. H.; BRYANT, D. Application of phylogenetic networks in evolutionary studies. **Molecular Biology and Evolution**, v. 23, n. 2, p. 254-267, 2006.

JUDD, W. S. et al. **Plant Systematics: A phylogenetic approach**. Universidade de Michigan: Editora Freeman, 2008.

KIMURA, M. Estimation of evolutionary distances between homologous nucleotide sequences. **Proc Natl Acad Sci USA**, v. 78, n. 1, p. 454-458, 1981.

KOERICH, L. B. et al. Low conservation of gene content in the *Drosophila* Y chromosome. **Nature**, v. 456, p. 949-951, 2008.

KOKUDAI, C. B. S.; SENE, F. M.; MANFRIN, M. H. Sympatry and Asymmetric Introgression Between the Cactophilic Species *Drosophila serido* e *Drosophila antonietae* (Diptera: Drosophilidae). **Ecology and Population Biology**, v.104, n.3, p.434-442, 2013.

LACERDA, L. D.; ARAUJO, D. S. D.; MACIEL, N. C. Dry coastal ecosystem of the tropical Brazilian coast. In: VAN DER MAAREL, E. (Ed.). **Dry coastal ecosystems: Africa, America, Asia and Oceania**. Amsterdam: Elsevier, 1993, p. 477-493.

LAVAGNINI-PIZZO, T. C. **História evolutiva de *Drosophila serido* ("cluster" *Drosophila buzzatii*)**. 2015. 161 f. Tese (Doutorado em Ciências) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2015.

LIBRADO, P.; ROZAS, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**, v. 25, p. 1451-1452, 2009.

MANFRIN, M. H.; DE BRITO, R. O. A.; SENE, F. M. Systematics and Evolution of the *Drosophila buzzatti* (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, p.57-75, 2006.

MARKOW, T. A.; O'GRADY, P. M. Evolutionary genetics of reproductive behavior in *Drosophila*: connecting the dots. **Annual review of genetics**, v.39, p. 263-291, 2005.

MATZKIN, L. M. et al. Functional genomics of cactus host shifts in *Drosophila mojavensis*. **Molecular Ecology**, v. 15, n. 14, p. 4635-4643, 2006.

MATZKIN, L. M. Population transcriptomics of cactus host shifts in *Drosophila mojavensis*. **Molecular Ecology**, v. 21, n. 10, p. 2428-2439, 2012.

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. 1073-1083, 2008.

MAYR, E. **Populações, espécies e evolução**. Ed. Nacional. São Paulo: Ed. Da Universidade de São Paulo, 1977.

MAYR, E. **Systematics and the Origin of Species**. Columbia University Press, 1942.

MORALES, A. C. **Análise histórico-evolutiva de populações das espécies cactofílicas *Drosophila serido* e *Drosophila antonietae* (Diptera, Drosophilidae)**. 2005. 110 f. Tese (Doutorado em Ciências) – Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2005.

OLIVEIRA, C. C. et al. Variations on a theme: diversification of cuticular hydrocarbons in a clade of cactophilic *Drosophila*. **BMC Evolutionary Biology**, v. 11, p. 179, 2011.

OLIVEIRA, M. T.; GARESSE, R.; KAGUNI, L. S. Animal models of mitochondrial DNA transactions in disease and ageing. **Experimental Gerontology**, v. 45, n. 7-8, p. 489-502, 2010.

PEAKALL, R.; SMOUSE, P. E. 2012. GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. **Bioinformatics**, 28(19): 2537–2539.

PELLEGRINO, K. C. M. et al. Phylogeography and species limits in the *Gymnodactylus darwinii* complex (Gekkonidae, Squamata): genetic structure coincides with river systems in the Brazilian Atlantic Forest. **Biological Journal of the Linnean Society**, v. 85, n. 1, p. 13-26, 2005.

POSADA, D.; BUCKLEY, T. R. Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian Approaches over Likelihood Ratio Tests. **Systematic Biology**, v. 53, n. 5, p. 793-808, 2004.

POSADA, D.; CRANDALL, K. A. MODELTEST: testing the model of DNA substitution. **Bioinformatics**, v. 14, n. 9, p. 817-818, 1998.

PRADO, D. E. **As caatingas da América do Sul**. In: Leal, I. R.; Tabarelli, M.; Silva, J. M. C. (Eds) Ecologia e Conservação da Caatinga. Ed. Universitária da UFPE, Recife, p. 3-74, 2003.

RAMBAUT, A. et al. Tracer version 1.6, Disponível em: <http://beast.bio.ed.ac.uk/Tracer>.2014.

RAMBAUT, A. FigTree: Tree figure drawing tool. Disponível em: <http://tree.bio.ed.ac.uk/software/figtree.2014>.

RESENDE, H. C. et al. Pliocene and Pleistocene events shaping the genetic diversity within the central corridor of the Brazilian Atlantic Forest. **Biological Journal of the Linnean Society**, v. 101, p. 949-960, 2010.

RIBEIRO, R. A. et al. Phylogeography of the endangered rosewood *Dalbergia nigra* (Fabaceae): insights into the evolutionary history and conservation of the Brazilian Atlantic Forest. **Heredity**, v. 106, n. 1, p. 46-57, 2011.

ROBIN, G. C. Q. et al. Reconstructing the diversification of a α -Esterases: comparing the gene clusters of *Drosophila buzzatii* and *D. melanogaster*. **Journal of Molecular Evolution**, v. 51, p. 149-160, 2000.

RUIZ, A.; WASSERMAN, M. Evolutionary cytogenetics of the *Drosophila buzzatii* species complex. **Heredity**, v. 70, n. 6, p. 582-596, 1993.

RUIZ, G. M.; FOFONOFF, P. W.; CARLTON, J. T.; WONHAM, M. J.; HINES, A. H. Invasion of Coastal Marine Communities in North America: Apparent Patterns, Processes, and Biases. **Annual Review of Ecology and Systematics**, vol.31, p.481-531, 2000.

SANTOS, M. H.; HASSON, E.; SENE, F. M.; MANFRIN, M. H. *Esterase-5* gene sequences from the *Drosophila buzzatti* cluster. **DIS**, v.92, p.44-56, 2009.

SCARANO, F. R. 2002. Structure, function and floristic relationships of plant communities in stressful habitats marginal to the Brazilian atlantic rain forest. **Annals of Botany** 90:517-524.

SILVA, A. F. G.; SENE, F. M. Morphological geographic variability in *Drosophila serido* (Diptera, Drosophilidae). **Revista Brasileira de Entomologia**, v. 35, p. 455-468, 1991.

TAMURA, K.; NEI, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. **Molecular Biology and Evolution**, v. 10, n. 3, p. 512-526, 1993.

TEMPLETON, A. R. **Genética de Populações e Teoria Microevolutiva**. Sociedade Brasileira de Genética, 2011.

TEMPLETON, A. R. Statistical phylogeography: methods of evaluating and minimizing inference errors. **Molecular Ecology**, v. 13, p. 789-809, 2004.

THOMPSON, J.; HIGGINS, D.; GIBSON, T. 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. 4673-4680, 1994.

TONINI, J. F. R.; COSTA, L. P.; CARNAVAL, A. C. Phylogeographic structure is Strong in the Atlantic Forest; predictive power or correlative palaeodistribution models, not always. **Journal of Zoological Systematics and Evolutionary Research**, v. 51, n. 2, p. 114-121, 2013.

TOSI, D.; SENE, F. M. Further studies on chromosomal variability in *Drosophila serido* (Diptera, Drosophilidae). **Revista Brasileira de Genética**, v. 1, p. 729-746, 1989.

VILELA, C.R.; SENE, F.M. 1977. Two new neotropical species of the *repleta* group of the Genus *Drosophila* (Diptera, Drosophilidae). **Revista Brasileira de Entomologia**, São Paulo, 30: 295-299.

WASSERMAN, M.; RICHARDSON, R. H. Evolution of Brazilian *Drosophila mulleri* complex species. **Journal of Heredity**, v. 78, n. 5, p. 282-286, 1987.

WERNECK, F. P. et al. Deep diversification and long-term persistence in the South American 'dry diagonal': integrating continente-wide phylogeography and distribution modeling of geckos. **Evolution**, v. 66, n. 10, p. 3014-3034, 2012b.

WERNECK, F. P. et al. Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on palaeodistribution modelling and palynological evidence. **Global Ecology and Biogeography**, v. 20, v. 2, p. 272-288, 2011.