

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

**ANÁLISE DO ELEMENTO TRANSPONÍVEL *copia* EM  
ESPÉCIES DE *Drosophila***

**DISSERTAÇÃO DE MESTRADO**

**Paloma Menezes Rubin**

**Santa Maria, RS, Brasil.**

**2008**

**2008**

**Mestre**

**RUBIN, Paloma Menezes**

**PPGBA/UFSM, RS**

**ANÁLISE DO ELEMENTO TRANSPONÍVEL *copia* EM  
ESPÉCIES DE *Drosophila***

por

**Paloma Menezes Rubin**

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Biodiversidade Animal, Área de Concentração em Biologia Evolutiva de Insetos, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Biodiversidade Animal.**

**Orientadora: Vera Lúcia da Silva Valente  
Co-orientador: Élgion L. da Silva Loreto**

**Santa Maria, RS, Brasil.**

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

A Comissão Examinadora, abaixo assinada,  
aprova a Dissertação de Mestrado

**ANÁLISE DO ELEMENTO TRANSPONÍVEL *copia*  
EM ESPÉCIES DE *Drosophila***

elaborada por  
**Paloma Menezes Rubin**

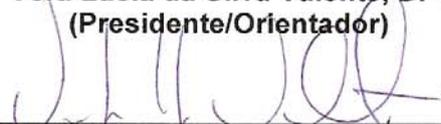
como requisito parcial para obtenção do grau de  
**Mestre em Biodiversidade Animal**

**COMISSÃO EXAMINADORA:**



---

Vera Lúcia da Silva Valente, Dr<sup>a</sup>.  
(Presidente/Orientador)



---

Victor Hugo Valiati, Dr<sup>o</sup>. (UNISINOS)



---

Fabiano Pimentel Torres, Dr<sup>o</sup>.

Santa Maria, 30 de abril de 2008.

À Helena, Daniel, Pedro  
e Felipe, por serem minha  
grande fonte de inspiração.

## AGRADECIMENTOS

À professora Vera pela oportunidade e pela honra de ter me aceitado como sua orientanda, mesmo sem me conhecer. Por seus incentivos valiosos, pelo carinho prestado desde o início deste trabalho, por sua amizade, e pela orientação sempre presente, mesmo estando em Porto Alegre.

Ao professor Élgion, que provavelmente não faz idéia da pessoa que me tornei após conhecê-lo. Sem seu suporte e orientação eu não teria chegado até aqui. Quando caí de pára-quadras no seu laboratório, seus conhecimentos e empolgação me deixaram apaixonada pelas moscas e pelos transposons. Sua ajuda, amizade e incentivo fizeram-me acreditar que eu poderia pesquisar; e principalmente, sua paciência não me deixou desistir nos momentos de sufoco. Os méritos são todos seus. Obrigada profe!

À Lenira por estar sempre presente e me fazer enxergar que pesquisas de qualidade também podem ser feitas com bom humor. Obrigada pela ajuda no nosso trabalho, e pela sua amizade.

À professora Claudia Carareto pelas idealizações do projeto e a colaboração com sugestões ao trabalho.

À Cristina Parada pelas membranas que foram re-hibridizadas.

À Ana Lauer Garcia pela identificação das espécies estudadas.

À Sonia Cechin e ao Sr. Paulo por estarem sempre dispostos a buscar soluções para os “pepinos” dos alunos do Curso de Pós-Graduação em Biodiversidade Animal.

Aos meus amores do LabDros, de ontem e de hoje! Começando pela Adri e pelo Newton, meus grandes incentivadores quando eu observava tudo, ainda assustada. O apoio e o exemplo de vocês foi fundamental. À Nina, pelas incansáveis conversas, pelas risadas demoradas, pela presença constante, pela ajuda com o inglês e pela amizade indispensável. À Maríndia e à Lizandra por estarem sempre dispostas a me ajudar, e ainda se tornarem minhas amigas. Aos meus companheiros de conversas produtivas e inteligentes: Mauro e Kari, que estão longe agora, mas pertinho. À Carolzinha, Juline, Ana, Brenda, Lucieli e Tális pelo carinho e pela amizade. Um agradecimento

especial àquelas pessoinhas que conviveram comigo todos os dias, e tornaram-se meus grandes amigos: Ronaldinho, Camila, Gabinho (além de amigo, agora compadre!), Felipe ten Caten, Ana Paula, Francis, Bruna e Josmael. Vocês são insubstituíveis. Ao Lucas e à Jéssica: muito obrigada pela ajuda de vocês, que passou de significativa a indispensável. E finalmente ao Felipe, meu companheiro de todas as horas, meu conselheiro, meu amigo: sabe que eu poderia escrever um capítulo inteiro de agradecimentos só pra você né?! Muito OBRIGADA pela sua inestimável colaboração neste trabalho, sem você, ele não teria ficado assim. O capítulo III é dedicado a você! Enfim, a todos vocês por terem me proporcionado os melhores dias de trabalho, e por fazer do laboratório uma extensão da minha casa. Tenho certeza de que seremos amigos por longos anos e não somente colegas de profissão.

À minha amigona Carol, por todos os momentos de “sofrimento” e “descarrego” que passamos juntas! Por você e o Thiago já serem parte da minha família e por me aturarem a tantos anos. Eu te amo sua lorpa!

Em especial à família que não é de “sangue”, mas que desempenha esse papel desde que iniciei a graduação: Jana, Rô, Pedro, Tiele e Cris, que SEMPRE estiveram do meu lado me apoiando, incansáveis. Principalmente ao Rô por dividir apartamento, ter a paciência que só ele sabe ter comigo e compreender que eu não sou sorridente o tempo inteiro. Aos novos componentes da família: Malusinha e Bernardo que já fazem parte do meu mundo. À Tai e ao João pelas tantas horas dedicadas ao ócio produtivo. Aos meus colegas queridos do mestrado: Camila, Fernando, Cassiano, André, Rodrigo e Leandro, obrigada pela amizade sincera. À Simone e ao Vargas por se fazerem presentes, mesmo distantes.

Ao CNPq pelo apoio financeiro.

Ao Júnior pelos quase sete anos de cumplicidade. Por você tentar me entender sempre e me fazer te entender (às vezes).

E, finalmente, agradeço às pessoas mais importantes: minha família. Vocês são as minhas jóias, meu esteio, o meu orgulho. Obrigada por todas as horas de pensamento positivo, todos os incentivos, por fazerem minha vida ter um sentido e por, principalmente, acreditarem em mim, de coração. À minha dindinha que só não é mãe por detalhe, pela força sempre! E à Tia Dalva, que está sempre cuidando dos meus. Eu amo vocês!!!

## Resumo

Dissertação de Mestrado  
Programa de Pós-Graduação em Biodiversidade Animal  
Universidade Federal de Santa Maria

### **Análise do Elemento Transponível *copia* em Espécies de *Drosophila***

AUTORA: PALOMA MENEZES RUBIN

ORIENTADORA: VERA LÚCIA DA SILVA VALENTE

CO-ORIENTADOR: ÉLGION L. DA SILVA LORETO

Data e Local da Defesa: Santa Maria, 30 de abril de 2008.

Elementos transponíveis (TEs) são segmentos de DNA que têm a capacidade de mover-se e replicar-se dentro dos genomas. O retrotransposon  *copia*  pertence à superfamília  *copia*  e foi primeiramente seqüenciado em *D. melanogaster*. Abordamos em parte deste trabalho uma busca por seqüências relacionadas ao elemento  *copia*  nos genomas de diversas linhagens de espécies do grupo  *willistoni* . A região estudada corresponde a porção da 5'LTR-URL, escolhida pela sua importância como seqüência regulatória e que pode nos fornecer informações filogenéticas relevantes. Dados  *in silico*  foram adicionados ao trabalho onde buscamos nos doze genomas de *Drosophila* disponíveis a seqüência completa de  *copia*  e a seqüência das LTRs que flanqueiam o elemento. A ampla distribuição do elemento  *copia*  no gênero *Drosophila* sugere indícios da ocorrência de vários casos de transferência horizontal, um deles entre *D. melanogaster* e *D. willistoni*. Para investigarmos esse caso amplificamos por PCR a região 5'LTR-URL de duas linhagens de *D. willistoni*, que apresentaram de 95 a 98% de similaridade com o elemento  *copia* , porém a mesma homologia não foi detectada em rastreamentos por *Southern Blot*. Algumas hipóteses podem ser levadas em consideração para explicarmos tais resultados: o polimorfismo de  *copia*  nos genomas hospedeiros, pequeno número de cópias em poucos indivíduos da população ou a seqüência do elemento  *copia*  estar em um vetor. No momento não podemos excluir nenhuma dessas hipóteses. Os resultados das buscas nos genomas seqüenciados mostraram uma ampla distribuição do elemento  *copia* , porém desigual, e algumas incongruências foram encontradas com relação à análise filogenética das espécies hospedeiras. Das espécies analisada em somente três do subgênero *Sophophora* e duas do subgênero *Drosophila* não foram encontradas seqüências relacionadas a  *copia* .

Palavras-chave: retrotransposons, filogenia, genomas  *in silico* .

## Abstract

Dissertação de Mestrado  
Programa de Pós-Graduação em Biodiversidade Animal  
Universidade Federal de Santa Maria

### **Analysis of *copia* transposable element in *Drosophila* species**

AUTORA: PALOMA MENEZES RUBIN

ORIENTADORA: VERA LÚCIA DA SILVA VALENTE

CO-ORIENTADOR: ÉLGION L. DA SILVA LORETO

Data e Local da Defesa: Santa Maria, 30 de abril de 2008.

Transposable elements (TEs) are segments of DNA that have the ability to move up and replicate themselves within the genome. The retrotransposon *copia* belongs to the superfamily *copia* and was first sequenced in *D. melanogaster*. We used in part of this work, which corresponded to a populational analysis of element *copia* in the genomes of species of the group *willistoni*, the portion of 5'LTR-URL for its importance as a regulatory sequence that can give us relevant phylogenetic information. Data *in silico* were added to the work which sought in the twelve genomes of *Drosophila* available the complete sequence of the element *copia* and the sequence of LTRs that surround the element. The wide distribution of element *copia* in the genus *Drosophila* suggests the occurrence of several cases of horizontal transfer, one of them between *D. melanogaster* and *D. willistoni*. To investigate this case we amplify by PCR the region 5'LTR-URL of two strains of *D. willistoni*, which showed a 95 to 98% similarity with the element *copia*, but the same homology was not detected in the search made by *Southern blot*. Some assumptions can be considered to explain these results: the *copia* polymorphism in the host genome, a small number of copies in a few individuals of the population or the sequence of the *copia* element could be in a vector. But the data are still inconclusive. The results of the searches showed a wide distribution of element *copia* in the genome, despite uneven, and some inconsistencies were found related to the phylogenetic analysis of the host species. Of the species examined, only three of subgenus *Sophophora* and two of subgenus *Drosophila* did not show sequences related to *copia* element.

Keywords: retrotransposons, phylogeny, genomes *in silico*.

# SUMÁRIO

<b>CAPÍTULO I</b> .....	<b>10</b>
<b>1. Introdução Geral</b> .....	<b>11</b>
<b>1.1. Elementos de Transposição</b> .....	<b>11</b>
1.1.1. Classificação dos TEs .....	13
1.1.2. Os Elementos da Classe I .....	14
1.1.3. Evolução dos TEs .....	16
1.1.4. Implicações Evolutivas dos Elementos de Transposição .....	17
1.1.5. O elemento <i>copia</i> .....	19
<b>1.2. O Gênero <i>Drosophila</i></b> .....	<b>20</b>
<b>1.3. Distribuição do Elemento <i>copia</i> em <i>Drosophila</i></b> .....	<b>21</b>
<b>CAPÍTULO II</b>	
<b>Investigação da Dinâmica Evolutiva do Elemento Transponível <i>copia</i> em espécies do grupo <i>willistoni</i> de <i>Drosophila</i></b> .....	<b>25</b>
<b>1. Introdução</b> .....	<b>26</b>
<b>1.1. O grupo <i>willistoni</i></b> .....	<b>26</b>
1.1.1. O subgrupo <i>willistoni</i> .....	27
1.1.2. Distribuição Geográfica .....	30
<b>2. Objetivo Geral</b> .....	<b>34</b>
2.1. Objetivos Específicos .....	34
<b>3. Materiais e Métodos</b> .....	<b>34</b>
3.1. Espécies Investigadas .....	34
3.2. Manutenção dos estoques de <i>Drosophila</i> .....	35
3.3. Extração de DNA .....	36
3.4. PCR e Clonagem .....	36
3.5. Seqüenciamento de DNA .....	37
3.6. Análise por <i>Southern Blot</i> .....	38
<b>4. Resultados e Discussão</b> .....	<b>38</b>
<b>CAPÍTULO III</b>	
<b>O Retrotransposon <i>copia</i> nos genomas disponíveis de <i>Drosophila</i>: uma análise de sua distribuição e história evolutiva</b> .....	<b>44</b>
<b>1. Justificativa</b> .....	<b>45</b>
<b>2. Objetivo Geral</b> .....	<b>45</b>

2.1. Objetivos Específicos .....	45
3. Materiais e Métodos .....	46
4. Resultados e Discussão .....	47
4.1. Relações Filogenéticas do Elemento <i>copía</i> X Relações Filogenéticas das Espécies Hospedeiras .....	55
<b>CAPÍTULO IV</b>	
Considerações Finais e Perspectivas .....	62
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	65

# **CAPÍTULO I**

# 1. Introdução Geral

## 1.1. Elementos de Transposição

Os elementos de transposição são seqüências de DNA incluídas na fração de DNA medianamente repetida do genoma que possuem a capacidade intrínseca de mudar sua localização. Eles compartilham as seguintes características: causam duplicação dos sítios de inserção, apresentam polimorfismo de sítios de inserção e apresentam variabilidade no número de cópias dentro e entre espécies (Capy *et al.*, 1998).

Na década de 1970, experimentos levaram à redescoberta desses elementos, que podem gerar mutações em células somáticas e germinativas. Esses experimentos definitivamente identificaram os “elementos controladores” sugeridos por Barbara McClintock, na década de 40, agora chamados de elementos transponíveis (TEs – Transposable Elements), como importantes constituintes dos genomas (revisão em Biémont e Vieira, 2006).

Os TEs podem apresentar em sua seqüência todas as instruções necessárias para sua transposição e cada hospedeiro pode ter muitos tipos diferentes de TEs que podem representar parte considerável de seus genomas. Os elementos transponíveis infectam os genomas hospedeiros por possuírem habilidade para se mover e replicar, gerando assim plasticidade. Embora os principais grupos de TEs sejam antigos e estejam presentes em todos os reinos, eles exibem extrema diversidade: há diversas famílias de TEs em plantas, que chegam a constituir 80% do seu DNA genômico total (SanMiguel e Bennetzen, 1998; Morgante, 2005 e Wicker *et al.*, 2007). Por exemplo, estima-se que 22% do genoma de *Drosophila melanogaster* (Kapitonov e Jurka, 2003), mais de 50% do genoma do milho (SanMiguel *et al.*, 1996) e 45% do genoma humano (*International Human Genome Sequencing Consortium*, 2001) são constituídos por TEs. Segundo Wicker *et al.* (2007), os TEs foram encontrados em todas as espécies estudadas até o momento, com exceção do *Plasmodium falciparum* e

provavelmente em espécies estritamente relacionadas a ele. Apesar de serem menos abundantes do que em plantas, os TEs em fungos e metazoários também representam uma parte significativa desses genomas (3-20% em fungos e 3-45% em metazoários), incluindo membros de muitas superfamílias (Daboussi e Capy, 2003; Hua-Van *et al.*, 2005 e Wicker *et al.*, 2007). No genoma de *D. melanogaster* foram identificadas mais de 130 famílias diferentes de TEs (Kapitonov e Jurka, 2003), com o número de cópias variando de uma a 146 (Kaminker *et al.*, 2002).

No final da década de 70 e início da década de 80, alguns autores consideravam os TEs como parasitas moleculares, que invadem novas espécies, se multiplicam dentro dos genomas (hospedeiros) e sobrevivem por longos períodos de tempo evolutivo e sua natureza egoísta impede que eles desempenhem um papel importante na evolução dos genomas. Entretanto, muitos estudos têm mudado essa visão, revelando as interações dos TEs, seu impacto na função e estrutura do genoma e na evolução e adaptação das populações (revisão em Becker e Lönnig, 2001; Kidwell e Lisch, 2001; Bowen e Jordan, 2002; Brookfield, 2005).

A abundância e a ubiquidade dessas seqüências geram uma série de questões sobre a natureza das relações entre TEs e seus hospedeiros, bem como sobre as implicações desses elementos na evolução dos genomas (Kidwell e Lisch, 2001). Embora seja aceito que os TEs podem comportar-se como parasitas genômicos e produzirem mutações neutras e deletérias, algumas mutações e mudanças genômicas, causadas pelos transposons, têm dado contribuições significativas à evolução dos seus hospedeiros. Um acúmulo de evidências moleculares tem demonstrado uma grande quantidade de maneiras pelas quais os TEs podem afetar a evolução dos organismos nos quais eles residem (McDonald, 1993; Kidwell e Lisch, 2001; Becker e Lönnig, 2001; Volff, 2006; Xing *et al.*, 2006).

### 1.1.1. Classificação dos TEs

Em 1989, Finnegan propôs o primeiro sistema de classificação para os TEs, que distinguia duas classes com base na sua intermediação da transposição: RNA (Classe I ou Retrotransposons) ou DNA (Classe II ou Transposons de DNA). O mecanismo de transposição da Classe I, comumente chamado “copiar e colar”, possui uma molécula de RNA como intermediário no processo; já o mecanismo de transposição da Classe II é chamado de “cortar e colar”, e o intermediário do processo é uma molécula de DNA.

Um novo modelo de classificação foi sugerido por Wicker *et al* (2007) mantendo as duas classes e aplicando critérios que utilizam os mecanismos de transposição e mecanismos enzimáticos. Segundo essa abordagem, os TEs continuam sendo separados em duas classes principais, e seguem a ordem hierárquica de taxonomia:

- Subclasse: para distinguir os TEs de Classe II que fazem cópias próprias para a sua inserção, dos que saem e deixam um sítio de doação para se inserirem em outros locais.
- Ordem: agrupa as maiores diferenças dos mecanismos de inserção, e conseqüentemente, de organização e enzimáticos.
- Superfamília: as superfamílias dentro de uma ordem compartilham a mesma estratégia de replicação, mas são distinguidas por características como a estrutura da proteína ou de domínios não codificantes. Além disso, diferem na presença e tamanho do sítio de duplicação que produzem no DNA alvo (TSD – Target Site Duplication), uma pequena repetição direta, que é gerada na região flangeadora do transposon, a partir de uma inserção.
- Família: são definidas pela conservação da seqüência de DNA. A similaridade em nível de proteína é geralmente elevada entre as famílias diferentes que pertencem à mesma superfamília.
- Subfamília: as subfamílias são definidas com base em dados filogenéticos, e pode, em casos específicos, servir para distinguir populações autônomas e não-autônomas internamente homogêneas.

### 1.1.2. Os Elementos de Classe I

Os TEs da classe I se transpõem por intermédio do RNA. Essa classe não possui subclasses, pois nenhum membro corta ou transfere pedaços do DNA para outro local. Ao invés disso, um RNA intermediário é transcrito como uma cópia genômica, e, em seguida, é transcrito reversamente para DNA pela enzima transcriptase reversa (RT, do inglês: Reverse Transcriptase), codificada pelo próprio elemento. Cada ciclo completo de replicação produz uma nova cópia do TE. Conseqüentemente os retrotransposons são os maiores contribuintes da fração repetida de DNA nos grandes genomas. Os retrotransposons podem ser divididos em cinco ordens, com base na sua organização, características do mecanismo de inserção e filogenia da RT: retrotransposons de LTRs, elementos *DIRS-like*, elementos *Penelope-like* (PLEs), LINEs e SINEs. Os retrotransposons de LTRs são menos abundantes em animais, mas é a ordem predominante em plantas. Eles variam de pouco mais de 100 pb até, excepcionalmente, 25 kb. As LTRs (do inglês: Long Terminal Repeats – longas repetições terminais) flanqueiam os elementos variando de pouco mais de 100 pb até mais de 5 kb, e iniciam sua seqüência com 5'–TG–3' e finalizam com 5'–CA–3'. Logo após sua integração ao DNA alvo, as LTRs produzem um TSD de 4-6 pb. Esses elementos geralmente contêm fases abertas de leitura (ORFs, do inglês: Open Reading Frames) para o gene *gag*, para partículas semelhantes a proteínas de estrutura viral e para o gene *pol* (do inglês: Polymerase). O gene *pol* codifica as proteínas requeridas para transposição, entre elas: proteinase aspártica (AP), RT, RNase H (RH) e integrase (INT). Ocasionalmente há a adição de uma ORF sem função definida. As duas principais superfamílias que ocorrem em metazoários são *copia* (superfamília do elemento *copia*, foco deste trabalho) e Gypsy. Estas superfamílias diferem na ordem da RT e INT do gene *pol*. Todas as superfamílias da ordem retrotransposons de LTRs usam um mecanismo similar de transposição (Wicker *et al.*, 2007).

O gene *gag* codifica uma proteína estrutural interna de vírus, a proteína Gag (do inglês: Group-specific Anti Gen), que é proteoliticamente processada em três proteínas maduras: matriz (MA), capsídeo (CA) e nucleocapsídeo

(NC), correspondendo aos componentes do capsídeo viral (Boeke e Stoye, 1997).

A Figura 1 mostra a organização estrutural de um retrotransposon de LTR, da superfamília *copia* modificado de Wicker *et al.* (2007):

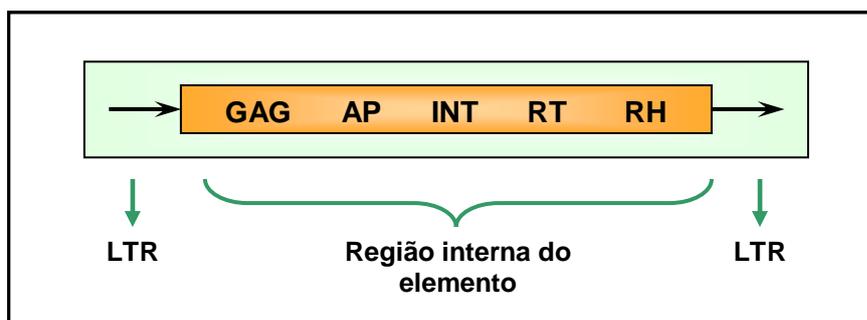


Figura 1: Organização estrutural dos TEs da superfamília *copia*. As flechas (em preto) correspondem à LTRs (Long Terminal Repeats) que flanqueiam os elementos da superfamília. No retângulo em laranja a representação da região interna dos elementos da superfamília (GAG, gene *gag*; AP, proteinase aspártica; INT, integrase; RT, transcriptase reversa e RH, RNase H).

As LTRs contêm normalmente seqüências reguladoras, tanto promotoras como terminadoras da transcrição do elemento e, ainda, sinais específicos para empacotamento, dimerização e transcrição reversa (Kumar e Bennetzen, 1999 e Wiker *et al.*, 2007).

Segundo SanMiguel *et al.* (1998), uma das características das LTRs é que ambas têm seqüências idênticas no momento da inserção do elemento. Com o tempo elas vão acumulando mutações e o número das substituições entre ambas pode servir para datar o momento da inserção.

As LTRs apresentam três domínios distintos: U3, R e U5 (Figura 2). A transcrição do elemento começa no sítio de iniciação da transcrição da 5'LTR, que ocorre entre os domínios U3 e R. O final da transcrição acontece na 3'LTR onde está o sítio de poliadenilação funcional, que é delimitado pelos domínios R e U5. O domínio U3 contém uma seqüência promotora para a RNA polimerase II, presente em ambas as LTRs, mas somente atua como forte promotor na 5'LTR. A seqüência da 3'LTR poderia atuar como um promotor fraco para os genes adjacentes no local da inserção (White *et al.*, 1994).

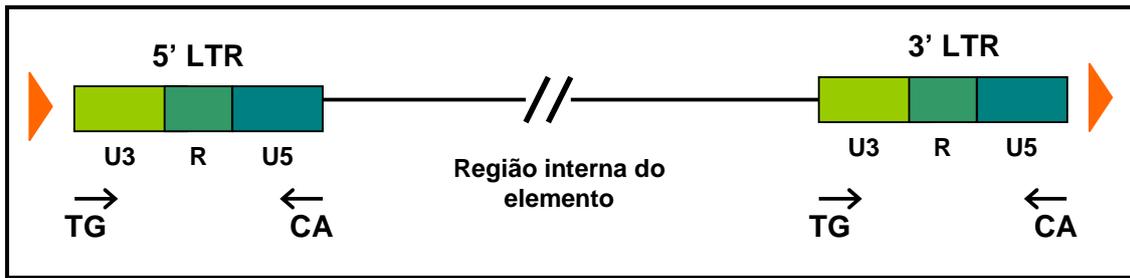


Figura 2: Organização geral das LTRs de um retrotransposon da superfamília *copia*. Os triângulos (em laranja) mostram a orientação do elemento (5' – 3'). Os retângulos (em verde) ilustram a organização das LTRs 5' e 3' que flanqueiam a região interna do elemento, com seus respectivos domínios U3, R e U5. A orientação das seqüências iniciadoras TG e terminadoras CA, também é ilustrada na figura (White *et al.*, 1994).

### 1.1.3. Evolução dos TEs

A origem dos TEs dentro de uma espécie pode ser resultado de transmissão vertical de cópias presentes em uma espécie ancestral durante o processo de especiação ou de transmissão horizontal, que pode ser definida como o processo pelo qual genes podem se mover entre espécies isoladas reprodutivamente. Transmissão horizontal tem sido tradicionalmente inferida quando as relações filogenéticas dos TEs não estão de acordo com a história evolutiva das espécies hospedeiras. Existem três métodos utilizados para detectar transmissão horizontal de TEs: (1) detecção de elementos com alto grau de similaridade de seqüência em táxons não relacionados e com divergência entre TEs muito menor do que a divergência entre genes nucleares das suas respectivas espécies hospedeiras; (2) detecção de diferenças topológicas entre a filogenia do TE e das espécies hospedeiras; (3) distribuição descontínua dos elementos entre táxons proximamente relacionados. Assim, as filogenias obtidas de TEs são freqüentemente comparadas com filogenias de suas espécies hospedeiras, na tentativa de entender a evolução destas seqüências móveis. As congruências entre as duas filogenias são encaradas como indicativos de transmissão vertical dos elementos, enquanto que as discrepâncias podem ser indicativos de transmissão horizontal (Silva *et al.* 2004 e revisão em Loreto *et al.*, 2008).

Entretanto, explicações alternativas precisam ser testadas, como polimorfismo ancestral com independente distribuição de cópias nas espécies descendentes, taxas diferentes de substituição em TEs nas diferentes espécies e perda estocástica de TEs em alguns táxons (Capy *et al.*, 1994).

Vários casos de transmissão horizontal de elementos entre espécies são relatados para Classe I (entre eles *copia*). No entanto, o mecanismo pelo qual a transmissão horizontal ocorre permanece obscuro. Alguns mecanismos e potenciais vetores dessa transferência são: alguns retrotransposons que produzem partículas vírus-*like* (VLPs), vírus, alguns parasitas ou parasitóides (vespas parasitas, ácaros), bactérias endossimbiontes e bacteriófagos (Silva *et al.* 2004 e revisão em Loreto *et al.*, 2008).

#### 1.1.4. Implicações Evolutivas dos Elementos de Transposição

Os elementos de transposição são apontados como geradores de variabilidade, podendo causar um repertório de efeitos mutacionais, desde a simples interrupção de seqüências codificantes, resultando em fenótipos nulos ou letais, até criar um novo padrão de expressão fenotípica, ao se inserir em uma região regulatória de um gene (Ko Harada *et al.*, 1990).

Assim, os elementos de transposição estão relacionados com a hipermutabilidade, a reestruturação dos genomas ao promoverem rearranjos cromossômicos, e a transmissão horizontal de informação genética. Todos estes mecanismos podem ser de grande importância como fonte de variabilidade para a evolução (Flavell *et al.*, 1980).

Devido a sua capacidade de transposição, os TEs são importantes agentes mutagênicos, com potencial para causar uma variedade de mudanças nos genomas de seus organismos hospedeiros. Espera-se que a maioria das novas mutações provocadas por TEs sejam deletérias para seu hospedeiro, assim como são as mutações produzidas por qualquer agente mutagênico. Entretanto, estão sendo descobertos muitos exemplos de interações TE -

hospedeiro proporcionando benefícios para os hospedeiros (Kidwell e Lisch, 2002).

Muitos exemplos de domesticação molecular têm sido descritos, onde os TEs são recrutados para desempenhar uma função no hospedeiro (McDonald, 1993; Britten, 1996; Nekrutenko e Li 2000; Kidwell e Lisch, 2001; Bundock e Hooykaas, 2005; Hammer *et al.*, 2005). Os TEs também podem agir como carregadores móveis de elementos regulatórios, como promotores, e ao integrarem-se dentro ou próximo de genes eles podem contribuir para uma diversificação funcional, alterando a expressão desses genes. Quatro categorias de recrutamento de TEs são descritas gerando: funções regulatórias, funções estruturais, funções enzimáticas e novas seqüências codificadoras (Kidwell e Lisch, 2001).

Segundo Kidwell e Lisch (2001), embora o conceito de TE como parasita molecular seja útil para o entendimento da natureza das relações hospedeiro-parasita, esta abordagem deveria ser expandida para considerar um contínuo de interações entre elementos e seus hospedeiros, variando de parasitismo em um extremo a mutualismo ou simbiose em outro.

Quando os elementos de transposição apresentam alta taxa de transposição, eles devem interferir no perfil do hospedeiro. Em função disto, os eventos de transposição parecem ser limitados no tempo e células ou tecidos específicos (Becker e Lönnig, 2001). A regulação da transposição é um mecanismo complexo e específico de cada elemento e inclui regulação de sua transcrição e tradução. Os próprios TEs possuem seqüências regulatórias que modulam sua transcrição. Além disso, fatores dos hospedeiros parecem ser importantes nessa regulação, como proteínas repressoras ou ativadoras e também mecanismos epigenéticos, como co-supressão mediada por RNA de interferência e metilação (Lozovskaya *et al.*, 1995; Labrador e Corces, 1997; Hurst e Werren, 2001; Castro e Carareto, 2004). A atividade dos TEs pode ainda ser induzida por fatores ambientais, em particular por estresse (Capy *et al.*, 2000).

### 1.1.5. O Elemento *copia*

Foi assim denominado por transcrever uma quantidade significativa de cópias de mRNA em *D. melanogaster* (Flavell *et al.*, 1981). O retrotransposon *copia* pertence à superfamília *copia* (Wicker *et al.*, 2007), e foi primeiramente seqüenciado em *D. melanogaster*. Identificado no alelo *white-apricot* do locus *white* deste genoma, possui aproximadamente 5,4 kb de comprimento e duas LTRs de 276 pb (na mesma orientação do elemento) flanqueando uma única ORF de 4.227 nucleotídeos, capaz de codificar uma poliproteína de 1.409 aminoácidos. Essa poliproteína gera produtos similares aos genes *gag* e *pol* dos retrovírus (Mount e Rubin, 1985).

Segundo Jordan e McDonald (1998), este elemento é ideal para se estudar a evolução dos retrotransposons, pois já é bem conhecido a nível molecular, e a sua distribuição no gênero *Drosophila* é razoavelmente caracterizada. Além disso, *copia* é encontrado em representantes de todas as principais radiações de *Drosophila*, o que sugere que tenha surgido cedo na história dos drosofilídeos (Martin *et al.*, 1983).

Abordamos em parte do trabalho, que correspondeu a uma análise populacional do elemento *copia* nos genomas de espécies do grupo *willistoni*, a porção da 5'LTR e sua região promotora que não é traduzida (URL do inglês: Untranslated Leaders Region) pela grande importância dessa seqüência regulatória por fornecer informações filogenéticas, bem como compreender a regulação dos retrotransposons (Capítulo II).

A extremidade 5'LTR-URL do elemento *copia* (Figura 3) contém seqüências responsáveis por controlar sua transcrição, atuando diretamente no processo de retrotransposição (Jordan e McDonald, 1998).

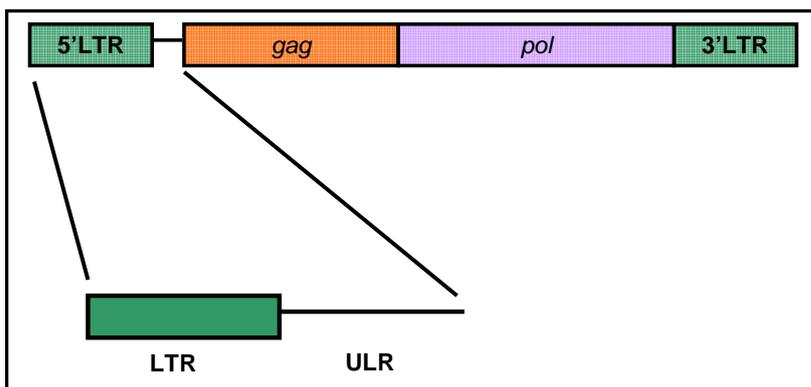


Figura 3: Região 5'LTR-URL, modificada de Jordan e McDonald (1998), com indicação da região estudada no presente trabalho. Os retângulos em verde correspondem à região das LTRs. Também está ilustrada a ORF (Open Reading Frames) de *copia*: gene *gag* (no retângulo laranja) e gene *pol* (no retângulo roxo).

Dados *in silico* foram adicionados ao trabalho. Buscamos, nos doze genomas de *Drosophila* disponíveis (Clark *et al.*, 2007) a seqüência completa de *copia* e a seqüência das LTRs que flanqueiam o elemento (Capítulo III).

## 1.2. O Gênero *Drosophila*

O gênero *Drosophila* faz parte da família Drosophilidae (ordem Diptera) e se destaca por abranger sozinho cerca de um terço das suas espécies. Encontra-se constituído por 1.149 espécies, subdivididas em oito subgêneros (Bächli, 2008). Dentre estes subgêneros, podem-se destacar os subgêneros *Drosophila* e *Sophophora* pela diversidade de espécies apresentada. Segundo Russo *et al.* (1995), esses dois subgêneros divergiram há cerca de 40 milhões de anos. Tamura *et al.* (2004), por outro lado ao corrigir distâncias sinônimas para os efeitos de viés na utilização de códons, obtém estimativas mais antigas, da ordem de 63 milhões de anos.

O subgênero *Sophophora*, estabelecido por Sturtevant (1939), compreende 332 espécies, subdivididas em oito grupos (Bächli, 2008) dentre os quais se destacam os grupos *melanogaster*, *obscura*, *saltans* e *willistoni*.

O grupo *melanogaster* é formado por 185 espécies (Bächli, 2008) classificadas em doze subgrupos, os quais se agrupam em três clados

principais: (1) subgrupo *ananassae*, (2) subgrupo *montium* e (3) subgrupo *melanogaster* e demais subgrupos (Kastanis *et al.*, 2003, Yang *et al.* 2004 e Da Lage *et al.* 2007).

Os grupos *saltans* e *willistoni* representam à radiação *Sophophora* ocorrida no Novo Mundo (Throckmorton, 1975). Em parte deste trabalho, utilizamos as espécies do grupo *willistoni* que serão listadas adiante.

O subgênero *Drosophila*, por outro lado, é o maior representante do gênero *Drosophila*, com pelo menos 721 espécies, subdivididas em 43 grupos (Bächli, 2008), as quais podem ser divididas em duas linhagens principais (Throckmorton, 1975): a seção *virilis-repleta* e a seção *quinaria-tripunctata*. Além disso, constitui o maior representante da família *Drosophilidae* em Região Neotropical, onde apresenta, pelo menos, 25 grupos de espécies (Val *et al.* 1981).

### 1.3. A distribuição do Elemento *copia* em *Drosophila*

Em 1999, Biémont e Cizeron publicaram uma revisão sobre a análise da distribuição de 43 TEs em 228 espécies *Drosophila*, dentre eles *copia*. As seqüências de *copia* foram identificadas por PCR e *Southern Blot* e encontradas em 52 espécies do gênero *Drosophila*: 22 espécies do grupo *melanogaster*, 7 espécies do grupo *willistoni*, 7 espécies do grupo *obscura*, 6 espécies do grupo *saltans*, 2 espécies do grupo *immigrans*, 1 espécie do grupo *mesophragmatica* e uma espécie do grupo *pinicola*. O resultado dessa revisão no grupo *willistoni* pode ser observado na Tabela 1.

Embora o elemento *copia* tenha sido encontrado por análise de *Souther Blot* em diversas espécies de *Drosophila*, apenas em 8 espécies do grupo *melanogaster*, 2 espécies do grupo *repleta*, *D. willistoni* e *Zaprionus tuberculatus* é que foi descrita a seqüência nucleotídica desse elemento ou pelo menos de parte dele (Biémont e Cizeron, 1999).

Tabela 1: Presença do TE *copia* em espécies do grupo *willistoni* (modificado de Biémont e Cizeron, 1999). A tabela mostra que *copia* foi encontrado por análise de *Souther Blot* nas espécies: *D. capricorni*, *D. fumipennis*, *D. nebulosa*, *D. succinea*, *D. equinoxialis*, *D. paulistorum*, *D. tropicalis* e *D. willistoni* que aparecem marcadas com +. A ausência do elemento (-) foi detectada em *D. bocainensis*, *D. nebulosa*, *D. insularis* e *D. pavlovskiana*.

Grupo	Subgrupo	Espécie	Presença de <i>copia</i>
<i>willistoni</i>	<i>Bocainensis</i>	<i>D. bocainensis</i>	-
		<i>D. capricorni</i>	+
		<i>D. fumipennis</i>	+
		<i>D. nebulosa</i>	+/-*
		<i>D. succinea</i>	+
	<i>Willistoni</i>	<i>D. equinoxialis</i>	+
		<i>D. insularis</i>	-
		<i>D. paulistorum</i>	+
		<i>D. pavlovskiana</i>	-
		<i>D. tropicalis</i>	+
	<i>D. willistoni</i>	+*	

\*Ausência de marcação *in situ* para cromossomos politênicos (Biémont e Cizeron, 1999).

Na busca por uma análise filogenética do elemento *copia* Jordan e McDonald (1998) descreveram duas famílias de *copia* para o gênero *Drosophila* e três subfamílias do TE no subgrupo *melanogaster* (*D. melanogaster*, *D. simulans*, *D. mauritiana*, *D. sechellia*, *D. yakuba*, *D. teissieri* e *D. erecta*) com base nas diferenças das seqüências das LTRs nas espécies.

Em 2006, Almeida e Carareto realizaram um novo estudo baseado em 24 espécies de *Drosophila* incluindo seqüências disponíveis no banco de dados do NCBI. Segundo as autoras, estudos da filogenia desse elemento no gênero *Drosophila* (grupo *repleta*) são difíceis de realizar, não somente por causa do baixo número de seqüências no grupo, mas também por serem seqüências parciais, a maioria delas da 5'LTR-URL. Este estudo reforçou a idéia de Jordan e McDonald (1998) que previam 2 grupos de seqüências de *copia* no gênero *Drosophila*: família *melanogaster* (dividida em 3 subfamílias) e família *repleta* (baseada em 2 espécies do grupo *repleta*) como mostra a árvore filogenética na Figura 5. As autoras ainda argumentam que devido a grande distribuição, e a ocorrência de heterogeneidade

de *copia* nos subgêneros *Drosophila* e *Sophophora* podem ser explicados por um ancestral comum do TE, e que cópias desse retrotransposon foram transmitidas verticalmente ao longo do tempo. Por outro lado, valores muito baixos de distâncias evolutivas foram encontrados em espécies de gêneros diferentes (do grupo *melanogaster* e *willistoni*, e *Z. tuberculatus*) e entre os diferentes grupos de espécies dentro do subgênero *Sophophora* (grupos *willistoni* e *melanogaster*), que pode ser indicativo de transferência horizontal entre *D. willistoni* / *Z. tuberculatus* / espécies do subgrupo *melanogaster*, porém, tal hipótese não pôde ser definitivamente comprovada pelas autoras.

Vários casos de transferência horizontal foram descritos para *copia*: entre espécies do subgrupo *melanogaster* (Jordan e McDonald, 1998), entre *D. melanogaster* e *D. simulans* (Sánchez-Gracia *et al.*, 2005) e ainda entre *D. melanogaster* e *D. willistoni* (Jordan *et al.*, 1999).

Vale salientar que, segundo Clark *et al.* (2007) o genoma de *D. willistoni* possui um dos maiores índices da presença de TEs, se comparado com os outros onze genomas seqüenciados de *Drosophila*.

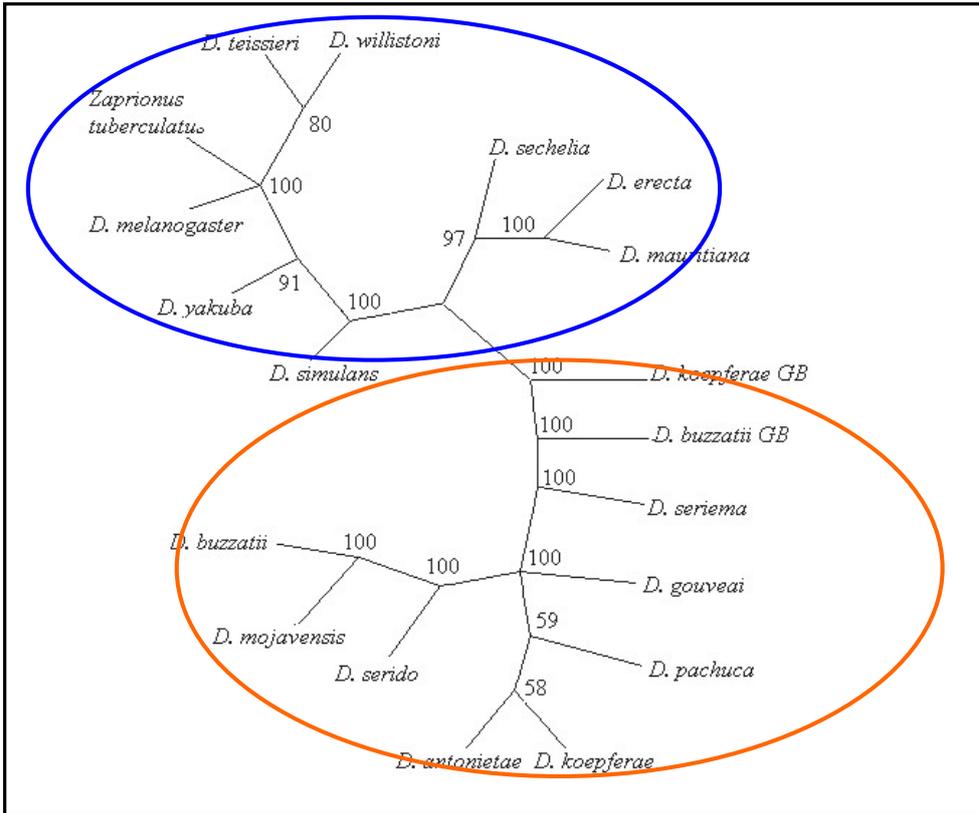


Figura 4: Árvore filogenética gerada a partir de seqüências da 5’LTR-URL segundo Almeida e Carareto (2006). A separação em duas famílias do elemento  *copia*  é evidenciado na elipse azul (seqüências das espécies do grupo  *melanogaster*  –  *D. sechellia* ,  *D. erecta* ,  *D. mauritiana* ,  *D. simulans* ,  *D. yakuba* ,  *D. melanogaster*  e  *D. teissieri*  –  *Z. tuberculatus*  e  *D. willistoni* ) e na elipse laranja (seqüências das espécies do grupo  *repleta* :  *D. serido* ,  *D. buzzatii* ,  *D. mojavensis* ,  *D. gouveai* ,  *D. seriema* ,  *D. koepferae* ,  *D. antonietae*  e  *D. pachuca* ).

## **CAPÍTULO II**

**Investigação da Dinâmica Evolutiva do Elemento  
Transponível  *copia* em espécies do grupo *willistoni* de  
*Drosophila***

# Investigação da Dinâmica Evolutiva do Elemento Transponível *copia* em espécies do grupo *willistoni* de *Drosophila*

## 1. Introdução

A ampla distribuição do elemento *copia* no gênero *Drosophila* sugere indícios da ocorrência de vários casos de transferência horizontal, um deles entre *D. melanogaster* e *D. willistoni* (Jordan *et al.*, 1999). Nesse caso, os autores observaram maior similaridade entre as LTRs do elemento *copia* de *D. melanogaster* e *D. willistoni* (99%) do que entre as seqüências *copia* de *D. melanogaster* e de sua espécie críptica de *D. simulans* (90%) o que, segundo os autores, seria indicativo da ocorrência do evento de transferência horizontal.

O presente trabalho foi estruturado buscando maiores informações sobre esse evento de transferência horizontal. Tendo em vista que, somente uma espécie do grupo *willistoni* (*D. willistoni*) foi investigada pelos autores, ao investigarmos as outras espécies do grupo, poderemos inferir com maior clareza a partir de qual genoma se deu à transferência horizontal, fortalecendo a dinâmica evolutiva deste grupo de espécies e do elemento *copia* nos drosofilídeos.

### 1.1. O Grupo *willistoni*

O grupo *willistoni* de *Drosophila* vem sendo estudado desde a década de 1940, contribuindo significativamente para a teoria evolutiva. A origem deste grupo deve ter ocorrido na América do Sul, talvez após uma migração a partir da América do Norte (Throckmorton, 1975).

Este grupo pertence ao subgênero *Sophophora* e é constituído por 23 espécies, subdivididas atualmente em três subgrupos (Bächli, 2008): *alagitans* (5 espécies), *bocainensis* (12 espécies) e *willistoni* (6 espécies).

As espécies abaixo fazem parte do grupo *willistoni*, e, estão aqui listadas por ser parte do presente estudo, na ordem cronológica em que foram descritas (revisão em Ehrman e Powell, 1982):

- *Drosophila nebulosa* Sturtevant 1916.
- *Drosophila willistoni* Sturtevant 1921.
- *Drosophila equinoxialis* Dobzhansky 1946.
- *Drosophila paulistorum* Dobzhansky e Pavan (Burla *et al.* 1949).
  - semi-espécie Andino-Brasileira Dobzhansky e Spassky 1959.
  - semi-espécie Orinocana Dobzhansky e Spassky 1959.
- *Drosophila insularis* Dobzhansky *et al.* 1957.

Fazem parte do grupo *willistoni* seis espécies crípticas: *D. willistoni*, *D. equinoxialis*, *D. tropicalis*, *D. insularis*, *D. pavlovskiana* e *D. paulistorum*, que são, em geral, muito similares em sua morfologia externa, sendo que os métodos utilizados em sua identificação restringem-se a inspeção da genitália dos machos (Spassky, 1957), ao estudo detalhado dos cromossomos politênicos (Rohde *et al.*, 2006) e a testes de cruzamento (Spassky *et al.*, 1971), embora recentemente, Garcia *et al.* (2006) tenham descrito a utilização da variação da mobilidade eletroforética da enzima fosfatase ácida-1 (*Acph1*) na diagnose rápida e eficiente de cinco das seis espécies crípticas pertencentes ao subgrupo.

#### 1.1.1. O subgrupo *willistoni*

Sob o ponto de vista taxonômico o subgrupo da *D. willistoni* está composto por seis espécies crípticas que estão incluídas na seguinte classificação (de acordo com revisão em Ehrman e Powell, 1982):

- Subgrupo: *willistoni*
- Grupo: *willistoni*
- Subgênero: *Sophophora*
- Gênero: *Drosophila*
- Família: Drosophilidae

- Ordem: Diptera
- Classe: Insecta
- Phylum: Arthropoda

A primeira descrição de uma espécie do subgrupo *willistoni* ocorreu em 1896. Dentre uma série de espécies novas coletadas por Samuel Wendell Williston na Ilha de São Vicente, nas pequenas Antilhas, uma foi denominada de *D. pallida*. Entretanto, como esse nome já pertencia a outro drosofilídeo, duas décadas depois, Sturtevant (1916), corrigiu o nome da nova espécie denominando-a de: *D. willistoni*. Em 1943, Dobzhansky e Pavan examinaram indivíduos resultantes de cruzamentos provenientes de descendentes de isolinhagens (fêmeas fecundadas vindas da natureza) de moscas coletadas no Estado de São Paulo e concluíram que naquele local ocorriam duas espécies de drosofilídeos muito semelhantes, sendo que uma delas correspondia à descrição de *D. willistoni*. Impossibilitados de compararem seus materiais com espécimes vivos ou mesmo preservados, a separação das duas espécies foi feita com base na pequena diferença de tamanho corporal entre elas. Assim, os autores supuseram que a espécie mais freqüentemente encontrada e correspondente à de menor tamanho corporal deveria se tratar da *D. willistoni*. A espécie maior recebeu o nome *D. paulista*. Posteriormente, coletas de *D. willistoni* vindas da Guatemala e do México permitiram comparações e cruzamentos que esclareceram que *D. willistoni* era a espécie de maior tamanho. Desse modo, o nome *D. paulista* passou a ser sinônimo de *D. willistoni*. A outra espécie, a de menor tamanho, somente foi descrita em 1949 e recebeu o nome de *D. paulistorum* Dobzhansky e Pavan (Burla *et al.* 1949). Dobzhansky e Mayr (1944), através de experimentos de isolamento sexual, com linhagens de *D. willistoni* oriundas de diferentes regiões geográficas, verificaram que as moscas de Tefé (Amazonas - Brasil) apresentavam isolamento praticamente total com as demais populações. Outros estudos, mais detalhados e de maior duração, confirmaram estes resultados e mostraram que as moscas de Tefé eram incapazes de produzir híbridos com as das demais localidades. Foram analisadas também, pequenas diferenças morfológicas e fisiológicas entre elas.

Estes achados levaram à descrição de uma terceira espécie críptica do subgrupo da *D. willistoni*, a *D. equinoxialis* Dobzhansky (Dobzhansky, 1946). A quarta espécie críptica do grupo *willistoni* foi descrita por Burla *et al.* (1949). Trata-se da *D. tropicalis*, coletada originalmente em Palma, Goiás, Brasil. Em 1954, Townsend

verificou que linhagens de *D. tropicalis* procedentes das Grandes Antilhas geravam descendentes machos híbridos estéreis quando eram cruzadas com linhagens vindas do Continente Sul Americano (Belém, Pará, Içana e Amazonas). Resultados referentes a diferenças na morfologia, citologia e isolamento sexual, além da esterilidade dos machos híbridos sugeriram que as Grandes Antilhas eram habitadas por uma subespécie diferente da que ocupa o continente. Essa subespécie recebeu o nome de *D. tropicalis cubana* Townsend (Townsend, 1954). Resultados de cruzamentos envolvendo linhagens provenientes de Palma (Goiás), de Tefé (Amazonas), de Trinidad e de Cuba, realizados por Winge (1965), levaram a proposta de elevação de *D. tropicalis cubana* para o nível de espécie, devido ao grau de isolamento reprodutivo e desapareamento dos cromossomos híbridos (Winge, 1965). Essa proposta, porém, não foi formalizada por uma descrição detalhada e, assim, a *D. tropicalis cubana* é considerada ainda uma subespécie de *D. tropicalis*.

Dobzhansky (1957), em um estudo sobre a variabilidade cromossômica de *D. willistoni* da América Central e das Antilhas, apresentou um mapa da área de distribuição geográfica das espécies crípticas do subgrupo da *D. willistoni* conhecidas (Figura 4). Nesse mesmo trabalho, também foi descrita uma nova espécie críptica do grupo, restrita a algumas das ilhas das Pequenas Antilhas. Esta nova espécie foi denominada de *D. insularis* (Dobzhansky 1957). O trabalho de Dobzhansky e Spassky (1959) demonstrou que *D. paulistorum*, por sua vez, é na realidade uma superespécie, constituída de pelo menos seis raças ou semi-espécies. Durante o estudo desse complexo foi verificado que o que era considerado inicialmente “raça Guiana”, apresentava, além de diferenças citológicas, em comparação com as demais raças, também isolamento reprodutivo suficiente para ser elevada ao *status* de espécie. Conseqüentemente, a última espécie do grupo críptico *willistoni* descrita recebeu o nome de *D. pavlovskiana* Kastritsis e Dobzhansky (Kastritsis e Dobzhansky, 1967). Assim, o subgrupo *willistoni* é composto por seis espécies crípticas: *Drosophila willistoni*, *D. paulistorum*, *D. equinoxialis*, *D. tropicalis*, *D. insularis* e *D. pavlovskiana*.

### 1.1.2. Distribuição geográfica

A distribuição geográfica das espécies que compõem o subgrupo da *D. willistoni*, que pode ser visualizada na Figura 4, é essencialmente Neotropical e foi revisada nos trabalhos de Winge (1971), Spassky *et al.* (1971), Dobzhansky e Powell (1975), Ehrman e Powell (1982) e Santos e Valente (1990). A espécie *D. willistoni* é a de maior distribuição geográfica do grupo e pode ser encontrada desde o sul da Flórida/EUA e centro norte do México, até o norte da Argentina, em uma grande área compreendida aproximadamente entre os paralelos 28°N e 37°S. Essa espécie também foi encontrada nas Bahamas, em toda a América Central, inclusive Grandes e Pequenas Antilhas e na América do Sul, especialmente a leste da cordilheira dos Andes, sendo talvez a espécie mais comum em toda a região equatorial, na Bolívia, no sul do Brasil, Uruguai e norte da Argentina.

*D. paulistorum* tem, assim como as demais espécies crípticas, sua área de distribuição geográfica incluída na de *D. willistoni*. É a segunda espécie com maior amplitude de área geográfica ocupada, estendendo-se desde o norte da Guatemala, na localidade de Tikal, ao sul do paralelo 18°N, até Porto Alegre no Estado do Rio Grande do Sul/Brasil, ao sul do paralelo 30°S. Entre esses dois limites, foi coletada na América Central (istmo) e América do Sul, principalmente a leste dos Andes, sendo muito comum em toda a região equatorial, e também no Peru e na Bolívia. A oeste da cordilheira foi encontrada apenas na Colômbia e no Equador. Essa espécie não habita as Antilhas, mas é encontrada na ilha de Trinidad (Spassky *et al.*, 1971; Winge, 1971; Dobzhansky e Powell, 1975; Ehrman e Powell, 1982; Santos e Valente, 1990).

A área ocupada por *D. equinoxialis* estende-se desde as Grandes Antilhas e sul do México, até quase o centro do Brasil. No México essa espécie foi encontrada em Cuernavaca, Estado de Morelos, pouco abaixo do paralelo 19°N. Nas grandes Antilhas foi encontrada em Cuba, Jamaica, Grand Cayman, Haiti e Porto Rico, mas não foi coletada nas Pequenas Antilhas. Além de bem distribuída em toda a América Central, essa espécie é também encontrada na América do Sul na porção oeste dos Andes, na Colômbia e Equador; tendo sido registrada em altas frequências a leste da Cordilheira, como na Venezuela, Guiana, ilhas de Trinidad, Peru e Brasil, mais

precisamente nos Estados do Amazonas, Acre, Roraima, Tocantins, Rondônia, Pará, Maranhão e Norte de Goiás. O ponto mais ao sul de distribuição da espécie parece ser um pouco abaixo do paralelo 12°S, no Estado de Tocantins (Spassky *et al.*, 1971; Winge, 1971; Dobzhansky e Powell, 1975; Ehrman e Powell, 1982; Santos e Valente, 1990).

O extremo norte de distribuição de *D. tropicalis* encontra-se em Tikal, Guatemala, ao sul do paralelo 18°N. Nas Grandes Antilhas foi coletada em Cuba, Jamaica, Grand Cayman, Haiti, República Dominicana e Porto Rico. As Pequenas Antilhas não abrigam essa espécie. A área ocupada também inclui a América Central, centro da América do Sul, e localidades a oeste dos Andes como a Colômbia e o Equador. A leste dos Andes foi coletada no Peru, Bolívia e Brasil. A localidade de Santa Cruz de La Sierra (Bolívia) representa o limite sul de distribuição da espécie, correspondente ao paralelo 18°S. Entretanto, existem registros da espécie na região amazônica e em muitas localidades do nordeste brasileiro (Spassky *et al.*, 1971; Winge, 1971; Dobzhansky e Powell, 1975; Ehrman e Powell, 1982; Santos e Valente, 1990).

Duas das espécies que compõem o subgrupo da *D. willistoni* são consideradas endêmicas: *D. insularis* e *D. pavlovskiana*. A espécie *D. insularis* é bastante restrita geograficamente, tendo sido encontrada apenas em quatro das Pequenas Antilhas, onde é simpátrica com *D. willistoni*: St. Kitts, Montserrat, Guadalupe e Santa Lucia. Finalmente, *D. pavlovskiana* é a que apresenta a menor distribuição geográfica de todas as crípticas. Até hoje foi encontrada apenas na Guiana e em apenas duas localidades compreendidas entre os paralelos 7° e 4°N (Spassky *et al.*, 1971; Winge, 1971; Dobzhansky e Powell, 1975; Ehrman e Powell, 1982; Santos e Valente, 1990).

A partir dos dados apresentados, é possível confirmar que o subgrupo da *D. willistoni* apresenta uma extensa área de distribuição geográfica na região Neotropical (Figura 4). Quatro dessas espécies (*D. willistoni*, *D. paulistorum*, *D. equinoxialis* e *D. tropicalis*) são amplamente distribuídas e apresentam muitas áreas de simpatria, enquanto duas delas (*D. insularis* e *D. pavlovskiana*) são endêmicas (Spassky *et al.*, 1971; Winge, 1971; Dobzhansky e Powell, 1975; Ehrman e Powell, 1982; Santos e Valente, 1990).

A extrema similaridade morfológica das espécies do subgrupo *willistoni*, o fato de serem abundantes na natureza e muitas vezes simpátricas, dificultam bastante o

trabalho de identificação dos espécimes baseados na morfologia externa das genitálias, embora os padrões de bandas dos cromossomos politênicos (Rohde *et al.*, 2006) e eletroforéticos da enzima Fosfatase Ácida – Acph (Garcia *et al.* 2006) tenham mostrado recentemente serem capazes de identificar a maior parte das espécies crípticas. Apesar disso, nosso grupo de pesquisa tem se dedicado ao estudo de vários aspectos da biologia, genética e evolução do grupo *willistoni* de *Drosophila*, o que motivou a realização do presente estudo.



Figura 5: Mapa das Américas onde estão representadas (linhas e símbolos coloridos) as áreas de distribuição geográfica de cada uma das seis espécies do subgrupo da *D. willistoni*, de acordo com os estudos de Spassky *et al.* (1971), Winge (1971), Dobzhansky e Powell (1975), Ehrman e Powell (1982), Santos e Valente (1990).

## 2. Objetivo Geral

O objetivo geral deste capítulo é investigar a história evolutiva do elemento *copia* em algumas espécies do grupo *willistoni* de *Drosophila*.

### 2.1. Objetivos Específicos

- Verificar a presença do elemento *copia* em espécies do grupo *willistoni* de *Drosophila*.
- Inferir as relações evolutivas do elemento *copia* nas diferentes espécies do grupo *willistoni* de *Drosophila*.

## 3. Materiais e Métodos

### 3.1. Espécies Investigadas

As espécies do grupo *willistoni* que foram investigadas e seus respectivos locais e data de coleta estão listadas na Tabela 2. A espécie *D. melanogaster* Meigen, 1830 (grupo *melanogaster*) foi utilizada como controle positivo nas análises, e a espécie *D. paramediostrata* Townsend e Wheeler, 1955 (grupo *tripunctata*) foi utilizada como controle negativo.

**Tabela 2: Espécies e linhagens estudadas e seu local de coleta.**

Espécies	Linhagem / Local de coleta
<i>D. willistoni</i>	Florianópolis/SC – coletada em 1997.
<i>D. willistoni</i>	Porto Alegre/RS – coletada em 1995.
<i>D. willistoni</i>	Tucson Stock Center/St Kitts – coletada em 2005.
<i>D. willistoni</i>	WIP4, Ipitanga/BA – coletada nos anos 60.
<i>D. willistoni</i>	17A2, Eldorado do Sul/RS – coletada em 1992.
<i>D. paulistorum</i>	semi-espécie Andino-Brasileira, POA.
<i>D. paulistorum</i>	semi-espécie Orinocana, POA.
<i>D. paulistorum</i>	Ribeirão Preto/SP – estoque antigo de laboratório.
<i>D. paulistorum</i>	Porto Alegre – estoque antigo de laboratório.
<i>D. equinoxialis</i>	México – estoque antigo de laboratório.
<i>D. insularis</i>	St. Kitts – estoque antigo de laboratório.
<i>D. nebulosa</i>	Santa Maria, RS – coleta 2006.
<i>D. melanogaster</i>	Harwich – estoque antigo de laboratório.
<i>D. paramediostriata</i>	Estoque antigo de laboratório.

### 3.2. Manutenção dos estoques de *Drosophila*

Os estoques de *Drosophila* foram mantidos a 19-20°C, em meio de cultura a base de farinha de milho, adaptado de Klein *et al.* (1999) no Laboratório de Biologia Molecular da Universidade Federal de Santa Maria. O meio é preparado a partir de uma mistura constituída de 2 kg de farinha de milho, 280 g de germe de trigo, 1 xícara de açúcar, 3 colheres de sopa de leite em pó, 1 colher de sopa de sal. Para cada 250 g desta mistura acrescentamos 750 mL de água e 1 grama de Nipagin dissolvido em 10 mL de etanol comercial. Cozinhase por aproximadamente 15 minutos. Depois se distribui o meio nos vidros de cultura e autoclava-se por 15 minutos.

As moscas foram transferidas, semanalmente, para frascos com meios de cultura novos, que dão condições de alimentação e ovoposição. O meio de cultura, de onde as moscas foram retiradas, foi estocado em câmara de

temperatura constante (19-20°C) para que os ovos eclodissem e as larvas se alimentassem, para quando adultas, serem novamente transferidas para frascos com meio de cultura novo.

### 3.3. Extração de DNA

As extrações de DNA foram feitas a partir de moscas adultas utilizando protocolo de fenol/clorofórmio segundo Sassi *et al.* (2005). As manipulações de ácidos nucleicos foram feitas segundo Sambrook *et al.* (1998).

### 3.4. PCR e Clonagem

Para a reação de PCR (Polymerase Chain Reaction) usamos o seguinte protocolo: 200 ng de DNA genômico adicionado a uma solução de 2.5 mM de MgCl<sub>2</sub>, 50 mM de KCl, 200 µM de cada desoxinucleotídeo, 100 pM de cada *primer* e 1,5 unidades de *Taq polimerase* (Gybcó-BRL) em 50 µl de volume total. Na amplificação das seqüências 5'LTR-URL do retrotransposon *copia* de aproximadamente 440 pb (Figura 6) foram utilizados os *primers* CopLTR e CopPCS, descritos por Jordan e McDonald (1998) (Tabela 3), da seguinte forma: um ciclo inicial a 94 °C por 5 min., seguido de 30 ciclos de 94 °C por 45 seg., 52 °C por 60 seg. e 72 °C por 60 seg., terminando com um alongamento de 72 °C por 5 minutos.

Os amplicons foram selecionados e clonados no vetor TA Clonig® Kit (Invitrogen). Das possíveis colônias recombinantes foi feita extração de plasmídeo por lise alcalina, e os clones foram confirmados através da PCR com *primers* M13F e M13R (Tabela 3), que se anelam ao vetor do kit. Logo após foram purificados com PEG 13%, NaCl 1,6M segundo protocolo descrito por Fujioka *et al.* (2000).

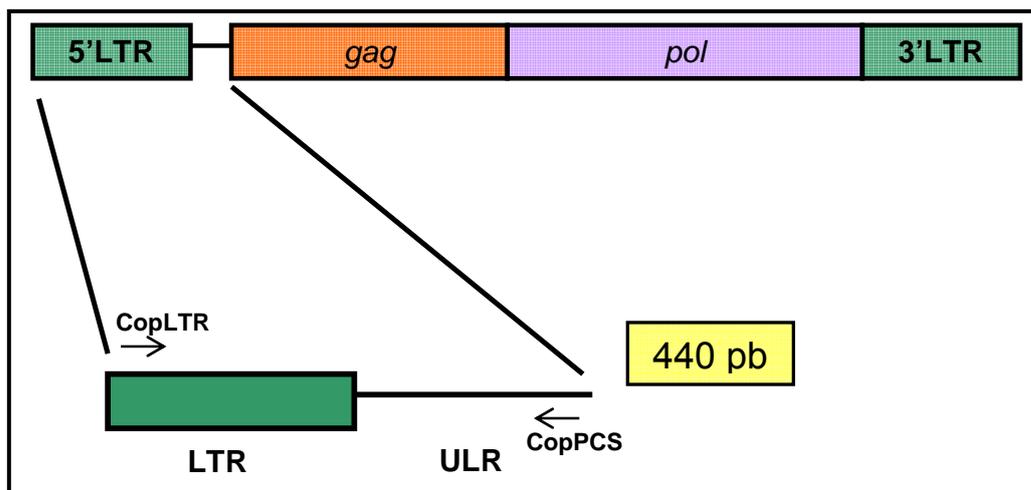


Figura 6: Ilustração da região 5'LRT-URL que foi amplificada no presente trabalho. O retângulo em amarelo refere-se ao tamanho da seqüência obtida. Os retângulos em verde correspondem à região das LTRs. Também estão ilustradas as ORFs (Open Reading Frames) de *copia*: gene *gag* (no retângulo laranja) e gene *pol* (no retângulo roxo). A região indicada é formada pela 5'LTR-URL e as flechas mostram a orientação dos primers CopLTR e CopPCS que amplificam essa região.

Tabela 3: Seqüência dos *primers* utilizados no trabalho.

<i>Primer</i>	Seqüência
CopLTR	5' – CTATTCAACCTACAAAATAACG – 3'
CopPCS	5' – TTACGTTTAGCCTTGTCAT – 3'
M13 <i>Forward</i>	5' – CTGGCCGTCGTTTTAC – 3'
M13 <i>Reverse</i>	5' – CAGGAAACAGCTATGAC – 3'

### 3.5. Seqüenciamento de DNA

O seqüenciamento de DNA foi realizado a partir de fragmentos de DNA dos possíveis clones, amplificados com os *primers* já descritos e purificados com PEG 13%, NaCl 1,6M em seqüenciador automático MegaBace 500 no Laboratório de Biologia Molecular da Universidade Federal de Santa Maria/RS. A reação de terminação em cadeia foi gerada com o kit DYEnamic ET® (GE Healthcare) e com os *primers* M13F e M13R. Ambas as fitas foram seqüenciadas.

Inicialmente, as seqüências de nucleotídeos obtidas foram confirmadas pelo uso do BLASTN (NCBI). As seqüências de cada clone foram submetidas separadamente e a análise de consenso realizada com base na confiabilidade apresentada pelos cromatogramas obtidos, pelo uso do programa Gap 4 do Staden Package (Staden, 1996).

### 3.6. Análise por *Southern Blot*

Para estimar o número de seqüências de *copia* em cada genoma, assim como, obter outra forma de evidência da existência destas seqüências nos genomas, o DNA das diferentes espécies estudadas foram submetidas à análise por *Southern Blot*.

Na análise por *Southern Blot* o DNA genômico (~6 µg) foi digerido com a enzima de restrição EcoRI. Os fragmentos de DNA foram separados por eletroforese em gel de agarose 0,8% e transferidos para uma membrana de náilon (Hybond-N+ - GE Healthcare). A sonda utilizada foi o plasmídeo PTZ18 contendo o elemento *copia* completo de *D. melanogaster*. A sonda foi marcada pelo método *random-prime*. A hibridização foi feita utilizando o Kit Alkphos® (GE Healthcare). Os filmes de auto-radiografia foram expostos à membrana primeiramente durante 2 horas, e, dependendo do sinal obtido, submetidos há tempos maiores.

## 4. Resultados e Discussão

Em um primeiro momento realizamos buscas por PCR de seqüências da '5LTR-URL do elemento *copia*, com os *primers* CopPCS e CopLTR, utilizando DNA genômico de algumas espécies do grupo *willistoni*. Foi utilizado como controle para amplificação positiva o plasmídeo contendo o elemento *copia* e o DNA genômico de *D. melanogaster*.

Observamos amplificação positiva para duas linhagens de *D. willistoni* (17A2 e Wip4), além da amplificação em *D. melanogaster* (Mount & Rubin, 1985) (Figura 7). Porém, nem todas as linhagens de *D. willistoni* amplificam para a região investigada, podendo ser indicativo de uma distribuição desigual dessa seqüência dentro dos genomas das diferentes linhagens da espécie.

Os amplicons de *D. willistoni* foram clonados e as colônias recombinantes foram submetidas à extração de plasmídeo. A partir das colônias extraídas foi realizada PCR com os *primers* M13F e M13R para confirmação os clones, e os mesmos submetidos à purificação com PEG 13%, NaCl 1,6M. Logo após os possíveis clones foram seqüenciados. Seqüenciamos ainda os produtos purificados da PCR, com os *primers* CopPCS e CopLTR,,

As seqüências nucleotídicas obtidas foram confirmadas pelo uso do BLASTN (NCBI). Foi encontrada homologia de 95 a 98% com a região 5'LTR-URL de *copia* de *D. melanogaster*, reforçando a hipótese de Jordan *et al.* (1999).

Empregamos uma estratégia para descartar a possibilidade de contaminação: a confirmação das espécies pelo padrão eletroforético da enzima Fosfatase Ácida – Acph1 (Garcia *et al.* 2006). A técnica foi realizada no Laboratório de *Drosophila* do Departamento de Genética da Universidade Federal do Rio Grande do Sul – UFRGS. Além disso, as moscas foram identificadas pela genitália do macho. O padrão de bandas pode ser observado na Figura 8. O resultado foi positivo para as espécies investigadas, ou seja, descartamos a hipótese de nossas amostras estarem contaminadas com outras espécies.

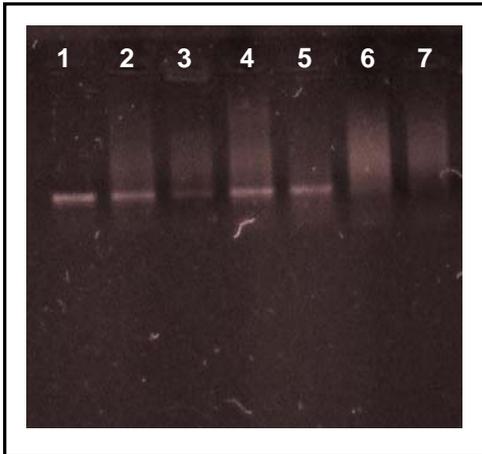


Figura 7: Foto de gel de agarose 0,8% representando a amplificação da região 5'LTR-URL de *copia* (1) plasmídeo PTZ18 contendo *copia*, (2) *D. melanogaster*, (3) *D. willistoni* Tucson Stock Center, (4) *D. willistoni* 17A2, (5) *D. willistoni* Wip4, (6) *D. paulistorum* Ribeirão Preto, (7) *D. nebulosa* Santa Maria.

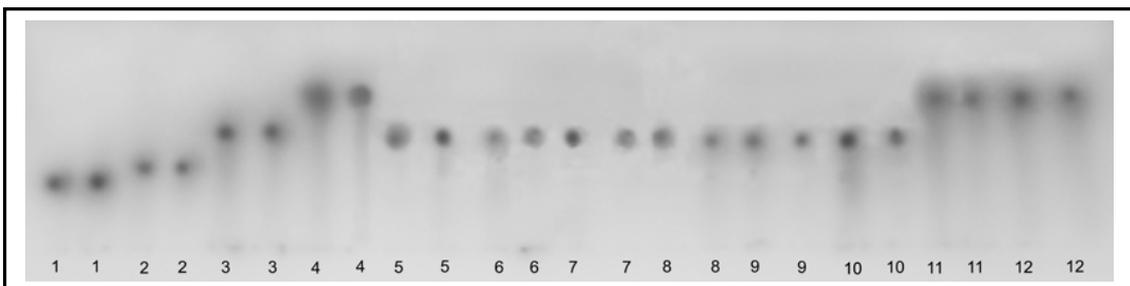


Figura 8: Padrão de bandas eletroforéticas observado em gel de poliacrilamida segundo técnica da enzima Fosfatase Ácida – Acph1 (Garcia *et al.* 2006). Em 1, 2, 3, 4, 6, 10, e 11 marcadores com espécies de estoques estabelecidos pelo Laboratório de *Drosophila*/UFRGS, respectivamente: *D. equinoxialis*, *D. tropicalis*, *D. willistoni* Wip4, *D. paulistorum* Ribeirão Preto, *D. willistoni* Wip4, *D. willistoni* Wip4 e *D. nebulosa*. Em 5, 7, 8, 9 e 12 as espécies e linhagens submetidas ao teste, respectivamente: *D. willistoni* POA, *D. willistoni* 17A2, *D. willistoni* Wip4, *D. willistoni* Florianópolis e *D. nebulosa* Santa Maria.

As mesmas espécies foram submetidas a um rastreamento por *Southern Blot* para aumentar o suporte dos dados obtidos por seqüenciamento. Além disso, outras linhagens de *D. willistoni* e outras espécies do grupo *willistoni* foram adicionadas à busca. Obtivemos sinal positivo para a maioria das espécies, porém a similaridade das seqüências não é a mesma encontrada na amplificação por PCR. Na Figura 9 podemos observar que o sinal da hibridização para *D. melanogaster* (3, controle positivo) é muito mais intenso do que para as espécies do grupo *willistoni* (4 a 12).

O que podemos afirmar é que as seqüências que obtivemos com alta similaridade ao elemento *copia* de *D. melanogaster* não são um componente comum presentes no genoma de *D. willistoni* e de outras espécies do grupo *willistoni*, já que nossas análises por *Southern Blot* mostram claramente que não obtivemos um sinal de hibridização forte.

Vale salientar que re-hibridizamos uma membrana inicialmente hibridizada com sondas para genes de cópia única (genes *white* e *yellow*) e fomos capazes de ter um sinal de hibridização positivo. Assim, se a maioria dos indivíduos da população possuísse uma cópia do elemento *copia* em alguma localização cromossômica definida, seríamos capazes de detectar este fato em nossos ensaios (Figura 10).

Algumas hipóteses foram levadas em consideração para explicarmos tais resultados, já que o fator contaminação foi excluído através da técnica da *Acp1*, que são:

- 1) Polimorfismo do elemento *copia* nos genomas hospedeiros. A exemplo do que foi descrito para o elemento *P* em linhagens de *D. sturtevantii* (subgrupo *sturtevantii*) (Almeida *et al.*, 2003) diferentes grupos de seqüências podem coexistir num mesmo genoma, e provavelmente refletirem a presença de múltiplas subfamílias do elemento. Isso pode ser explicado por eventos distintos de TH (Silva e Kidwell, 2000). Além disso, a maioria das seqüências descritas para o elemento *copia* se baseiam principalmente da região das LTRs, podendo ser indicativo da distribuição desigual desse elemento no gênero *Drosophila*. Esta hipótese explicaria nossas ampliações das LTRs em seqüências de 95 a 98% de similaridade com *copia* de *D. melanogaster*, e a discrepância com a análise de *Southern Blot*, pois se seqüências polimórficas do elemento estiverem presentes nos genomas seria possível sua amplificação por PCR, mas uma detecção de hibridização muito fraca por *Southern Blot*.
- 2) Pequeno número de cópias do elemento em poucos indivíduos da população. Outra possibilidade é que apenas alguns indivíduos da população, nas linhagens que apresentaram amplificação por PCR, possuam o elemento e em sítios polimórficos. Assim, seria possível termos um resultado positivo para PCR e negativo para *Southern Blot*, uma vez que, o número de cópias do elemento, ainda em sítios

polimórficos, não permitiria geral um sinal suficientemente forte para ser detectado por *Southern Blot*.

- 3) A seqüência do elemento  *copia*  num vetor. O fato da seqüência do elemento  *copia*  estar no genoma de um vetor é outra possibilidade. Segundo Loreto  *et al.*  (2008) os eventos de transferência horizontal podem ser considerados um dos passos primordiais para o ciclo de vida de um TE. Vários possíveis casos de TH foram descritos para retrotransposons e alguns vetores podem ser peças chave para o entendimento da evolução de alguns genomas. O ácaro  *Proctolaelaps regalis* , por exemplo, foi proposto como vetor específico da TH do elemento de  *P*  de  *D. willistoni*  para  *D. melanogaster*  (Houck  *et al.* , 1991). Os ácaros sugam e perfuram as larvas e ovos das moscas, e podem, através desse mecanismo, transferir o DNA dos TEs para o hospedeiro (Loreto  *et al.* , 2008). Bactérias simbióticas intracelulares, como  *Wolbachia*  e spiroplasmas, também podem consideradas como possíveis vetores, uma vez que eles vivem dentro células germinativas. Sendo assim, se o elemento  *copia*  estiver sendo transferido horizontalmente, utilizando um vetor como intermédio, poderíamos explicar sua amplificação por PCR e o baixo sinal obtido pela hibridização em  *Southern Blot* .

Devido aos resultados controversos encontrados no presente trabalho e com alguns os trabalhos encontrados na literatura (Biémont e Cizeron, 1999 e Jordan  *et al.* , 1999) realizaremos novas investigações e análises da distribuição do elemento  *copia*  no grupo  *willistoni* , para, num primeiro momento, confirmarmos as sugestões das hipóteses apresentadas neste trabalho, e posteriormente podermos inferir as relações filogenéticas desse elemento no gênero  *Drosophila*  e comparar com a análise filogenética das espécies hospedeiras.

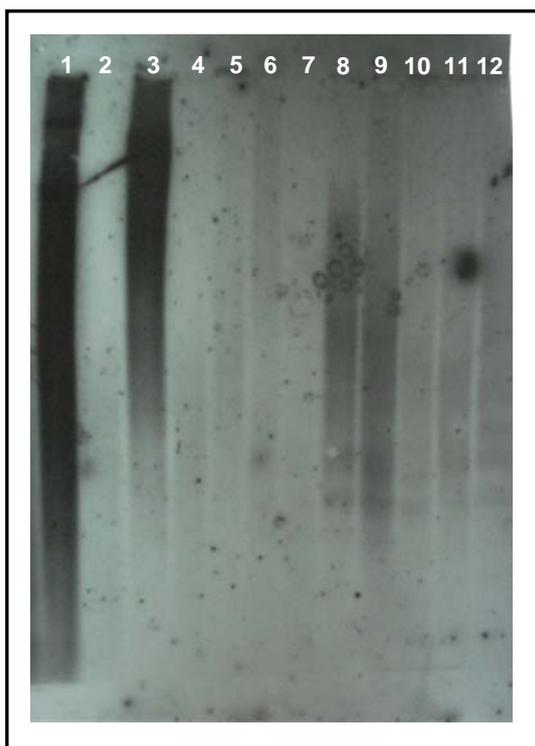


Figura 9: Fotografia de *Southern Blot*. 1) plasmídeo PTZ18 contendo *copia*, 2) *D. paramediostriata* (controle negativo), 3) *D. melanogaster*, 4) *D. willistoni* Wip4, 5) *D. willistoni* 17A2, 6) *D. willistoni* Tucson Stock Center, 7) *D. paulistorum* Morro Santana, 8) *D. willistoni* POA, 9) *D. paulistorum* Orinocana, 10) *D. paulistorum* Andino-brasileira, 11) *D. insularis* e 12) *D. equinoxialis*.

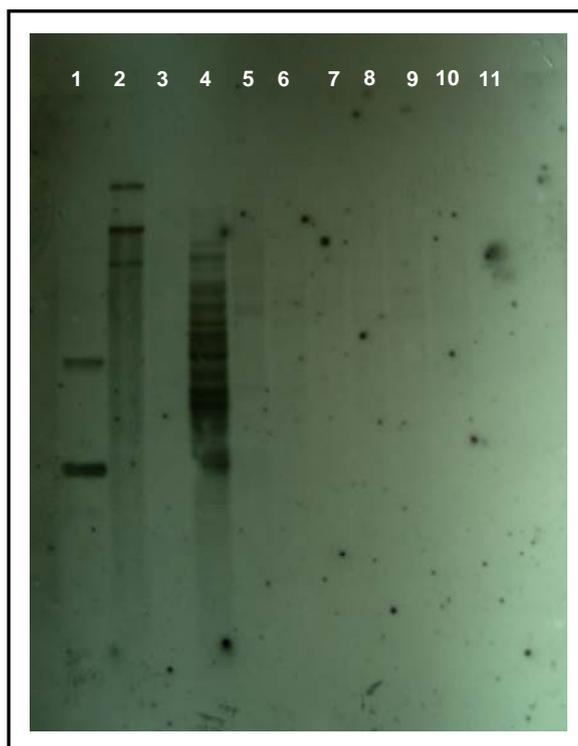


Figura 10: Fotografia de *Souther Blot*. Em 1) plasmídeo com inserto do gene *white*, 2) plasmídeo PTZ18 contendo *copia*, 3) *D. inmigrans*\* (utilizada como controle negativo), 4) *D. melanogaster*, 5) *D. nebulosa*, 6) *D. willistoni* Wip4, 7) *D. willistoni* EM1.00 *white*, 8) *D. willistoni* Q14.F11, 9) *D. willistoni* Ey10.00, 10) *D. willistoni* EM1.00 *yellow* e 11) *D. willistoni* Q14.

\*Argentina – coletada em 2006.

## **CAPÍTULO III**

**O Retrotransposon *copia* nos genomas disponíveis de *Drosophila*: uma análise de sua distribuição e história evolutiva**

# O Retrotransposon *copia* nos genomas disponíveis de *Drosophila*: uma análise de sua distribuição e história evolutiva

## 1. Justificativa

A divulgação, em banco de dados *on line*, do seqüenciamento e da análise comparativa de doze genomas de *Drosophila* (Clark *et al.*, 2007) estimulou a investigação de seqüências do elemento *copia* no presente trabalho. Adicionamos, assim, uma análise *in silico* da distribuição de *copia* no gênero *Drosophila* utilizando ferramentas de bioinformática, que poderá contribuir para o entendimento da dinâmica evolutiva desse elemento.

## 2. Objetivo Geral

O presente capítulo tem como objetivo geral ampliar as análises evolutivas do retroelemento *copia* através das seqüências encontradas *in silico* nos genomas disponíveis do gênero *Drosophila*.

### 2.1. Objetivos específicos

- Avaliar a distribuição do retrotransposon *copia* nos 12 genomas do gênero *Drosophila* disponíveis em bancos de dados *on line*. Atualmente existem disponíveis os genomas de: *D. ananassae*, *D. erecta*, *D. grimshawi*, *D. melanogaster*, *D. mojavensis*, *D. persimilis*, *D. pseudoobscura*, *D. sechellia*, *D. simulans*, *D. willistoni*, *D. virilis* e *D. yakuba*.

- Caracterizar os elementos  *copia*  encontrados nos genomas disponíveis do gênero  *Drosophila*  através da investigação do tamanho, número de cópias, conservação e presença de LTRs e ORFs.
- Avaliar o padrão evolutivo dos elementos  *copia*  encontrados nos diferentes genomas de  *Drosophila* .

### 3. Materiais e Métodos

Procuramos por seqüências completas homólogas ao retroelemento  *copia* . Foram realizadas buscas nos genomas de  *Drosophila*  atualmente disponíveis:  *D. grimshawi* ,  *D. virilis* ,  *D. mojavensis* ,  *D. willistoni* ,  *D. persimilis* ,  *D. pseudoobscura* ,  *D. ananassae* ,  *D. erecta* ,  *D. yakuba* ,  *D. melanogaster* ,  *D. simulans*  e  *D. sechellia* .

Para cada espécie, buscas foram realizadas usando a ferramenta BLAT (<http://genome.ucsc.edu/cgi-bin/hgBlat>; Kent, 2002), com exceção do genoma de  *D. willistoni* , que não estava disponível nessa ferramenta e, onde buscas foram feitas utilizando a ferramenta BLAST (<http://flybase.bio.indiana.edu/blast/>; Grumbling e Strelets, 2006).

A seqüência usada como sonda foi o elemento  *copia*  completo descrito em  *D. melanogaster*  (número de acesso: X02599) (Mount e Rubin, 1985). Para cada espécie, todas as seqüências encontradas com  *Score*  > 1.000 foram alinhadas usando o programa Clustal W (<http://www.ebi.uk/clustalw>; Higgins  *et al.* , 1994). Posteriormente, a partir das seqüências obtidas, buscamos pelas LTRs que flanqueiam o elemento, através de blast local.

Para manipulação e edição das seqüências foi utilizado o programa GeneDoc 2.6.001 (Nicholas e Nicholas, 1997). As seqüências encontradas foram nomeadas da seguinte forma: de três a quatro letras referentes às iniciais do nome da espécie em que foi encontrada, acrescidas de um número para cada seqüência diferente.

Foram realizadas buscas de possíveis ORFs em todas as seqüências obtidas, através do programa ORF Finder (<http://www.ncbi.nlm.nih.gov/projects/gorf>) para avaliarmos a integridade das mesmas nos diferentes genomas

Análises filogenéticas entre as seqüências encontradas nos diferentes genomas foram construídas, i) com base na seqüência de DNA do elemento completo, ii) a partir da seqüência de DNA das LTRs e iii) na seqüência de aminoácidos que foi deduzida a partir da seqüência de DNA, utilizando o programa GeneDoc 2.6.001 (Nicholas e Nicholas, 1997). Para as filogenias foram utilizados dois métodos de inferência filogenética: (1) Neighbor Joining (NJ) (Saitou e Nei, 1987) usando *bootstrap* de 1.000 réplicas no programa Mega 4.0 (Tamura *et al.*, 2007); (2) Máxima Parcimônia (MP) com parâmetros *default* e *bootstrap* de 1.000 réplicas no programa Mega 4.0 (Tamura *et al.*, 2007).

A distância-P entre as seqüências foi calculada pelo programa Mega 4.0 (Tamura *et al.*, 2007).

Para análise filogenética empregamos uma seqüência parcial do retroelemento *copia* que corresponde à posição 2.109 a 2.647 de *copia* de *D. melanogaster*, perfazendo um total de 539 nucleotídeos. Esta parte do elemento foi escolhida por permitir um alinhamento mais confiável. As seqüências de DNA com inserções e deleções, que geravam mudanças no quadro de leitura original dos aminoácidos, foram editadas para corresponderem à cadeia de aminoácidos do elemento *copia* de *D. melanogaster*.

#### 4. Resultados e Discussão

O resultado das buscas por seqüências completas de *copia* nos doze genomas de *Drosophila* pode ser observado na Tabela 4 e na Tabela 5. Os cálculos da distância-P entre as seqüências de DNA e a seqüência de aminoácidos podem ser observados na Tabela 6. As análises por MP e NJ resultaram em árvores com similar topologia. A árvore de NJ foi escolhida para apresentação, pois apresentou maior suporte estatístico. Na Figura 12 (A e B) apresentamos as árvores condensadas de NJ correspondentes, respectivamente, a seqüência de DNA e a seqüência de aminoácidos.

No genoma de *D. ananassae* foram encontradas dez seqüências relacionadas a *copia* (Tabela 4). O comprimento das seqüências variaram de 1.027 a 6.421 pb e

em duas seqüências (ana1 e ana2) foram encontradas LTRs 5' e 3' (Tabela 5). Vale salientar que, de todas as diferentes seqüências relacionadas a *copia* encontradas nos genomas analisados, apenas em ana1 observa-se uma ORF provavelmente completa, com 4.077 pb e codificando uma proteína de 1.358 aminoácidos, e, ainda, TSDs de quatro nucleotídeos. Segundo Ortiz e Loreto (2008), essas características estruturais podem ser indicativas de uma recente mobilização do elemento. Outro dado que corrobora esta hipótese é o fato de algumas seqüências (ana1, ana2, ana3, ana4) compartilharem uma similaridade integral (100% de similaridade de nucleotídeos) (Tabela 6). Em ana9 e ana10 observamos 13,3% de divergência com *copia* de *D. melanogaster*, porém, apresentam uma divergência de 28,9% com as outras seqüências relacionadas a *copia* no próprio genoma (Tabela 6). Em ana9 e ana10 observamos 13,3% de divergência com *copia*, porém, uma divergência de 28,9% entre as seqüências relacionadas a *copia* no próprio genoma (Tabela 6). Nas seqüências ana1, ana2, ana3 e ana4 a divergência das seqüências dos aminoácidos é menor com relação aos aminoácidos do elemento *copia* (10,3 a 11,8%) (Tabela 6).

No genoma de *D. melanogaster* foram encontradas trinta seqüências de *copia* (Tabela 4), em sua maioria, compartilhando 100% de similaridade entre elas conforme previamente descrito (Mount e Rubin, 1985).

No genoma de *D. mojavensis* foram encontradas quatro seqüências relacionadas a *copia* (Tabela 4), que apresentaram tamanhos variando de 2.962 a 3.752 pb (Tabela 5). Duas seqüências (moj2 e moj3) apresentaram possíveis LTRs e em nenhuma seqüência se observou ORFs completas (Tabela 5). As seqüências que foram utilizadas para as análises filogenéticas, moj1 e moj3, apresentaram 35,6% de divergência entre si e, 22,8 e 28,9%, respectivamente, com *copia* de *D. melanogaster* (Tabela 6).

No genoma de *D. persimilis* encontramos seis seqüências relacionadas a *copia* (Tabela 4) variando entre 3.249 a 7.241 pb, e as seqüências pers1, pers2 e pers5 apresentaram prováveis LTRs (Tabela 5). Não encontramos nenhuma ORF potencialmente completa. Estas seqüências apresentam de 0 a 2% de divergência se comparadas entre si e de 17,8 a 20% de divergência com *copia* de *D. melanogaster* (Tabela 6). A similaridade entre as seqüências desse genoma pode ser um indicativo de que a mobilização do elemento tenha ocorrido há pouco tempo.

No genoma de *D. sechellia* foram encontradas trinta e oito seqüências relacionadas ao elemento *copia* (Tabela 4). Essas seqüências apresentaram

tamanhos variando de 1.033 a 5.518 pb, e seis dessas seqüências apresentam LTRs 5' e 3' (sech03, sech17, sech23, sech24, sech28 e sech33), entretanto nenhuma ORF completa foi observada (Tabela 5). Devido ao grande número de seqüências obtidas no genoma de *D. sechellia* optamos por agrupar as seqüências com mais de 99% de similaridade (que corresponde a sech02, sech03, sech05, sech06, sech09, sech10, sech12, sech13, sech14, sech17, sech23, sech29, sech33, sech34, sech35 e sech38) para melhorar o desempenho do alinhamento e da árvore filogenética gerada. O referido consenso de seqüências foi nomeado como csech. O valor de divergência das seqüências desse genoma tem em média de 0 a 4,4%, com exceção de sech21 que apresentou valores maiores. Já com relação a *copia* de *D. melanogaster*, esses valores foram de 2,2 e 4,4%, com exceção de sech21 que apresentou divergência de 15,6% (Tabela 6). Há duas evidências da possível mobilização do elemento *copia* ocorrida provavelmente há pouco tempo no genoma de *D. sechellia*: i) o grande número de seqüências relacionadas ao elemento *copia* de *D. melanogaster* encontradas nesse genoma e, ii) os valores de divergência entre as seqüências de DNA e as seqüências de aminoácidos no próprio genoma e comparados ao o elemento *copia* de *D. melanogaster* (Tabela 6).

No genoma de *D. simulans* foram encontradas vinte e três seqüências relacionadas ao elemento *copia* (Tabela 4). No entanto, em todas as seqüências encontramos diversas inserções e as seqüências estavam relativamente degeneradas. Não conseguimos encontrar LTRs, TSDs ou ORFs. Devido a estes fatos, estas seqüências não foram incluídas na análise filogenética.

No genoma de *D. yakuba* encontramos cinco seqüências relacionadas ao elemento *copia* (Tabela 4). As seqüências analisadas (yak1, yak2 e yak3) variam de 3.033 a 6.945 pb, sendo que yak4 e yak5 apresentaram inserções e foram desconsideradas dessa análise. Não encontramos LTRs e ORFs completas (Tabela 5). As seqüências analisadas apresentaram uma divergência que variou de 0 a 2,2% entre elas, e uma divergência de 20 a 22,2% com relação *copia* de *D. melanogaster* (Tabela 6). Já a divergência nas seqüências de aminoácidos é menor, e fica entre 17,6 e 20,6% (Tabela 7).

Não foram encontradas seqüências relacionadas ao elemento *copia* nas espécies *D. erecta*, *D. grimshawi*, *D. pseudoobscura*, *D. willistoni* e *D. virilis* (Tabela 4).

**Tabela 4: Resultado das buscas de elementos relacionados a *copia* nos genomas de *Drosophila* disponíveis segundo presença e ausência, e número de cópias.**

Genoma	Presença de <i>copia</i>	Nº de cópias (Score >=1000)
<i>D. ananassae</i>	Sim	10
<i>D. erecta</i>	Não	-
<i>D. grimshawi</i>	Não	-
<i>D. melanogaster</i>	Sim	30
<i>D. mojavensis</i>	Sim	4
<i>D. persimilis</i>	Sim	6
<i>D. pseudoobscura</i>	Não	-
<i>D. sechellia</i>	Sim	38
<i>D. simulans</i>	Sim	23
<i>D. willistoni</i>	Não	-
<i>D. virilis</i>	Não	-
<i>D. yakuba</i>	Sim	5

**Tabela 5: Resultados obtidos a partir de buscas realizadas nos genomas de *Drosophila*. Descrição da localização das seqüências, tamanho da seqüência (em pares de base – pb), tamanho das LTRs 5' e 3' (em pb), tamanho da(s) ORF(s) (em pb) e tamanho da possível proteína (em aminoácidos – aa).**

Espécie	Seqüência	BLAT			Tamanho elemento(pb)	Tamanho LTR		ORFs >200bp Tamanho (pb)	Proteína Tamanho (aa)
		Localização	5'	3'		5'	3'		
<i>D. ananassae</i>	ana1	scaffold_12966	250699	257119	6421	236	335	4077	1358
	ana2	scaffold 12905	500361	505620	5260	218	237	1197 2964	398 967
	ana3	scaffold 5799	2819	6278	3460	-	-	2562	853
	ana4	scaffold 12330	1997	5102	3106	-	-	2445	814
	ana5	scaffold 12929	684825	686816	1992	-	-	1062	353
	ana6	scaffold 13082	2516800	2519761	2962	-	-	-	-
	ana7	scaffold 1251	608	1634	1027	-	-	905	304
	ana8	scaffold 8486	517	1546	1030	-	-	638	212
	ana9	scaffold 12966	5202	8377	3176	-	-	3150	1050
	ana10	scaffold 12966	89623	92690	3068	-	-	3045	1014
<i>D. mojavensis</i>	moj1	scaffold 4868	1713	5464	3752	-	-	438 1047 966 375	145 348 321 124
	moj2	scaffold 4386	16680	22004	5325	286	286	1725 609	574 202
	moj3	scaffold 4623	3632	8454	4823	278	278	867 2757	288 918
	moj4	scaffold 4833	1269	4230	2962	-	-	1191	396

<i>D. persimilis</i>	pers1	super_0	11766923	11770732	3810	250	149	756	251
								2514	837
	pers2	super_147	44151	48315	4165	250	250	786	261
								510	169
								561	186
	pers3	super_54	425576	443542	17967	-	-	1656	551
								2304	767
								504	167
	pers4	super_709	4212	8172	3961	-	-	633	210
								1266	421
pers5	super_72	144515	148445	3931	249	249	693	230	
							762	253	
pers6	super_840	876	4124	3249	-	-	1032	343	
							1215	404	
<i>D. yakuba</i>	yak1	3R_random	8480	15424	6945	-	-	732	243
								690	229
	yak2	U	4535719	4539452	3734	-	-	828	275
								-	-
								906	301
	yak3	U	3158777	3161809	3033	-	-	906	301
								477	158
								399	132
	yak4	U	3704122	4043448	339327	-	-	297	98
								990	329
1016								338	
yak5	U	17091729	17939310	847582	-	-	333	110	
							417	138	
							357	118	
<i>D. sechellia</i>	sech1	super_10813	1	1334	1334	-	-	327	108
								261	86
	sech2	super_116	85043	90209	5167	-	-	270	89
								180	59
								705	234
	sech3	super_1185	22	5153	5132	238	238	153	50
								195	64
								924	307
	sech4	super_1296	3451	7284	3834	-	-	261	86
								441	146
1830								609	
sech5	super_13516	1	1033	1033	-	-	918	305	
							111	36	
							3333	1110	
sech6	super_1501	1	4231	4231	-	-	801	266	
							129	42	
							192	63	
sech7	super_1581	1	4346	4346	-	-	1230	409	
							2061	686	
							107	36	
sech8	super_1639	1	2365	2365	-	-	974	324	
							279	92	
							174	57	
							3048	1015	
							1113	370	
							678	225	
							1725	574	
							171	56	
							114	37	
							111	36	
							1245	414	
							252	83	
							234	77	

							234	77
							192	63
sech9	super_1731	1	3599	3599	-	-	1245	414
							252	83
							264	87
							1323	440
							2802	933
sech10	super_1833	1	3845	3845	-	-	480	159
							105	34
							105	34
							2751	916
sech11	super_18	742555	747742	5088	-	-	447	148
							105	34
							153	50
							639	212
sech12	super_209	16350	21107	4758	-	-	411	136
							1674	557
							675	224
sech13	super_209	21208	23614	2407	-	-	2223	740
sech14	super_2195	1	3587	3587	-	-	492	163
							783	260
							1638	545
							828	275
sech15	super_2586	173	2769	2597	-	-	414	137
							102	33
							129	42
sech16	super_2607	1	1327	1327	-	-	1053	350
							105	34
							720	239
							2514	837
sech17	super_26	331521	336654	5134	240	239	777	258
							105	34
							153	50
							123	40
							105	34
sech18	Super_2788	1	2088	2088	-	-	195	64
							1230	409
							462	153
							480	159
sech19	super_42	185982	187400	1419	-	-	414	137
							108	35
							153	50
							195	64
sech20	Super_434	1116	2691	1576	-	-	150	49
							138	45
							489	162
							1254	417
sech21	Super_4601	1	2220	2220	-	-	150	49
							414	137
							306	101
sech22	Super_461	14176	15866	1691	-	-	150	49
							288	95
							108	35
							189	62

							138	45
							351	116
							687	228
sech23	Super_467	2472	7615	5144	240	239	174	57
							1986	661
							939	312
							366	121
							321	106
sech24	Super_479	4960	10477	5518	214	215	135	44
							156	51
							2655	884
							156	51
sech25	Super_5484	2	2024	2023	-	-	630	209
							531	176
							111	36
							189	62
sech26	super_57	256780	259871	3092	-	-	1230	409
							111	36
							153	50
							441	146
							414	137
sech27	super_57	259428	261249	1822	-	-	414	137
							111	36
							153	50
							1002	333
							2241	746
sech28	super_67	95046	100285	5240	238	238	777	258
							117	38
							105	34
							285	94
sech29	Super_6866	1	1782	1782	-	-	1254	417
							1053	350
sech30	Super_7249	1	1726	1726	-	-	480	159
sech31	Super_7475	1	1698	1698	-	-	1551	516
sech32	Super_7595	1	1683	1683	-	-	102	33
							1317	438
							327	108
sech33	super_80	12794	18035	5242	240	240	2013	670
							1995	664
							1413	470
sech34	Super_878	1083	6262	5180	-	-	186	61
							867	288
sech35	Super_8980	1	1530	1530	-	-	1440	479
sech36	super_91	1	2607	2607	-	-	129	42
							2123	707
							129	42
							735	244
sech37	Super_932	1	2289	2289	-	-	117	38
							129	42
							129	42
							153	50
							492	163
sech38	super_92	1	4329	4329	-	-	648	215
							888	295
							1812	603

As seqüências de LTRs 5' e 3' encontradas (ana1, ana2, moj2, moj3, pers1, pers2, pers5, sech03, sech17, sech23, sech24, sech28 e sech33) foram alinhadas e o alinhamento utilizado para gerar uma árvore filogenética. Os cálculos da distância-P para as LTRs podem ser observados na Tabela 7. As análises por MP e NJ resultaram em árvores com similar topologia. A árvore de NJ foi escolhida para apresentação, pois apresentou maior suporte estatístico. Na Figura 12-A apresentamos a árvore condensada de NJ correspondente a seqüência das LTRs 5' e 3'.

Os valores das divergências apresentadas, se compararmos as três metodologias – seqüência de DNA do elemento completo, seqüência de DNA das LTRs e na seqüência de aminoácidos – são bem distintos. Por exemplo, para ana1, a divergência de sua seqüência de DNA em relação a *copia* era de 22,2% e a seqüência de aminoácidos de 10,3% (Tabela 6). Quando consideramos a seqüência da região das LTRs esta porcentagem sobe para 54,1% (Tabela 7) com relação ao elemento *copia* de *D. melanogaster*. Assim, podemos observar um fenômeno interessante: o indicativo da ocorrência de um processo de diferenciação das seqüências das LTRs do genoma de *D. ananassae*, que estão bem conservadas, e com mais da metade das seqüências diferentes das do elemento *copia* descrito para *D. melanogaster*. Isso pode ser explicado pela região das LTRs apresentar seqüências que são responsáveis pela mobilização dos TEs nos genomas hospedeiros. A diferença entre as seqüências das LTRs nas diferentes espécies sugere que elas possam determinar a identidade do elemento e ser de grande importância para o entendimento das relações entre os TEs e seus hospedeiros.

#### 4.1. Relações Filogenéticas do Elemento *copia* X Relações Filogenéticas das Espécies Hospedeiras

A filogenia das seqüências relacionadas a *copia* nos genomas das espécies de *Drosophila* estudadas são coincidentes em muitos aspectos se comparados à filogenia das espécies hospedeiras (Figuras 11 e 12).

Quando comparamos as filogenias geradas a partir das seqüências de DNA e das seqüências de aminoácidos do elemento *copia* no genoma de *D. sechellia* (Figura 11–A e B) percebemos que elas são muito semelhantes e se agrupam, demonstrando coerência com a filogenia das espécies hospedeiras (Figura 11–C), onde *D. sechellia* e *D. melanogaster* fazem parte do mesmo subgrupo.

A primeira incongruência observada na comparação das seqüências do elemento *copia* com a filogenia das espécies hospedeiras diz respeito aos genomas de *D. yakuba* e *D. ananassae*. As seqüências relacionadas a *copia* no genoma de *D. yakuba* estão relativamente mais distantes da seqüência do elemento *copia* de *D. melanogaster* do que as seqüências de *D. ananassae*, que estão mais próximas (Figura 11–A e B). Porém, na filogenia das espécies (Figura 11–C), *D. yakuba* está mais próxima de *D. melanogaster* (fazem parte do subgrupo *melanogaster*) e *D. ananassae* está mais distante, compartilhando apenas do mesmo grupo de *D. melanogaster* e *D. yakuba*.

Uma segunda incongruência ocorre nas seqüências relacionadas ao elemento *copia* no genoma de *D. persimilis*. As seqüências relacionadas ao elemento *copia* nesta espécie, que pertence ao subgênero *Sophophora*, estão filogeneticamente mais distantes das seqüências relacionadas ao elemento *copia* das suas espécies aparentadas e, mais próximas das seqüências do genoma de *D. mojavensis*, que pertence ao subgênero *Drosophila* (Figura 11–A e B). Na filogenia baseada na seqüência de aminoácidos (Figura 11–B) fica evidente que *D. persimilis* e *D. mojavensis* compartilham, inclusive, de um ancestral comum das seqüências do elemento *copia*.

Outras relações mais complexas entre as seqüências do elemento *copia* nos diferentes genomas dificultaram a compreensão do cenário evolutivo deste elemento. Um exemplo disso são as seqüências ana9 e ana10, de *D. ananassae*, que estão em um ramo totalmente separado das seqüências das

espécies com quem divide o mesmo grupo *melanogaster* (*D. yakuba* e *D. melanogaster*). Outro exemplo relevante diz respeito às seqüências de DNA relacionadas ao elemento *copia* de *D. mojavensis*, que agrupam próximas as seqüências de *D. yakuba* (moj1) e próximas ao clado ana9-ana10 de *D. ananassae* (moj3). Quando consideramos a filogenia dos aminoácidos, essas seqüências (moj1 e moj3) ficam agrupadas e próximas das seqüências de *D. persimilis*, que, como já mencionamos, são de subgêneros diferentes segundo a filogenia das espécies (Figura 11–A e B). Levando em consideração a discrepância das seqüências encontradas em *D. mojavensis*, podemos sugerir que estas seqüências possam representar famílias diferentes do retrotransposon *copia* encontradas no mesmo genoma, podendo também representar um polimorfismo ancestral.

Na filogenia baseada nas seqüências das LTRs do retroelemento *copia* encontramos novos indícios de que o elemento possa ter compartilhado de um ancestral comum nos subgêneros *Drosophila* e *Sophophora* (Almeida e Carareto, 2006) e que pode ser observado na árvore filogenética da Figura 12–A. As seqüências das LTRs relacionadas ao elemento *copia* nos diferentes genomas possuem graus de parentesco semelhantes a das espécies hospedeiras (Figura 12 – A e B). Um exemplo disso é entre as seqüências de *D. sechellia* e *D. melanogaster*, que estão agrupadas no mesmo clado. Em *D. mojavensis* as seqüências das LTRs separam-se das seqüências das LTRs das espécies do subgênero *Sophophora*, concordando também com a filogenia das espécies. Incongruências são encontradas nessas seqüências, no que diz respeito à ancestralidade das espécies hospedeiras. As seqüências de *D. ananassae*, por exemplo, dividem ancestralidade com as seqüências de *D. persimilis*, que faz parte de outro grupo de espécies (*obscura*) na filogenia das espécies hospedeiras (Figura 12–B).

Algumas considerações, porém, devem feitas no que diz respeito aos genomas em que não encontramos seqüências relacionadas ao elemento *copia*. O primeiro caso no subgrupo *melanogaster*, onde encontramos seqüências relacionadas ao elemento em *D. sechellia*, *D. yakuba* e *D. simulans* (além de *D. melanogaster*), mas não encontramos em *D. erecta*. Essa espécie (*D. erecta*) pode ter sofrido perda estocástica do elemento *copia* se levarmos em consideração a transmissão vertical nas espécies relacionadas ao subgrupo *melanogaster*. Outro caso parecido acontece no grupo *obscura*, onde encontramos seqüências

relacionadas ao elemento *copia* no genoma de *D. persimilis*, mas não encontramos no genoma de sua espécie-irmã *D. pseudoobscura*.

No caso de *D. willistoni*, não obtivemos seqüências relacionadas ao elemento *copia* nas buscas realizadas no genoma disponível, contradizendo os dados de Biémont e Cizeron (1999) para dados obtidos com experimentos de hibridização, e, ainda, os dados obtidos de Jordan *et al.* (1999), para amplificação da região da 5'LTR-URL. Não descartamos a possibilidade de a linhagem utilizada no seqüenciamento ser muito antiga, o que explicaria o fato de não encontramos seqüências relacionadas ao retroelemento, que pode ter "invadido" mais recentemente esse genoma.

Portanto, das nove espécies investigadas do subgênero *Sophophora*, somente em três delas não encontramos seqüências relacionadas ao elemento *copia*, e nas três espécies investigadas do subgênero *Drosophila*, seqüências relacionadas ao elemento foram encontradas apenas no genoma de *D. mojavensis*, demonstrando uma distribuição desigual desse elemento nos genomas, e sendo necessário, assim, uma análise mais detalhada das seqüências encontradas para entendermos sua história evolutiva em Drosophilidae.

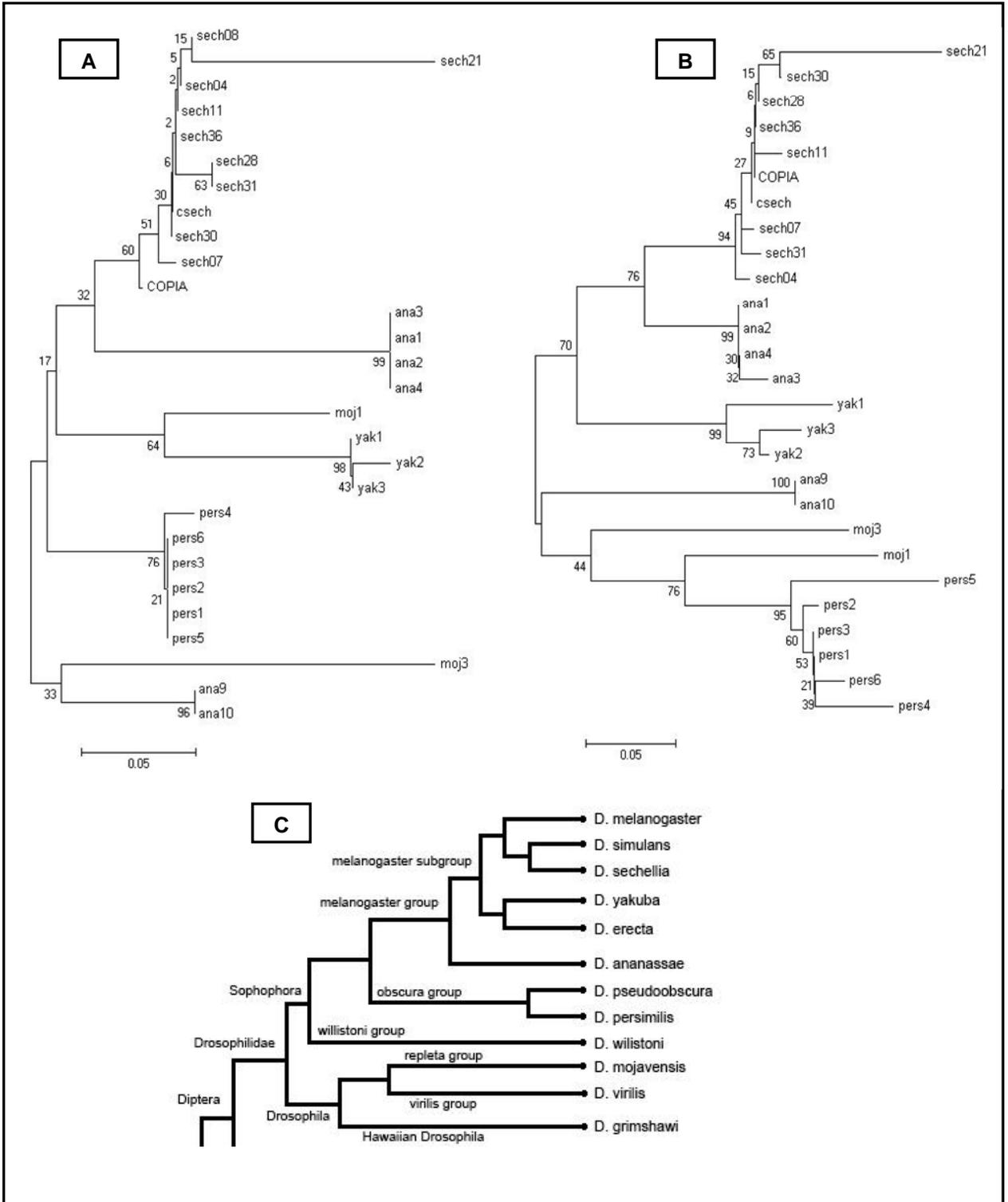


Figura 11: Árvores Filogenéticas geradas a partir de NJ, em A) Seqüências de DNA do elemento *copia* encontradas nos genomas, e B) Seqüências de aminoácidos encontradas nos genomas (*bootstrap* de 1000 réplicas). Em C Filogenia das espécies hospedeiras (modificada de <http://flybase.bio.indiana.edu/blast>).

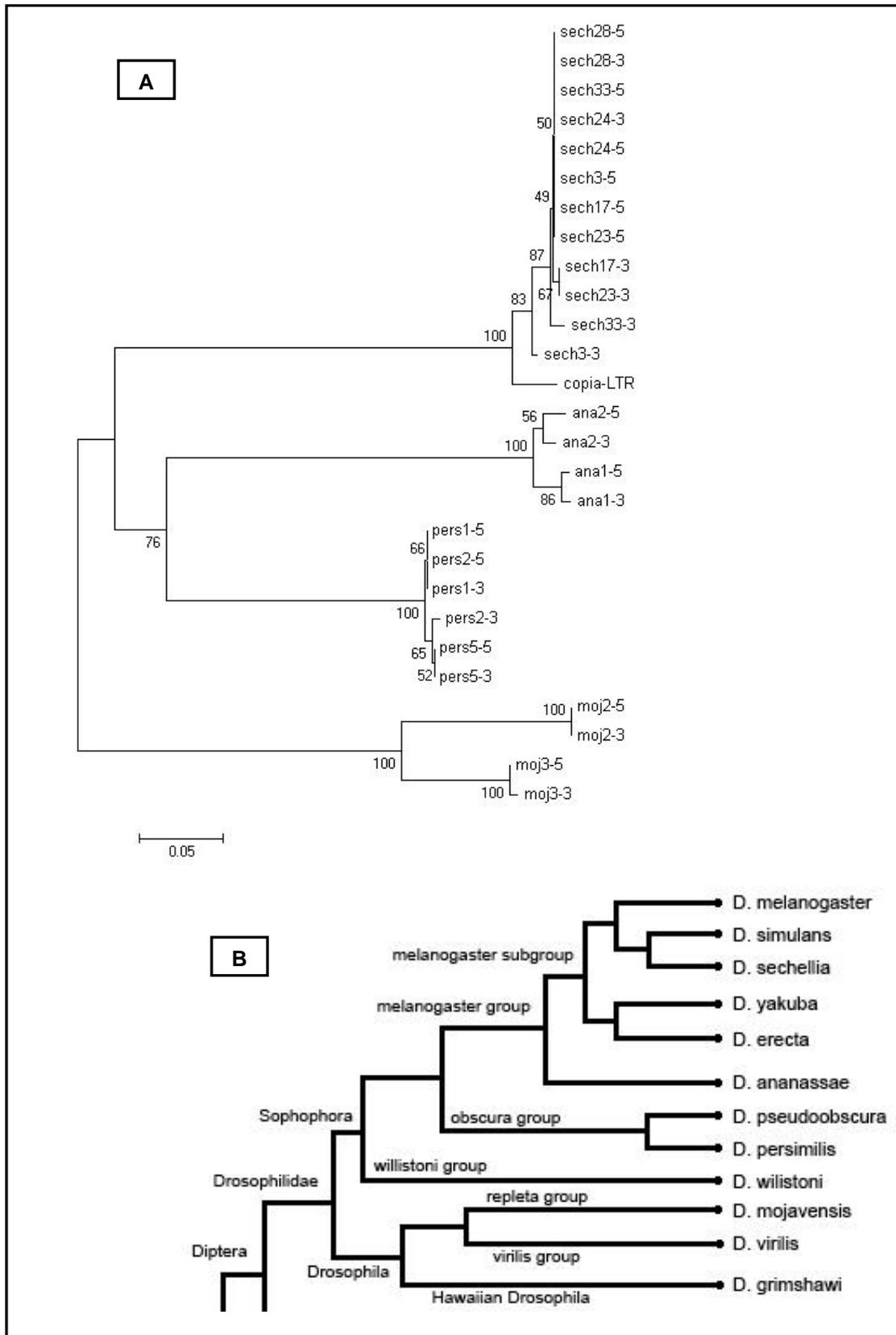


Figura 12: Árvore Filogenética geradas a partir de NJ, em A) Seqüências de DNA das LTRs do elemento *copia* encontradas nos genomas (*bootstrap* de 1000 réplicas). Em B) Filogenia das espécies hospedeiras (modificada de <http://flybase.bio.indiana.edu/blast>).

**Tabela 6: Porcentagem de divergência de nucleotídeos, calculada por distância-P, (abaixo e esquerda na diagonal) e aminoácidos (acima e direita na diagonal) encontrada entre as seqüências dos genomas e as seqüências do elemento *copia*.**

	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
1-copia		10,3	10,3	11,8	10,3	23,5	23,5	27,9	25,0	23,5	23,5	23,5	27,9	29,4	25,0	0,0	1,5	1,5	1,5	8,8	0,0	1,5	1,5	0,0	20,6	17,6	19,1
2-ana1	22,2		0,0	1,5	0,0	22,1	22,1	25,0	23,5	23,5	23,5	23,5	26,5	29,4	25,0	10,3	10,3	11,8	11,8	19,1	10,3	11,8	11,8	10,3	22,1	19,1	20,6
3-ana2	22,2	0,0		1,5	0,0	22,1	22,1	25,0	23,5	23,5	23,5	23,5	26,5	29,4	25,0	10,3	10,3	11,8	11,8	19,1	10,3	11,8	11,8	10,3	22,1	19,1	20,6
4-ana3	22,2	0,0	0,0		1,5	23,5	23,5	26,5	25,0	25,0	25,0	25,0	27,9	30,9	26,5	11,8	11,8	13,2	11,8	20,6	11,8	13,2	13,2	11,8	23,5	20,6	22,1
5-ana4	22,2	0,0	0,0	0,0		22,1	22,1	25,0	23,5	23,5	23,5	23,5	26,5	29,4	25,0	10,3	10,3	11,8	11,8	19,1	10,3	11,8	11,8	10,3	22,1	19,1	20,6
6-ana9	13,3	28,9	28,9	28,9	28,9		0,0	23,5	26,5	26,5	27,9	26,5	27,9	30,9	27,9	23,5	23,5	25,0	25,0	30,9	23,5	25,0	23,5	23,5	26,5	23,5	25,0
7-ana10	13,3	28,9	28,9	28,9	28,9	0,0		23,5	26,5	26,5	27,9	26,5	27,9	30,9	27,9	23,5	23,5	25,0	25,0	30,9	23,5	25,0	23,5	23,5	26,5	23,5	25,0
8-moj1	22,2	26,7	26,7	26,7	26,7	22,2	0,0		27,9	16,2	17,6	16,2	19,1	23,5	17,6	27,9	27,9	29,4	29,4	36,8	27,9	29,4	29,4	27,9	27,9	26,5	27,9
9-moj3	28,9	33,3	33,3	33,3	33,3	26,7	22,2	35,6		23,5	25,0	23,5	27,9	25,0	25,0	25,0	23,5	26,5	32,4	25,0	26,5	23,5	25,0	25,0	32,4	32,4	32,4
10-pers1	11,1	26,7	26,7	26,7	26,7	17,8	26,7	22,2	26,7		1,5	0,0	4,4	8,8	1,5	23,5	23,5	22,1	25,0	32,4	23,5	25,0	25,0	23,5	27,9	25,0	26,5
11-pers2	11,1	26,7	26,7	26,7	26,7	17,8	17,8	22,2	26,7	0,0		1,5	5,9	10,3	2,9	23,5	23,5	22,1	25,0	32,4	23,5	25,0	25,0	23,5	27,9	25,0	26,5
12-pers3	11,1	26,7	26,7	26,7	26,7	17,8	17,8	22,2	26,7	0,0	0,0		4,4	8,8	1,5	23,5	23,5	22,1	25,0	32,4	23,5	25,0	25,0	23,5	27,9	25,0	26,5
13-pers4	13,3	28,9	28,9	28,9	28,9	20,0	17,8	20,0	28,9	2,2	2,2	2,2		13,2	5,9	27,9	27,9	26,5	29,4	36,8	27,9	29,4	29,4	27,9	29,4	26,5	27,9
14-pers5	11,1	26,7	26,7	26,7	26,7	17,8	20,0	22,2	26,7	0,0	0,0	0,0	22,2		10,3	29,4	29,4	27,9	30,9	38,2	29,4	30,9	29,4	29,4	30,9	27,9	30,9
15-pers6	11,1	26,7	26,7	26,7	26,7	17,8	17,8	22,2	26,7	0,0	0,0	0,0	22,2	0,0		25,0	25,0	23,5	26,5	33,8	25,0	26,5	26,5	25,0	27,9	26,5	27,9
16-csech	2,2	20,0	20,0	20,0	20,0	15,6	17,8	24,4	31,1	13,3	13,3	13,3	15,6	13,3	13,3		1,5	1,5	1,5	8,8	0,0	1,5	1,5	0,0	20,6	17,6	19,1
17-sech04	2,2	20,0	20,0	20,0	20,0	15,6	15,6	24,4	31,1	13,3	13,3	13,3	15,6	13,3	13,3	0,0		2,9	2,9	10,3	1,5	2,9	2,9	1,5	22,1	19,1	20,6
18-sech07	4,4	22,2	22,2	22,2	22,2	17,8	15,6	26,7	28,9	11,1	11,1	11,1	13,3	11,1	11,1	2,2	2,2		2,9	10,3	1,5	2,9	2,9	1,5	22,1	19,1	20,6
19-sech08	2,2	20,0	20,0	20,0	20,0	15,6	17,8	24,4	31,1	13,3	13,3	13,3	15,6	13,3	13,3	0,0	0,0	2,2		10,3	1,5	2,9	2,9	1,5	22,1	19,1	20,6
20-sech11	2,2	20,0	20,0	20,0	20,0	15,6	15,6	24,4	31,1	13,3	13,3	13,3	15,6	13,3	13,3	0,0	0,0	2,2	0,0		8,8	8,8	10,3	8,8	29,4	26,5	26,5
21-sech21	15,6	33,3	33,3	33,3	33,3	28,9	15,6	37,8	44,4	26,7	26,7	26,7	28,9	26,7	26,7	13,3	13,3	15,6	13,3	13,3		1,5	1,5	0,0	20,6	17,6	19,1
22-sech28	4,4	22,2	22,2	22,2	22,2	17,8	28,9	26,7	31,1	15,6	15,6	15,6	17,8	15,6	15,6	2,2	2,2	4,4	2,2	2,2	15,6		2,9	1,5	22,1	19,1	20,6
23-sech30	2,2	20,0	20,0	20,0	20,0	15,6	17,8	24,4	31,1	13,3	13,3	13,3	15,6	13,3	13,3	0,0	0,0	2,2	0,0	0,0	13,3	2,2		1,5	20,6	17,6	17,6
24-sech31	4,4	22,2	22,2	22,2	22,2	17,8	15,6	26,7	31,1	15,6	15,6	15,6	17,8	15,6	15,6	2,2	2,2	4,4	2,2	2,2	15,6	0,0	2,2		20,6	17,6	19,1
25-sech36	2,2	20,0	20,0	20,0	20,0	15,6	17,8	24,4	31,1	13,3	13,3	13,3	15,6	13,3	13,3	0,0	0,0	2,2	0,0	0,0	13,3	2,2	0,0	2,2		7,4	10,3
26-yak1	20,0	31,1	31,1	31,1	31,1	26,7	15,6	20,0	37,8	24,4	24,4	24,4	22,2	24,4	24,4	22,2	22,2	24,4	22,2	22,2	35,6	24,4	22,2	24,4	22,2		2,9
27-yak2	22,2	33,3	33,3	33,3	33,3	26,7	26,7	22,2	35,6	26,7	26,7	26,7	24,4	26,7	26,7	24,4	24,4	26,7	24,4	24,4	37,8	26,7	24,4	26,7	24,4	2,2	
28-yak3	20,0	31,1	31,1	31,1	31,1	26,7	26,7	20,0	37,8	24,4	24,4	24,4	22,2	24,4	24,4	22,2	22,2	24,4	22,2	22,2	35,6	24,4	22,2	24,4	22,2	0,0	2,2

**Tabela 7: Porcentagem de divergência de nucleotídeos, calculada por distância-P, encontrada entre as seqüências das LTRs de *copia* dos genomas e as seqüências do elemento *copia*.**

	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		
1-copiaLTR																												
2-ana2-5'	53,6																											
3-ana2-3'	53,0	2,2																										
4-ana1-5'	54,1	4,4	2,2																									
5-ana1-3'	54,1	5,5	3,3	1,1																								
6-moj2-5'	56,9	57,5	57,5	58,0	58,0																							
7-moj2-3'	56,9	57,5	57,5	58,0	58,0	0,0																						
8-moj3-5'	52,5	56,4	56,4	57,5	57,5	17,1	17,1																					
9-moj3-3'	53,0	56,9	56,9	58,0	58,0	16,6	16,6	0,6																				
10-sech3-5'	5,0	52,5	51,9	53,0	53,0	58,0	58,0	54,1	54,7																			
11-sech3-3'	6,6	51,4	50,8	51,9	51,9	57,5	57,5	53,6	54,1	1,7																		
12-sech17-5'	5,0	52,5	51,9	53,0	53,0	58,0	58,0	54,1	54,7	0,0	1,7																	
13-sech17-3'	5,5	53,0	52,5	53,6	53,6	57,5	57,5	53,6	54,1	0,6	2,2	0,6																
14-sech28-5'	5,0	52,5	51,9	53,0	53,0	58,0	58,0	54,1	54,7	0,0	1,7	0,0	0,6															
15-sech28-3'	5,0	52,5	51,9	53,0	53,0	58,0	58,0	54,1	54,7	0,0	1,7	0,0	0,6	0,0														
16-sech23-5'	5,0	52,5	51,9	53,0	53,0	58,0	58,0	54,1	54,7	0,0	1,7	0,0	0,6	0,0	0,0													
17-sech23-3'	5,5	53,0	52,5	53,6	53,6	57,5	57,5	53,6	54,1	0,6	2,2	0,6	0,0	0,6	0,6	0,6												
18-sech24-5'	5,0	52,5	51,9	53,0	53,0	58,0	58,0	54,1	54,7	0,0	1,7	0,0	0,6	0,0	0,0	0,0	0,0	0,6										
19-sech24-3'	5,0	52,5	51,9	53,0	53,0	58,0	58,0	54,1	54,7	0,0	1,7	0,0	0,6	0,0	0,0	0,0	0,6	0,0										
20-sech33-3'	6,1	53,0	52,5	53,6	53,6	58,6	58,6	54,7	55,2	1,1	2,8	1,1	1,7	1,1	1,1	1,1	1,7	1,1	1,1									
21-sech33-5'	5,0	52,5	51,9	53,0	53,0	58,0	58,0	54,1	54,7	0,0	1,7	0,0	0,6	0,0	0,0	0,0	0,6	0,0	0,0	1,1								
22-pers1-5'	45,3	39,8	39,2	39,2	39,2	50,3	50,3	45,3	45,9	44,8	43,1	44,8	45,3	44,8	44,8	44,8	45,3	44,8	44,8	45,3	44,8	44,8	45,3	44,8				
23-pers1-3'	45,3	39,8	39,2	39,2	39,2	50,3	50,3	45,3	45,9	44,8	43,1	44,8	45,3	44,8	44,8	44,8	45,3	44,8	44,8	45,3	44,8	44,8	45,3	44,8	0,0			
24-pers2-5'	45,3	39,8	39,2	39,2	39,2	50,3	50,3	45,3	45,9	44,8	43,1	44,8	45,3	44,8	44,8	44,8	45,3	44,8	44,8	45,3	44,8	44,8	45,3	44,8	0,0	0,0		
25-pers2-3'	46,4	40,3	39,8	39,8	39,8	50,3	50,3	45,3	45,9	45,9	44,2	45,9	46,4	45,9	45,9	45,9	46,4	45,9	45,9	46,4	45,9	45,9	46,4	45,9	1,1	1,1	1,1	
26-pers5-5'	45,9	39,8	39,2	39,2	39,2	50,8	50,8	45,9	46,4	45,3	43,6	45,3	45,9	45,3	45,3	45,3	45,9	45,3	45,3	45,9	45,3	45,3	45,9	45,3	0,6	0,6	0,6	0,6
27-pers5-3'	45,9	39,8	39,2	39,2	39,2	50,8	50,8	45,9	46,4	45,3	43,6	45,3	45,9	45,3	45,3	45,3	45,9	45,3	45,3	45,9	45,3	45,3	45,9	45,3	0,6	0,6	0,6	0,0

**CAPITULO IV**  
**Considerações Finais e Perspectivas**

## Considerações Finais e Perspectivas

### 1. Referentes ao Capítulo II

Devido aos resultados controversos encontrados no presente trabalho em relação a alguns trabalhos encontrados na bibliografia (Biémont e Cizeron, 1999 e Jordan *et al.*, 1999), novas investigações serão necessárias. Pretendemos testar as hipóteses que levantamos com relação a discrepância entre nossos resultados de PCR e *Southern Blot* da seguinte forma:

- Hipótese 1: “Se existirem divergências entre a região da 5’LTR-URL e outras regiões do elemento”. Para testarmos esta hipótese pretendemos desenhar *primers* degenerados para regiões da TR do elemento  *copia* usando as seqüências obtidas no capítulo III.
- Hipótese 2: “Pode existir um polimorfismo de número de cópias e sítios de inserção nas populações que amplificam por PCR, sendo que poucos indivíduos apresentam o elemento”. Para testar esta hipótese, faremos 50 isolinhagens e testaremos todas estas isolinhagens por PCR.
- Hipótese 3: “O elemento  *copia* poderia estar no genoma de um vetor (*Wolbachia* ou *spiroplasma*)”. Esta hipótese é mais difícil de testar, entretanto, alguns antibióticos e tratamento com frio são descritos como capazes de eliminar estes endoparasitas.

Acreditamos que estes ensaios se fazem necessários para podermos resolver as inconsistências entre nossos dados de PCR e *Southern* com os dados da literatura.

## 2. Referentes ao Capítulo III

Pretendemos ampliar os dados das buscas realizadas nos doze genomas disponíveis, a princípio com algumas estratégias:

- Realizando novas buscas, e utilizando como sonda as seqüências obtidas nas primeiras buscas. Como descrito, as buscas foram feitas somente com a seqüência do elemento *copia* de *D. melanogaster*. Imaginamos que, ao ampliar as buscas, empregando as seqüências já obtidas, novas seqüências podem ser encontradas, o que poderia refinar nossa análise filogenética.
- Realizarmos testes de dS e dN (divergências encontradas em sítios sinônimos e não-sinônimos, respectivamente) pois se o elemento *copia* foi transmitido de uma espécie para outra por TH, é esperado que o dS de *copia* entre as espécies seja significativamente menor que os dS encontrado para genes nucleares, refletindo o menor tempo de divergência das seqüências em relação a seqüência ancestral.
- Ampliar as análises das LTRs.

## REFERÊNCIAS BIBLIOGRÁFICAS

Almeida LD, Langeani F e Carareto CMA (2003). Geographic polymorphism of P element in populations of *Drosophila sturtevanti*. *Genet Mol Biol* 26.2:175-179.

Almeida LD e Carareto CMA (2006) Sequence heterogeneity and phylogenetic relationship between the *copia* retrotransposon in *Drosophila* species of the *repleta* and *melanogaster* groups. *Genet. Sel. Evol.* 38: 535–550

Bächli G (2008) Taxodros: The database on taxonomy of Drosophilidae, v1, 03. Database 2008/01. Last accessed on 08/01/2008.

Becker H e Lönnig W (2001) Transposons: eukaryotic. In: *Nature encyclopedia of life sciences*, Nature Publishing Group, Londres, Reino Unido, 18:529-539.

Biémont C e Cizeron G (1999) Distribution of transposable elements in *Drosophila* species. *Genetica* 105:43-62.

Biémont C e Vieira C (2006) Genetics: junk DNA as an evolutionary force. *Nature* 443:521-524.

Boeke JD e Stoye JP (1997) Retrotransposons, endogenous retroviruses, and the evolution of retroelements. In: Coffin JM, Hughes SH e Varmus HE (eds) *Retroviruses*. Cold Spring Harbor Laboratory Press, New York, pp 343–435.

Bowen NJ e Jordan IK (2002) Transposable elements and the evolution of eukaryotic complexity. *Curr Issues Mol Biol* 4:65-76.

Britten RJ (1996) DNA sequence insertion and evolutionary variation in gene regulation. *Proc Natl Acad Sci USA* 93:9374-9377.

Brookfield JF (2005) The ecology of the genome - mobile DNA elements and their hosts. *Nature Rev Genet* 6:128-134.

Bundock P e Hooykaas P (2005) An *Arabidopsis* *hAT*-like transposase is essential for plant development. *Nature* 436:282-284.

Burla H, Da Cunha AB, Cordeiro AR, Dobzhansky T, Malogolowkin C, Pavan C (1949) The *willistoni* group of sibling species of *Drosophila*. *Evolution* 3:300–314.

Capy P, Anxolabéhère D e Langin T (1994) The strange phylogenies of transposable elements: are horizontal transfers the only explanation? *Trends Genet* 10:7-12.

Capy P, Bazin C, Higuët D e Langin T (1998) *Dynamics and evolution of transposable elements*. Landes Bioscience, Austin, Texas, 197 pp.

Capy P, Gasperi G, Biemont C e Bazin C (2000) Stress and transposable elements: co-evolution or useful parasites? *Heredity* 85:101-106.

Castro JP e Carareto CM (2004) *Drosophila melanogaster* P transposable elements: mechanisms of transposition and regulation *Genetica* 121:107-118.

Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman T, Kellis M, Gelbart W, *et al* (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450:203–218

Daboussi M e Capy P (2003) Transposable elements in filamentous fungi. *Annu. Rev. Microbiol.* 57:275–299.

Da Lage J-L, Kergoat GJ, Maczkowiak F, Silvain J-F, Cariou M-L e Lachaise D (2007) A phylogeny of Drosophilidae using the *Amyrel* gene: questioning the *Drosophila melanogaster* species group boundaries. *J. Zool. Syst. Evol. Res.* 45:47–63.

Dobzhansky T (1946) Complete reproductive isolation between two morphologically similar species of *Drosophila*. *Ecology* 27:205–211.

Dobzhansky T (1957) Chromosomal variability in island and continental populations of *Drosophila willistoni* from central America and the West Indies. *Evolution* 11:280–293.

Dobzhansky T e Spassky B (1959) *Drosophila paulistorum*, a cluster of species in statu nascendi. *Proc Natl Acad Sci* 45:419–428.

Dobzhansky T, Powell JP (1975) The *willistoni* group of sibling species of *Drosophila*. In: King RC (ed.), *Handbook of Genetics*. Plenum Press, New York, pp 589–622.

Ehrman L e Powell JR (1982) The *Drosophila willistoni* species group. In: Ashburner M, Carson HL and Thompson JN (eds) *The Genetics and Biology of Drosophila*. Academic Press, New York, pp 193-225.

Finnegan DJ (1989) Eukaryotic transposable elements and genome evolution. *Trends Genet* 5:103-107.

Flavell AJ, Ruby SW, Toole JJ, Robertso BE e Rubin GM (1980) Translation and developmental regulation of RNA encoded by the eukaryotic transposable element *copia*. *Proc Natl Acad Sci USA* 77.12:7107-7111.

Flavell AJ, Levis R, Simon MA, Rubin GM (1981) The 5' termini of RNAs encoded by the transposable element *copia*, *Nucleic Acids Res.* 9:6279–6291.

Fujioka M, Jaynes JB, Bejsovec A e Weir M (2000) Production of transgenic *Drosophila*. *Methods Mol. Biol.*: 136:353-63.

Garcia ACL, Rohde C, Audino GF, Valente VLS, Valiati VH. (2006) Identification of the sibling species of the *Drosophila willistoni* subgroup through the electrophoretic mobility of acid phosphatase-1. *J. Zool. Syst. Evol. Res* 44:212-216.

Grumblin G, Strelets V (2006) FlyBase: anatomical data, images and queries. *Nucleic Acids Res* 34:D484-488.

Hammer SE, Strehl S e Hagemann S (2005) Homologs of *Drosophila P* transposons were mobile in zebrafish but have been domesticated in a common ancestor of chicken and human. *Mol Biol Evol* 22:833-844.

Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG e Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673-4680.

Houck MA, Clark JB, Peterson KR, Kidwell MG (1991). Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. *Science* 253:1125-1129.

Hua-Van A, Le Rouzic A, Maisonhaute C e Capy P (2005) Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. *Cytogenet. Genome Res* 110:426-440.

Hurst GDD e Werren JH (2001) The role of selfish genetic elements in eukaryotic evolution. *Nat Rev Genet* 2:597-606.

International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921.

Jordan IK e McDonald JF (1998) Evolution of  *copia* retrotransposon in *Drosophila melanogaster* species subgroup. *Mol. Biol. Evol.* 15:1160-1171.

Jordan IK e McDonald JF (1998) Interelement selection in the regulatory region of the  *copia* retrotransposon. *J. Mol. Evol.* 47:670-676.

Jordan IK, Matyunina VL e McDonald JF (1999) Evidence for the recent horizontal transfer of long terminal repeat Retrotransposon. *Proc. Natl. Acad. Sci. USA.* 96.22: 12621-12625.

Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, Ashburner M e Celniker SE (2002) The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol* 3:0084.1-0084.20.

Kapitonov VV e Jurka J (2003) Molecular paleontology of transposable elements in the *Drosophila melanogaster* genome. *Proc Natl Acad Sci USA* 100:6569-6574.

Kastanis P, Eliopoulos E, Goulielmos GN, Tsakas S e Loukas M (2003) Macroevolutionary relationships of species of *Drosophila melanogaster* group based on mtDNA sequences. *Mol Phylogenet Evol* 28:518-528.

Kastritsis CD e Dobzhansky T (1967) *Drosophila pavlovskiana*, a race or a species? Amer. Midl. Nat. 78:244-247.

Kent WJ (2002) BLAT-the BLAST-like alignment tool Genome Res 12:656-664.  
Kidwell MG e Lisch DR (2001) Perspective: transposable elements, parasitic DNA, and genome evolution. Evolution Int J Org Evolution 55:1-24.

Kidwell MG e Lisch DR (2002) Transposable elements as sources of genomic variation. In: Craig NL, Craigie R, Gellert M e Lambowitz AM (eds) Mobile DNA II. ASM Press, Washington, pp 59-90.

Klein CC, Essi L, Golombieski RM e Loreto ELS (1999) Disgenesia do híbrido em populações naturais de *Drosophila melanogaster*. Ciência e Natura 21: 7-20.

Ko Harada, Yukuhiro K e Mukai T (1990) Transposition Rates of Movable Genetic Elements in *Drosophila melanogaster*. Proc Natl Acad Sci USA 87.8:3248-3252.

Kumar A e Bennetzen J (1999). Plant retrotransposons. Annu. Rev. Genet. 33:479-532.

Labrador M e Corces VG (1997) Transposable element-host interactions: regulation of insertion and excision. Annu Rev Genet 31:381-404.

Loreto ELS, Carareto CMA e Capy P (2008) Revisiting horizontal transfer of transposable elements in *Drosophila*. Uncorrected proof – Heredity 1:10.

Lozovskaya E, Hartl DL e Petrov DA (1995) Genomic regulation of transposable elements in *Drosophila*. Curr Opin Genet Dev 5:768-773.

Martin G, Wiernasz D e Scheld P (1983) Evolution of *Drosophila* repetitive-dispersed DNA. J. Mol Evol. 19: 203-213.

McDonald JF (1993) Evolution and consequences of transposable elements. Curr Opin Genet Dev. 3:855-864.

Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A e Rafalski A (2005) Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. Nat Genet 37:997-1002.

Mount S e Rubin GM (1985) Complete nucleotide sequence of the *Drosophila* transposable element *copia*: homology between *copia* and retroviral proteins. Mol. Cell. Biol. 5:1630-1638.

Nekrutenko A e Li WH (2001) Transposable elements are found in a large number of human protein-coding genes. Trends Genet 17:619-621.

Nicholas, KB e Nicholas HB Jr (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author. <http://www.psc.edu/biomed/genedoc>.

Ortiz, MF e Loreto ELS (2008) Characterization of new hAT transposable elements in 12 *Drosophila* genomes. *Genetica* DOI 10.1007/s10709-008-9259-5.

Rohde C, Garcia ACL, Valiati VH e Valente VLS (2006) Chromosomal evolution of sibling species of the *willistoni* group of *Drosophila*. I. Chromosomal arm IIR (Muller's element B). *Genetica* 126:77-88.

Russo AMC, Takezaki N e Nei M (1995) Molecular phylogeny and divergence times of drosophilid species. *Mol. Biol. Evol.* 12:391-404.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning - A Laboratory Manual*. 2ed. Cold Spring Harbour Laboratory Press, New York.

Sánchez-Gracia A, Maside X e Charlesworth B (2005) High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet* 21:200-203.

SanMiguel P e Bennetzen JL (1998) Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Annals of Botany* 82:37-44.

SanMiguel P, Tikhonov A, Jin Y, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z e Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274:765-768.

SanMiguel P, Gaut BS, Tikhoniv A, Nakajima Y e Bennetzen, JL (1998) The paleontology of intergene retrotransposons in maize. *Nature Genet.* 20:43-45.

Santos RA e Valente VLS (1990) On the occurrence of *Drosophila paulistorum* Dobzhansky & Pavan (Diptera, Drosophilidae) in urban environment: ecological and cytological observations. *Evol.Biol.* 4: 253-268.

Sassi, A. K., Heredia, F, Loreto, E. L. da S., Valente, V. L. da S., Rodhe, C. Transposable elements P and gypsy in natural populations of *Drosophila willistoni*. *Genetics and Molecular Biology, Ribeirao Preto SP*, v. 28, n. 4, p. 734-739, 2005.

Silva JC e Kidwell MG (2000). Horizontal transfer and selection in the evolution of P elements. *Mol Biol Evol* 17: 1542–1557.

Silva JC, Loreto EL e Clark JB (2004) Factors that affect the horizontal transfer of transposable elements. *Curr Issues Mol Biol* 6:57-71.

Spassky B, Richmond RC, Perez-Salas S, Pavlovsky OA, Mourao CA, Hunter AS, Hoenigsberg HF, Dobzhansky T, Ayala FJ (1971) Geography of sibling species related to *Drosophila willistoni*, and the semi-species of the *Drosophila paulistorum* complex. *Evolution* 25:129–143.

Spassky B (1957) Morphological differences between sibling species of *Drosophila*. Univ Texas Publ 5721:48–61.

Staden R (1996) The Staden sequence analysis package. Mol Biotechnol 5:233-241.

Sturtevant AH (1916) Notes on North American Drosophilidae with descriptions of 23 new species. Ann. Entomol. Soc. Am. 9:323–343.

Sturtevant AH (1939) On the subdivision of the genus *Drosophila*. Proc Natl Acad Sci USA 25:137-141.

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

Tamura K, Dudley J, Nei M e Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Mol. Biol. Evol. 24:1596-1599. 2007.

Throckmorton LH (1975) The phylogeny, ecology and geography of *Drosophila*. In: King RC (ed) Handbook of genetics. V 3: Invertebrates of genetic interest. Plenum, New York, pp 421-469.

Val FC, Marques MD e Vilela CR (1981) Drosophilidae of Neotropical region. In: Ashburner M, Carson HL e Thompson JN (eds) The Genetics and Biology of *Drosophila*. V 3. Academic Press, Orlando, pp123-168.

Volff JN (2006) Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. Bioessays 28:913-922.

White SE, Habera LF e Wessler SR (1994) Retrotransposons in the flanking regions of normal plant genes: A role for copia-like elements in the evolution of gene structure and expression. Proc Natl Acad Sci USA 91:11792-11796.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P e Schulman AH (2007) A unified classification system for eukaryotic transposable elements. Nature Rev Genet 8:973-982.

Winge H (1965) Interspecific hybridization between the six cryptic species of *Drosophila willistoni* group. Heredity 20:9-19.

Winge H. (1971) Níveis de Divergência Evolutiva no Grupo Críptico da *Drosophila willistoni*. Tese. Universidade do Rio Grande do Sul, Porto Alegre, Brasil.

Xing J, Wang H, Belancio VP, Cordaux R, Deininger PL e Batzer MA (2006) Emergence of primate genes by retrotransposon-mediated sequence transduction. Proc Natl Acad Sci USA 103:17608-17613.

Yang Y, Zhang Y, Qian Y e Zeng Q (2004) Phylogenetic relationships of *Drosophila melanogaster* species group deduced from spacer regions of histone gene H2A-H2B. *Mol Phylogenet Evol* 30:336-343.