

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCÊNCIAS
BACHARELADO EM CIÊNCIAS BIOLÓGICAS

PEDRO DORNELLES DE CASTRO

**CARACTERIZAÇÃO GENÔMICA DO GENE *ODYSSEUS* E ELEMENTOS
TRANSPONÍVEIS ASSOCIADOS NO SUBGRUPO *WILLISTONI* DE *DROSOPHILA*
(DIPTERA, DROSOPHILIDAE)**

Porto Alegre

2022

PEDRO DORNELLES DE CASTRO

**CARACTERIZAÇÃO GENÔMICA DO GENE *ODYSSEUS* E ELEMENTOS
TRANSPONÍVEIS ASSOCIADOS NO SUBGRUPO *WILLISTONI* DE *DROSOPHILA*
(DIPTERA, DROSOPHILIDAE)**

Trabalho de Conclusão de curso apresentado
como requisito parcial para obtenção do título de
Bacharel em Ciências Biológicas na
Universidade Federal do Rio Grande do Sul.

Orientadora: Profa. Dra. Maríndia Deprá

Coorientador: Me. Henrique da Rocha Moreira
Antoniolli

Porto Alegre

2022

CIP - Catalogação na Publicação

de Castro, Pedro Dornelles
CARACTERIZAÇÃO GENÔMICA DO GENE ODYSSEUS E
ELEMENTOS TRANSPONÍVEIS ASSOCIADOS NO SUBGRUPO
WILLISTONI DE DROSOPHILA (DIPTERA, DROSOPHILIDAE) /
Pedro Dornelles de Castro. -- 2022.
22 f.
Orientadora: Maríndia Deprá.

Coorientador: Henrique Antonioli.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Instituto
de Biociências, Bacharelado em Ciências Biológicas,
Porto Alegre, BR-RS, 2022.

1. Evolução. 2. Especiação. 3. Odysseus. 4.
Drosophila. I. Deprá, Maríndia, orient. II.
Antonioli, Henrique, coorient. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

BANCA EXAMINADORA:

Profa. Dra. Maríndia Deprá
Instituto de Biociências - UFRGS

Me. Henrique da Rocha Moreira Antonioli
Instituto de Biociências - UFRGS

Profa. Dra. Andréia Carina Turchetto Zolet
Instituto de Biociências - UFRGS

Me. William Vilas Boas Nunes
UNESP - Câmpus de São José do Rio Preto

RESUMO

A esterilidade em machos híbridos é uma das principais barreiras de isolamento reprodutivo pós-zigóticas a surgir durante o processo de especiação. No entanto, as bases genéticas da esterilidade híbrida ainda são pouco conhecidas. Neste sentido, *Odysseus (OdsH)* foi o primeiro gene a ser caracterizado como gene de especiação, devido ao seu papel na diminuição de aptidão em machos híbridos em espécies irmãs do clado da *Drosophila simulans*, espécie de origem afrotropical pertencente ao grupo *melanogaster*. O gene *OdsH* surgiu da duplicação do gene *unc-4* e possivelmente evoluiu a taxas mais aceleradas do que sua duplicata nessas espécies por meio de seleção positiva. De origem Neotropical, o subgrupo *willistoni* compreende seis espécies crípticas, as quais se apresentam em distintos níveis taxonômicos e com diferentes graus de isolamento reprodutivo, fornecendo excelentes organismos modelo para o estudo da especiação. Para verificar o status de assinatura evolutiva de *OdsH* nessas espécies irmãs, foi feita a caracterização genômica da região que compreende esses genes e análise de pressão seletiva de ortólogos *OdsH* no subgrupo *willistoni*, além da busca por elementos repetitivos. Os valores de dN/dS se mostraram bastante heterogêneos entre as espécies, entretanto, não se encontrou sinal de seleção positiva. Foi detectada a presença dos transposons *Jockey* e *Mariner* no primeiro íntron em duas subespécies de *D. paulistorum*, e a ausência de elementos repetitivos nas outras espécies. A análise de expressão transcricional e dos híbridos gerados fornecerão mais insights sobre o papel de *OdsH*-transposon no processo de especiação das subespécies de *D. paulistorum*.

Palavras-chave: assinatura evolutiva; *Drosophila paulistorum*; especiação; *jockey*; *mariner*.

ABSTRACT

Sterility in hybrid male is one of the main post-zygotic barriers to emerge during the speciation process. However, the genetic basis of hybrid sterility is still poorly understood. In this sense, *Odysseus* (*OdsH*) was the first gene to be characterized as a speciation gene, due to its role in decreasing fitness in hybrid males from sibling species of the *Drosophila simulans* clade, which belong to the afro-tropical *melanogaster* group. The *OdsH* gene arose from the duplication of the *unc-4* gene and is believed to have evolved at faster rates than its duplicate in these species, through positive selection. The neotropical *willistoni* subgroup comprises six cryptic species, which present distinct taxonomic levels and degrees of reproductive isolation. Thus, these species are excellent model organisms to study speciation. To verify the evolutionary signature status of *OdsH* in these sibling species, genomic characterization and selective pressure analysis of *OdsH* orthologs in the *willistoni* subgroup were performed, including the search of repetitive elements. The values of dN/dS were very heterogeneous, however, there was no sign of positive selection. The presence of *Jockey* and *Mariner* transposons was detected in the first intron in two *D. paulistorum* subspecies, while repetitive sequences were absent in this region in the other species. Further analyses of transcriptional expression and hybrids produced in this study will give insights on the role of *OdsH*-transposon in the process of speciation in *D. paulistorum* subspecies.

Keywords: *Drosophila paulistorum*; evolutionary signature; *jockey*; *mariner*; speciation.

SUMÁRIO

1. INTRODUÇÃO	8
2. MÉTODOS	10
2.1. Busca in silico e anotação dos genes unc-4 e OdsH	10
2.2. Teste de variação de dN/dS entre linhagens e testes de seleção natural	10
2.3. Identificação da linhagem 14030-0771.06, cruzamentos e a geração de híbridos	11
3. RESULTADOS	13
3.1. Ocorrência de unc-4 e OdsH no subgrupo willistoni	13
3.2. Elementos transponíveis associados ao OdsH	13
3.3. Testes de seleção	14
3.4. Testes de cruzamentos e identificação da linhagem 14030-0771.06	17
4. DISCUSSÃO	18
5. REFERÊNCIAS	20

1. INTRODUÇÃO

A especiação é um dos temas centrais da biologia evolutiva. Tal processo tipicamente envolve o surgimento de barreiras pré e pós-zigóticas entre populações, o que permite a manutenção da distinção genética e fenotípica dessas populações em proximidade geográfica (SEEHAUSEN *et al.*, 2014). Estudar as bases genéticas subjacentes a essas barreiras reprodutivas, e o subsequente isolamento reprodutivo das populações, é importante para termos um melhor entendimento do processo de especiação e da diversidade biológica.

Nesse sentido, vários estudos têm identificado genes envolvidos na manutenção do isolamento reprodutivo através da redução do valor adaptativo de híbridos, os chamados “genes de especiação” (WU; TING, 2004). Este é o caso do gene *Odysseus* (*OdsH*), responsável por causar esterilidade em machos híbridos de *Drosophila simulans* quando feita a sua introgressão em *D. mauritiana*. O gene *OdsH* originou-se por duplicação do gene *unc-4* e evoluiu a taxas mais aceleradas que sua duplicata no clado de *D. melanogaster*, apresentando notável divergência em suas sequências, funções e expressão (TING *et al.*, 1998).

Enquanto *unc-4* possui expressão em células neurais do Sistema Nervoso Central, *OdsH* é expresso em células reprodutivas masculinas (TING *et al.*, 2004). Através de estudos de expressão e *knockout* de gene, SUN *et al.* (2004) demonstraram que a provável função de *OdsH* está relacionada a um modesto aumento da fertilidade em machos jovens. Indícios de uma provável interação de *OdsH* com elementos transponíveis (TEs) também foram reportados por FRANCHINI *et al.* (2004) ao detectarem uma cópia do retroelemento *blastopia* inserida na região do primeiro íntron. Esta sequência apresentou cerca de 3,2 mil pares de base (pb) de comprimento, sendo uma cópia quase completa do retroelemento – neste caso, a cópia canônica possui cerca de 3,8 mil pb (FROMMER *et al.*, 1994).

Quanto à taxonomia, tanto *D. simulans* quanto *D. mauritiana* e *D. melanogaster* pertencem ao grupo de espécies *melanogaster*, incluído no subgênero *Sophophora* de *Drosophila* (BÄCHLI, 2022; SUVOROV *et al.*, 2022). No entanto, este grupo pertence à linhagem de *Sophophora* do Velho Mundo (O'GRADY; KIDWELL, 2002). O subgrupo *willistoni* - alvo desse estudo, por sua vez, pertence ao clado neotropical de *Sophophora* e apresenta ampla distribuição predominantemente neotropical (O'GRADY; KIDWELL, 2002; ZANINI *et al.*, 2015a, 2015b), compreendendo seis espécies crípticas: *Drosophila equinoxialis*, *D.*

insularis, *D. pavlovskiana*, *D. tropicalis*, *D. willistoni* e *D. paulistorum* (BÄCHLI, 2022), sendo esta última um complexo de seis subespécies (DOBZHANSKY; SPASSKY, 1959; ZANINI *et al.*, 2018). Este complexo, por sua vez, inclui as subespécies *D. paulistorum amazonian*, *D. paulistorum andeanbrazilian*, *D. paulistorum centroamerican*, *D. paulistorum interior*, *D. paulistorum orinocan* e *D. paulistorum transitional*.

As espécies do subgrupo possuem diferentes graus de isolamento reprodutivo que variam de isolamento pré-zigótico completo até isolamento pós-zigótico incompleto, com a formação de híbridos (DOBZHANSKY *et al.*, 1969). Há décadas, esse subgrupo se apresenta como organismo modelo promissor para estudos que abordem processos evolutivos na região neotropical, principalmente devido ao fato de ser constituído de espécies e subespécies crípticas e pela sobreposição na sua distribuição geográfica. Apesar do conhecimento acerca das interações e do papel do *OdsH* no isolamento reprodutivo em espécies do grupo *melanogaster*, pouco se sabe a respeito da presença tanto deste gene – e elementos repetitivos a ele associados – quanto do seu efeito nas espécies do subgrupo *willistoni*.

O objetivo principal deste trabalho foi, portanto, caracterizar a região do gene *OdsH* nos genomas sequenciados do subgrupo *willistoni* disponíveis nos bancos de dados. Essa caracterização consistiu em identificar a presença de elementos de transposição e testar a pressão seletiva atuante. Além disso, foi feita a identificação a nível de subespécie para a linhagem 14030-0771.06 de *D. paulistorum* do Bloomington Stock Center. Esses dados serão fundamentais para a continuidade do estudo que pretende avaliar a expressão transcricional desse gene em híbridos do complexo *D. paulistorum*, os quais já foram estabelecidos e vêm sendo mantidos em laboratório.

2. MÉTODOS

2.1. Busca in silico e anotação dos genes *unc-4* e *OdsH*

As espécies do subgrupo *willistoni* com genoma sequenciado são: *Drosophila equinoxialis*, *D. insularis*, *D. paulistorum*, *D. tropicalis* e *D. willistoni* (DROSOPHILA 12 GENOMES CONSORTIUM, 2007; KIM *et al.*, 2021); as montagens estão depositadas no GenBank (NCBI). Para a identificação de *unc-4* e *OdsH* em espécies do subgrupo *willistoni* nas quais os respectivos genes não estão anotados, utilizou-se a ferramenta BLAST (NCBI) para identificar regiões de homologia, usando como sonda sequências de mRNA de *Drosophila melanogaster* dos dois genes, depositadas no GenBank (NCBI) (códigos de acesso: NM_078665.3 e NM_133014.3; para *OdsH* e *unc-4*, respectivamente). Foram selecionados os *hits* com maior *score* e menor *e-value*. Entre as espécies do subgrupo *willistoni*, somente *D. willistoni* possui ambos os genes anotados; nesse caso, as respectivas sequências foram recuperadas do NCBI.

As sequências obtidas foram alinhadas no programa MAFFT 7 (KATO; STANDLEY, 2013) e editadas no programa AliView (LARSSON, 2014), obtendo-se uma sequência consenso contendo ambos os genes (íntrons e éxons) e a região intergênica. Para todas as espécies, a região que compreende o primeiro íntron foi submetida à ferramenta CENSOR (KOHANY *et al.*, 2006) para busca de elementos repetitivos por homologia. As sequências classificadas como elementos transponíveis foram submetidas à busca por domínios conservados na *Conserved Domain Database* – CDD (NCBI). Devido ao fato de os éxons 1 e 4 possuírem poucas regiões de homologia, somente os éxons 2 e 3 foram utilizados nas análises posteriores. A região escolhida compreende o homeodomínio (165 pb), mais 12 pb a montante e 99 pb a jusante.

2.2. Teste de variação de dN/dS entre linhagens e testes de seleção natural

O alinhamento das sequências de *OdsH* foi usado na reconstrução da filogenia por Máxima Verossimilhança, realizada no programa MEGA X (KUMAR *et al.* 2018). O melhor modelo de substituição nucleotídica foi escolhido no programa ModelTest-NG (DARRIBA *et al.*, 2020), implementado na plataforma PhyloSuite 1.1.2 (ZHANG *et al.*, 2020), de acordo com o critério BIC. Para estimar os valores de suporte de ramos, foram utilizadas 500 réplicas de *bootstrap*.

O número de substituições não-sinônimas por sítios não-sinônimos (dN) e o número de substituições sinônimas por sítios sinônimos (dS) foi estimado por inferência Bayesiana em análises para-a-par no software PAML (YANG, 2004).

Após, foram realizados testes para identificar o tipo de pressão seletiva atuante nos ramos da filogenia de *OdsH* das espécies estudadas. Para isso, dois modelos de análise por verossimilhança foram aplicados às sequências de estudo e cada modelo foi testado comparando o valor de log likelihood ($2\Delta l$) com uma distribuição χ^2 . Os testes foram realizados no programa CODEML do pacote PAML (YANG, 2004). Comparou-se o modelo *one-ratio*, que assume uma única razão ω para todos os ramos da filogenia, com o modelo *free-ratio*, que assume uma razão ω diferente para cada ramo da filogenia.

2.3. Identificação da linhagem 14030-0771.06, cruzamentos e a geração de híbridos

As linhagens DP-Andina e 14030-0771.06, de *D. paulistorum andeanbrazilian* e *D. paulistorum* ssp. respectivamente, foram utilizadas como modelo de análise para geração de híbridos. Essas linhagens são mantidas em laboratório em câmara de temperatura ($21\pm 1^\circ\text{C}$) e umidade (60% rh) constante, com fotoperíodo 12h:12h. Os adultos são repicados semanalmente para um frasco contendo meio de cultura estéril para garantir o desenvolvimento de todos os estágios do ciclo de vida e manutenção da população estoque. Essas duas linhagens são oriundas de Florianópolis, Brasil, e do Bloomington Stock Center, respectivamente. Esta última teve seu genoma sequenciado por KIM *et al.* (2021) e é a mesma linhagem de *D. paulistorum* empregada aqui na caracterização do gene *OdsH* e análises posteriores.

Apesar de ter seu genoma sequenciado e ser comercializada pelo *Drosophila* Stock Center, a linhagem 14030-0771.06 não tem sua caracterização a nível de subespécie definida. Décadas de estudos acumulados em nosso grupo de pesquisa possibilitaram encontrar marcadores moleculares adequados para a definição das subespécies de *D. paulistorum* (ZANINI *et al.* 2018). Assim, buscas por BLASTn foram realizadas para localizar sequências correspondentes a cinco genes nucleares (*Adh*, *Ddc*, *Hb*, *kl-3* e *per*) no genoma da subespécie 14030-0771.06, utilizando sequências previamente caracterizadas para *D. paulistorum* depositadas no GenBank (NCBI) (códigos de acesso: EU532122, MG010062, MG010076,

EU532110, U51080; respectivamente) como sondas. Foram incluídas também sequências pertencentes às seis subespécies de *D. paulistorum*, sequenciadas por ZANINI *et al.* (2018), *D. equinoxialis* (espécie-irmã do complexo *D. paulistorum*) e *D. willistoni* (espécie-irmã às demais) como grupo externo.

Cada marcador teve suas sequências alinhadas utilizando o algoritmo ClustalW no software MEGA X (KUMAR *et al.* 2018), sendo removidas regiões de alinhamento espúrio no programa trimAl (CAPELLA-GUTIERREZ *et al.*, 2009). O melhor modelo de substituição nucleotídica foi escolhido separadamente para cada gene pelo critério de AIC, no IQ-TREE 2 (MINH *et al.*, 2020). A matriz final, concatenada e particionada, foi montada utilizando o programa FASconCAT (KÜCK; MEUSEMANN, 2010). A posição filogenética da subespécie foi, então, inferida por inferência bayesiana no programa MrBayes 3.2.7 (RONQUIST *et al.* 2012), com modelo de árvore *birth-death process*, utilizando 1 milhão de gerações de MCMC iniciais, com amostragem de árvores a cada 100 gerações e *burn-in* de 25%. A convergência da análise foi dada como satisfatória quando o desvio-padrão médio das frequências atingiu valor menor do que 0,01.

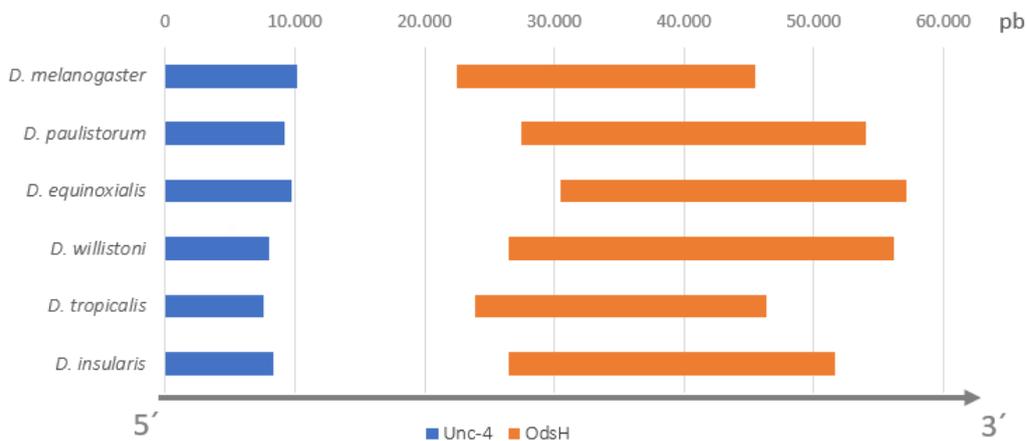
Os testes de cruzamentos entre as linhagens foram realizados com escolha única, onde uma fêmea virgem é colocada em um frasco contendo meio de cultura junto a apenas um macho da outra subespécie. Os cruzamentos foram observados por três horas e, após esse período, as fêmeas que realizaram cópula foram realocadas em novos frascos contendo o meio de cultura para realizarem a oviposição. Estes frascos foram mantidos em laboratório, em câmara com as mesmas condições de temperatura e umidade da linhagem estoque. Foram realizados 20 testes de cruzamentos. Esses híbridos posteriormente serão avaliados quanto a expressão transcricional relativa do gene *OdsH* por PCR quantitativo. Tanto os primers para análise do gene *OdsH* [desenhados para esse trabalho com o auxílio do programa Primer3 (<https://primer3.ut.ee/>)] quanto dos normalizadores Rp49 (LALONDE *et al.* 2006), Alpha-tubulin e EF1 (PONTON *et al.* 2012) já foram sintetizados.

3. RESULTADOS

3.1. Ocorrência de *unc-4* e *OdsH* no subgrupo *willistoni*

Em todas as espécies, *unc-4* e *OdsH* se encontram em tandem e preservam a mesma estrutura éxon-íntron. A região intergênica variou de 12.302 pares de base (pb) para *D. melanogaster* e 20.645 pb para *D. equinoxialis*. O tamanho de *unc-4* variou de 7.571 pb em *D. tropicalis* a 10.140 pb em *D. melanogaster*, enquanto *OdsH* variou de 22.487 pb em *D. tropicalis* a 29.728 em *D. willistoni* (Fig. 1).

Figura 1. Representação do tamanho e localização de *unc-4* (azul) e *OdsH* (laranja). O comprimento das barras corresponde ao tamanho dos genes, em pares de base. A seta na parte inferior indica o sentido de transcrição (5'-3').



Fonte: o autor.

3.2. Elementos transponíveis associados ao *OdsH*

A busca por elementos transponíveis (TEs) no primeiro íntron de *OdsH* resultou na detecção dos elementos *Mariner-29* e *Jockey-2* em *D. paulistorum andeanbrazilian* e *D. paulistorum ssp.*, respectivamente (Tabela 1). As demais espécies apresentaram poucas regiões homólogas à elementos repetitivos. Foi detectado domínio conservado da transposase IS630 na cópia de *Mariner-29*; já a cópia de *Jockey-2* apresentou domínios da *Exonuclease Endonuclease Phosphatase* (EEP) e da retrotransposase de elementos não LTR (*RT_nonLTR*).

Tabela 1. Descrição dos resultados obtidos na caracterização do primeiro íntron do gene *Odysseus* (*OdsH*).

Espécie	Transposon	Íntron			Transposon	
		Tamanho	Início	Fim	Início	Fim
<i>D. paulistorum andeanbrazilian</i>	<i>Mariner-29</i>	22.423	16.278	17.916	1	1.611
<i>D. paulistorum</i> ssp.	<i>Jockey-2</i>	25.552	12.181	16.747	1	4.551

Valores apresentados em pares de base (pb).

3.3. Testes de seleção

As sequências de aminoácidos de *unc-4* e *OdsH* de parte dos éxons 2 e 3 preditas para as cinco espécies estudadas do subgrupo *willistoni* foram alinhadas com as respectivas sequências de *D. melanogaster* (Fig. 2). Os homólogos *unc-4* apresentaram sequência idêntica para todas as seis espécies, indicando que todas as substituições ao longo da região selecionada do gene foram sinônimas. Já os homólogos de *OdsH* apresentaram 26 substituições ao longo dos 92 aminoácidos. Quando comparadas somente as espécies do subgrupo *willistoni* ocorrem 13 substituições ao longo da sequência selecionada, sendo cinco substituições dentro do homeodomínio.

Figura 2. Alinhamento das sequências de aminoácidos de *OdsH* e *unc-4* preditas das cinco espécies de *Drosophila*. A cabeça de seta indica a posição do íntron e a linha indica o homeodomínio.

		exon 2	▼	exon 3	
OdsH_mel	RRGR	TNFNSWQLRELERVFQGSHPYDIFMREALATKLDLMEGRIAVWFQNRRAKWRKQE		HTKKGPGRPAHNAHPQSCSGDPIPLSELRAREL	
OdsH_pau	..C.EQ..LS.HH.....VR...R.....K.	P...Q..T...A...I..Q...R	
OdsH_equ	..C.EQ..LS.HH.....VR...R.....K.	P...Q..T...A...I..Q...R	
OdsH_wil	..C.EQ..LS.HH.....MR...R.....K.	Q.KT...A...I..Q..Q..R	
OdsH_trp	..C.EQ..LS.HHN...MYV...MR...R.....K.	QLRT...A...PV..Q...R	
OdsH_ins	..C.EQ..LS.HH.....Y...MR...R.....K.	Q...A...I..Q...S	
unc4_mel	..S.E...A.SA.....MR...K.S.V.....V..R.	Q..T...E...PN..K.K.R	

Fonte: o autor.

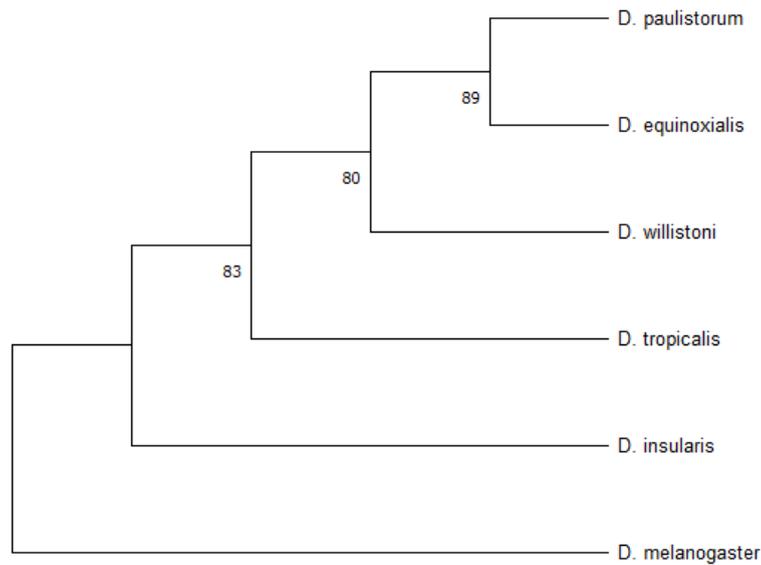
As análises de dN/dS apresentaram ampla variação no valor de ω , variando de 0,001 entre *D. equinoxialis* e *D. paulistorum* a 1,027 entre *D. equinoxialis* e *D. willistoni* (Tabela 2). Valores de $\omega > 1$ indicam a ocorrência de seleção positiva, onde substituições não sinônimas teriam uma taxa de fixação maior do que as substituições sinônimas; enquanto valores de $\omega < 1$ indicam seleção purificadora, o que é observado para a maioria das comparações entre linhagens realizadas aqui – com exceção de *D. equinoxialis* e *D. willistoni*.

Tabela 2. Número de substituições não-sinônimas por sítio não-sinônimo (dN); número de substituições sinônimas por sítio sinônimo (dS) e razão dN/dS em *OdsH* para espécies do subgrupo *willistoni*.

Par de espécies	dN	dS	dN/dS
equ-pau	0.0001	0.0653	0.0010
equ-trp	0.0518	0.0826	0.6274
equ-ins	0.0198	0.2305	0.0858
equ-wil	0.0200	0.0194	1.0272
pau-trp	0.0518	0.1498	0.3462
pau-ins	0.0198	0.3122	0.0633
pau-wil	0.0210	0.0766	0.2735
wil-trp	0.0478	0.1225	0.3903
wil-ins	0.0197	0.2628	0.0749
ins-trp	0.0409	0.3517	0.1164

O modelo K2P, com 4 categorias e distribuição gama, foi estimado como o melhor modelo de substituição nucleotídica de acordo com o critério BIC. As sequências de *OdsH* formaram um agrupamento com *D. melanogaster* como grupo externo. *Drosophila insularis* foi a primeira espécie a divergir no subgrupo *willistoni*, seguida de *D. tropicalis* e *D. willistoni*. *Drosophila equinoxialis* e *D. paulistorum* apresentaram-se como espécies-irmãs (Fig. 3). A topologia da árvore reconstruída está de acordo com os dados mais recentes reportados na literatura (SUVOROV *et al.*, 2022).

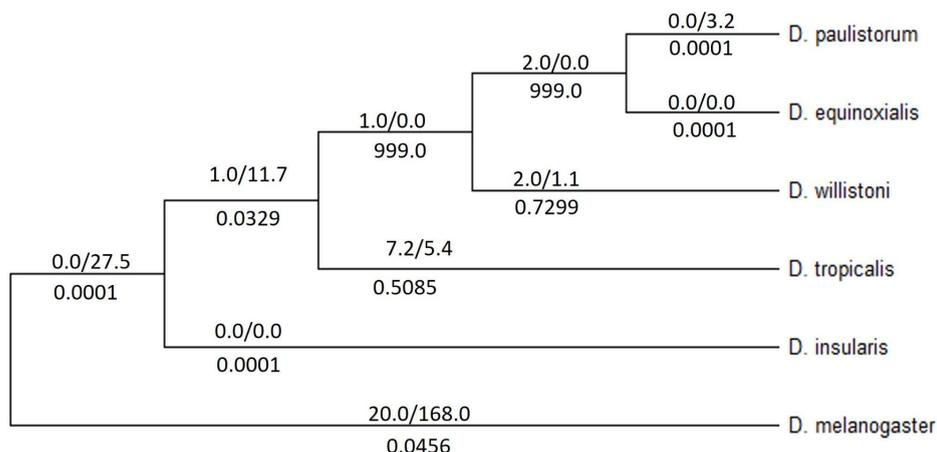
Figura 3. Relações filogenéticas entre as sequências de *OdsH* encontradas em espécies do subgrupo *willistoni*, inferidas por máxima verossimilhança. Os valores em cada nó representam o suporte de ramo, baseados em 500 réplicas de *bootstrap*.



Fonte: o autor.

O modelo *one-ratio*, que assume um parâmetro ω para a árvore, obteve valor de log likelihood de ($l_0 = 715,32$), enquanto o modelo *free-ratio*, que usou 10 parâmetros ω , obteve ($l_1 = 702,06$). Portanto o valor de $2\Delta l = 26,52$ foi comparado com a distribuição χ^2 com nove graus de liberdade, dado pela diferença no número de parâmetros ω entre os dois modelos. A diferença entre os dois modelos é significativa ($0,01 > P > 0,001$), indicando que o modelo de razão ω variável é o que melhor se aplica aos dados. O número de substituições sinônimas e não-sinônimas e o valor da razão ω em cada ramo da árvore é mostrado na figura 4.

Figura 4. Filogenia utilizada para os testes de dN/dS. Os dois números mostrados acima de cada ramo são, respectivamente, as substituições sinônimas e não-sinônimas estimadas para cada ramo. Abaixo de cada ramo está o valor da razão ω para cada ramo.

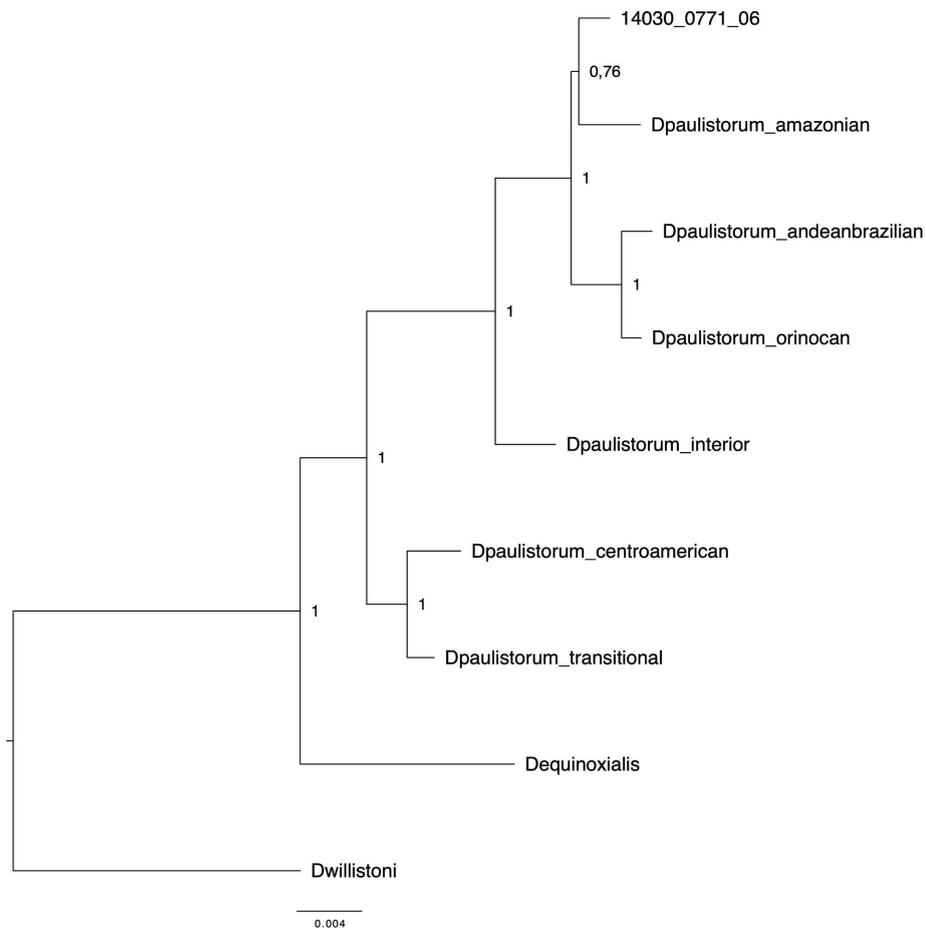


Fonte: o autor.

3.4. Testes de cruzamentos e identificação da linhagem 14030-0771.06

Sequências dos genes nucleares *Adh*, *Ddc*, *Hb*, *kl-3* e *per* foram obtidas para a linhagem 14030-0771.06 a partir de buscas por BLASTn no seu genoma. Essas sequências foram utilizadas para comparação com sequências previamente empregadas para diferenciação das 6 subespécies de *D. paulistorum*. A topologia reconstruída pela inferência bayesiana (Fig. 5) recuperou a monofilia do complexo *D. paulistorum* (PP = 1,0), com *D. equinoxialis* como espécie-irmã (PP = 1,0) e *D. willistoni* como grupo externo (PP = 1,0). A linhagem 14030-0771.06 agrupou com a subespécie *D. paulistorum amazonian* (PP = 0,76). O cruzamento DP-Andino x 14030-0771.06 foi único capaz de gerar híbridos.

Figura 5. Árvore de máxima credibilidade de clados, reconstruída por inferência bayesiana, representando a posição filogenética da linhagem 14030-0771.06 dentro do complexo de subespécies *Drosophila paulistorum*. Números próximos aos nós representam a sua probabilidade posterior (PP).



Fonte: o autor.

4. DISCUSSÃO

Neste estudo, foi realizada a caracterização genômica da região que compreende os genes *unc-4* e *Odysseus (OdsH)* nas espécies do subgrupo *willistoni* do gênero *Drosophila*. O gene *OdsH* é parálogo ao gene *unc-4* e foi descrito ao causar esterilidade em híbridos no grupo *melanogaster*. Assim, este trabalho fornece as bases para o entendimento da função e da interação deste gene com o isolamento reprodutivo no complexo de subespécies *D. paulistorum*, do subgrupo *willistoni*.

Sobre a estrutura dos ortólogos de *OdsH*, é interessante notar que a sequência de aminoácidos de seus éxons 1 e 4 apresenta muitos polimorfismos, inclusive quando são comparadas espécies pertencentes ao mesmo subgrupo. Essa questão é frequentemente subestimada em estudos que tratam do *OdsH*, contudo analisar especificamente esses dois éxons é essencial para termos um melhor entendimento da estrutura desse gene.

Os valores de dN/dS obtidos por análise par-a-par, encontrados para as espécies do subgrupo *willistoni*, são, em geral, muito baixos. Com exceção das espécies *D. willistoni* e *D. equinoxialis* – que apresentaram valor de $\omega > 1$ – todos os outros pares de espécies obtiveram valores de $\omega < 1$, sugerindo que *OdsH* está sendo sujeito a seleção purificadora nesse clado. Também foram observados valores de $\omega < 1$ nas análises por ramos. Embora o ramo do ancestral de *D. willistoni* + *equinoxialis* + *D. paulistorum* e *D. paulistorum* + *D. equinoxialis* apresentaram valores de $\omega > 1$, o número de substituições não-sinônimas nesses ramos foram muito baixos (1,0 e 2,0, respectivamente), sendo pouco informativo. Baixos valores de dN/dS para *OdsH* também foram observados por WEN *et al.* (2006) nos clados das espécies *D. lini* e *D. kikkawai* – espécies do grupo *melanogaster* de *Drosophila* – sugerindo que *OdsH* também foi submetido a uma forte seleção purificadora nesses dois clados.

A presença de elementos de transposição em regiões gênicas e intergênicas já é bem documentada na literatura. No entanto, a presença do transposon *mariner* e do retrotransposon *Jockey* no primeiro íntron de *D. paulistorum* se apresenta como uma intrigante interação a ser elucidada. De fato, ANTONIOLLI *et al.* (2022, submetido à publicação) observaram não apenas que a fração de elementos repetitivos no genoma dessa espécie é bastante elevada (cerca de 41% e 39% em *D. paulistorum andeanbrazilian* e *D. paulistorum ssp.*, respectivamente), mas

também que o retrotransposon *Jockey* é um dos que apresentam maior número de cópias idênticas nestes genomas. Altas taxas de similaridade entre cópias de transposons são indicativos de mobilização recente destas sequências (SCHÖN *et al.*, 2021). Além disso, a história evolutiva do *Jockey* é marcada pela alta frequência de transferências horizontais em drosofilídeos (TAMBONES *et al.*, 2019).

Os resultados aqui reportados são congruentes com os encontrados para o elemento *blastopia*, também presente no primeiro íntron do gene *OdsH*, em *D. melanogaster*, ao apresentar sequências com mais de 99% de similaridade (FRANCHINI *et al.*, 2004). Ainda, a ausência de elementos repetitivos no primeiro íntron nas demais espécies do subgrupo *willistoni* é outro indicativo de que esta região genômica foi alvo de mobilização recente de elementos transponíveis (TEs). A ação, tanto de seleção positiva nessa região, quanto da atuação desses TEs na regulação de *OdsH*, ainda será avaliada. O envolvimento de TEs na geração de novidades evolutivas ou estabelecimento de novas redes regulatórias tem sido relatado (DEPRÁ *et al.*, 2009; CHUONG *et al.*, 2017; LORETO *et al.*, 2018).

Os cruzamentos realizados com as subespécies *D. paulistorum amazonian* e *D. paulistorum andeanbrazilian* resultaram na produção de híbridos. Esse resultado é congruente com os resultados reportados por EHRMAN (1960), a qual obteve fêmeas férteis e machos estéreis no cruzamento entre essas duas subespécies. O presente estudo, portanto, revalida os experimentos de EHRMAN (1960) e estabelece os protocolos para obtenção de híbridos que serão posteriormente utilizados em análises transcricionais, tendo como alvo o *OdsH*.

Por fim, a formação de novas espécies ao longo da evolução certamente não pode ser explicada somente pela ação de um gene isolado, uma vez que fatores geográficos e ecológicos complexos possuem papel fundamental na maior parte dos casos de especiação. Entretanto, pouco se sabe sobre as bases genéticas dos processos de isolamento pré e pós-zigóticos atuantes em organismos em estágios de especiação. Estudar clados de espécies em diferentes estágios de isolamento reprodutivo – o caso do subgrupo *willistoni* – nos permitirá obter insights sobre os processos que moldam a formação de novas espécies.

5. REFERÊNCIAS

- ANTONIOLLI, H.R.M.; DEPRÁ, M.; VALENTE, V.L.S. Patterns of genome size evolution versus fraction of repetitive elements *in statu nascendi* species: the case of the *willistoni* subgroup of *Drosophila* (Diptera, Drosophilidae). Submetido à publicação, em 2022.
- BÄCHLI, G. **TaxoDros: the database on taxonomy of Drosophilidae**, 2022. Disponível em <<https://www.taxodros.uzh.ch/>>. Acesso em: 08 out 2022.
- CAPELLA-GUTIERREZ, S.; SILLA-MARTINEZ, J.M.; GABALDON, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. **Bioinformatics**, v. 25, p. 1972–1973, 2009.
- CHUONG, E.; ELDE, N.; FESCHOTTE, C. Regulatory activities of transposable elements: from conflicts to benefits. **Nature Reviews Genetics**, v. 18, p. 71–86, 2017.
- DARRIBA, D.; POSADA, D.; KOZLOV, A.M.; STAMATAKIS, A.; MOREL, B.; FLOURI, T. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. **Molecular Biology and Evolution**, v. 37, n. 1, p. 291–294, 2020.
- DEPRÁ, M.; VALENTE, V.L.S.; MARGIS, R.; LORETO, E.L.S. The *hobo* transposon and *hobo*-related elements are expressed as developmental genes in *Drosophila*. *Gene*, v. 448, n. 1, p. 57–63, 2009.
- DOBZHANSKY, T.; SPASSKY, B. *Drosophila paulistorum*, a cluster of species *in statu nascendi*. **Proceedings of the National Academy of Sciences**, v. 45, n. 3, p. 419–428, 1959.
- DOBZHANSKY, T.; PAVLOVSKY, O. Spontaneous origin of an incipient species in the *Drosophila paulistorum* complex. **Proceedings of the National Academy of Sciences**, v. 55, n. 4, p. 727–733, 1966.
- DOBZHANSKY, T.; PAVLOVSKY, O.; EHRMAN, L. 1969. Transitional Populations of *Drosophila paulistorum*. **Evolution**, v. 23, n. 3, p. 482–492, 1969.
- DROSOPHILA 12 Genomes Consortium. Evolution of genes and genomes on the *Drosophila* phylogeny. **Nature**, v. 450, p. 203–218, 2007.
- EHRMAN, L. A genetic constitution frustrating the sexual drive in *Drosophila paulistorum*. **Science**, v. 131, n. 3410, p. 1381–1382, 1960.
- FRANCHINI, L.F.; GANKO, E.W.; MCDONALD, J.F. Retrotransposon-Gene Associations Are Widespread Among *D. melanogaster* Populations. **Molecular Biology and Evolution**, v. 21, n. 7, p. 1323–1331, 2004.

FROMMER, G.; SCHUH, R.; JACKE, H. Localized expression of a novel microRNA-like element in the blastoderm of *Drosophila melanogaster* is dependent on the anterior morphogen *bicoid*. **Chromosoma**, v. 103, p. 82–89, 1994.

KATOH, K.; STANDLEY, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. **Molecular Biology and Evolution**, v. 30, n. 4, p. 772–780, 2013.

KIM, B.Y.; WANG, J.R.; MILLER, D.E.; BARMINA, O.; DELANEY, E.; THOMPSON, A.; et al. Highly contiguous assemblies of 101 drosophilid genomes. **eLife**, 10:e66405, 2021.

KOHANY, O.; GENTLES, A.J.; HANKUS, L.; JURKA, J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. **BMC Bioinformatics**, v. 7, n. 474, 2006.

KÜCK, P.; MEUSEMANN, K. FASconCAT: Convenient handling of data matrices. **Molecular Phylogenetics and Evolution**, v. 56, n. 3, p. 1115–1118, 2010.

KUMAR, S.; STECHER, G.; LI, M.; KNYAZ, C.; TAMURA, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. **Molecular Biology and Evolution**, v. 35, n. 6, p. 1547–1549, 2018.

LALONDE, M.; JANSSENS, H.; YUN, S.; et al. A role for Phospholipase D in *Drosophila* embryonic cellularization. **BMC Developmental Biology**, v. 6, n. 60, 2006.

LARSSON, A. AliView: a fast and lightweight alignment viewer and editor for large datasets. **Bioinformatics**, v. 30, n. 22, p. 3276–3278, 2014.

LORETO, E.L.S.; DEPRÁ, M.; DIESEL, J.F.; PANZERA, Y.; VALENTE-GAIESKY, V.L.S. *Drosophila* relics *hobo* and *hobo*-MITEs transposons as raw material for new regulatory networks. **Genetics and Molecular Biology**, v. 41, suppl. 1, p. 198–205, 2018.

MINH, B.Q.; SCHMIDT, H.A.; CHERNOMOR, O.; SCHREMPF, D.; WOODHAMS, M.D.; VON HAESLER, A.; LANFEAR, R. IQTREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. **Molecular Biology and Evolution**, v. 37, n. 5, p. 1530–1534, 2020.

NATIONAL Center For Biotechnology Information – NCBI. Disponível em: <<https://ncbi.nlm.nih.gov/>>. Acesso em: 08 out 2022.

O'GRADY, P.M.; KIDWELL, M.G. Phylogeny of the subgenus *Sophophora* (Diptera: Drosophilidae) based on combined analysis of nuclear and mitochondrial sequences. **Molecular Phylogenetics and Evolution**, v. 22, n. 3, p. 442–453, 2002.

PÉREZ-SALAS, S.; RICHMOND, R. C.; PAVLOVSKY, O. A.; EHRMAN, L.; DOBZHANSKY, T. The interior semispecies of *Drosophila paulistorum*. **Evolution**, v. 24, p. 519–527, 1970.

PONTON, F.; CHAPUIS, M.P.; PERNICE, M.; SWORD, G.A.; SIMPSON, S.J. Evaluation of potential reference genes for reverse transcription-qPCR studies of physiological responses in *Drosophila melanogaster*. **Journal of Insect Physiology**, v. 57, n. 6, p. 840–850, 2011.

RONQUIST, F.; TESLENKO, M.; VAN DER MARK, P.; AYRES, D.L.; DARLING, A.; HÖHNA, S.; LARGET, B.; LIU, L.; SUCHARD, M.A.; HUELSENBECK, J.P. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. **Systematic Biology**, v. 61, n. 3, p. 539–42, 2012.

SCHÖN, I.; RODRIGUEZ, F.; DUNN, M.; MARTENS, K.; SHRIBAK, M.; ARKHIPOVA, I.R. A Survey of Transposon Landscapes in the Putative Ancient Asexual Ostracod *Darwinula stevensoni*. **Genes**, v. 12, n. 13, 401, 2021.

SEEHAUSEN, O.; BUTLIN, R. K.; KELLER, I.; WAGNER, C. E.; BOUGHMAN, J. W.; HOHENLOHE; et al. Genomics and the origin of species. **Nature Reviews Genetics**, v. 15, n. 3, p. 176–192. 2014.

SUN, S.; TING, C-T.; WU, C-I. The normal function of a speciation gene, *Odysseus*, and its hybrid sterility effect. **Science**, v. 305, n. 5680, p. 81-83, 2004.

SUVOROV, A.; KIM, B.Y.; WANG, J.; ARMSTRONG, E.E.; PEEDE, D.; D'AGOSTINO, E.R.R.; et al. Widespread introgression across a phylogeny of 155 *Drosophila* genomes. **Current Biology**, v. 32, p. 111–123, 2022.

TAMBONES, I.L.; HAUDRY, A.; SIMÃO, M.C.; CARARETO, C.M.A. High frequency of horizontal transfer in Jockey families (LINE order) of drosophilids. *Mobile DNA*, v. 10, n. 43, 2019.

TING, C.-T.; TSAUR, S.-C.; WU, M.-L.; WU, C.-I. A rapidly evolving homeobox at the site of a hybrid sterility gene. **Science**, v. 282, n. 5393, p. 1501–1504, 1998.

Wen, S.-Y., Shimada, K., Kawai, K., & Toda, M. J. (2006). Strong Purifying Selection on the *Odysseus* Gene in Two Clades of Sibling Species of the *Drosophila montium* Species Subgroup. **Journal of Molecular Evolution**, 62(5), 659–662.

WU, C.-I.; TING, C.-T. Genes and speciation. **Nature Reviews Genetics**, v. 5, n. 2, p. 114–122, 2004.

YANG, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. **Molecular Biology and Evolution**, v. 24, n. 8, p. 1586–1591, 2007.

ZANINI, R.; DEPRÁ, M.; VALENTE, V.L.S. On the geographic distribution of the *Drosophila willistoni* group (Diptera, Drosophilidae) - updated geographic distribution of the Neotropical *willistoni* subgroup. **Drosophila Information Service**, v. 98, p. 39-43, 2015a.

ZANINI, R.; DEPRÁ, M.; VALENTE, V. L.S. Can sibling species of the *Drosophila willistoni* subgroup be recognized through combined microscopy techniques? **Revista Brasileira de Entomologia**, 59(4), 323–331, 2015

ZANINI, R.; MÜLLER, M. J.; VIEIRA, G. C.; VALIATI, V. H.; DEPRÁ, M.; VALENTE, V.L.S. Combining morphology and molecular data to improve *Drosophila paulistorum* (Diptera, Drosophilidae) taxonomic status. **Fly**, v. 12, n. 2, p. 81–94, 2018.

ZHANG D, GAO F, JAKOVLIĆ I, ET AL. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. **Molecular Ecology Resources**, v. 20, p. 348–355, 2020.