

Julcimary Ricci

Análise da ocorrência de elementos de transposição em regiões reguladoras dos genes da família *Cyp* em espécies de *Drosophila*

São José do Rio Preto
2009

Julcimary Ricci

Análise da ocorrência de elementos de transposição em regiões reguladoras dos genes da família *Cyp* em espécies de *Drosophila*

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Genética , junto ao Programa de Pós-Graduação em Genética, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Orientadora: Prof^a. Dr^a. Claudia Marcia Aparecida Carareto

São José do Rio Preto
2009

Ricci, Julcimary.

Análise da ocorrência de elementos de transposição em regiões reguladoras dos genes da família *Cyp* em espécies de *Drosophila*/
Julcimary Ricci - São José do Rio Preto : [s.n.], 2009.

128 f. : il. ; 30 cm.

Orientadora: Cláudia Marcia Aparecida Carareto
Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas

1. Genética. 2. Genes *Cyps*. 3. Elementos de transposição. 4. DNAREP1_DM. 5. resistência a inseticidas. 6. regulação da expressão gênica. 7. grupo *melanogaster* I. Carareto, Cláudia Marcia Aparecida. II. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU - 575

Ficha catalográfica elaborada pela Biblioteca do IBILCE
Campus de São José do Rio Preto - UNESP

JULCIMARY RICCI

Análise da ocorrência de elementos de transposição em regiões reguladoras dos genes da família *Cyp* em espécies de *Drosophila*

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Genética, junto ao Programa de Pós-Graduação em Genética, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

BANCA EXAMINADORA

Prof^ª. Dr^ª. Claudia Marcia Aparecida Carareto
Professor Titular
UNESP – São José do Rio Preto
Orientadora

Prof. Dr. Ricardo De Marco
Professor Doutor
USP – São Carlos

Prof. Dr. André Luís Laforga Vanzela
Professor Assistente Doutor
UEL – Londrina

São José do Rio Preto, 27 de Abril de 2009.

O presente trabalho foi realizado no Departamento de Biologia do Instituto de Biociências, Letras e Ciências Exatas de São José do Rio Preto, da Universidade Estadual Paulista, sob a orientação da Profa. Dra. Claudia Marcia Aparecida Carareto, com bolsa de Mestrado da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

Dedico este estudo à meus pais Arlindo e Sonia, por me ensinarem a viver honestamente, a enfrentar situações difíceis sempre com sinceridade e amor, a ser ter humildade e sempre respeitar as pessoas, e por terem me dado a condição de poder chegar até esta etapa. As minhas irmãs Bel e Meire, pelo carinho, apoio e respeito em todos momentos. Ao meu amor Fernando, pela compreensão, paciência, pela força e principalmente pelo amor. A minha orientadora Profa. Claudia Carareto pela paciência e pelo imenso aprendizado.

Obrigada!!!

AGRADECIMENTOS

À minha orientadora, **Profa. Dra. Claudia Carareto**, pela oportunidade e confiança, pela paciência e persistência, pela imensa dedicação, responsabilidade e honestidade em atuar como orientadora, e pelo grande aprendizado.

À **Dra. Cristina Vieira**, pela colaboração no desenvolvimento deste trabalho, pelas críticas e sugestões, e pelo aprendizado em tão pouco tempo de convivência.

À **Profa. Dra. LÍlian Madi-Ravazzi**, pela realização dos bioensaios utilizados nas análises populacionais.

Ao doutorando e colaborador deste projeto **Ba. Fabrício Ramon Lopes**, pela imensa paciência, atenção e, principalmente, pelo aprendizado das técnicas de bioinformática.

Ao **Prof. Dr. Luís Gustavo Galego**, pela confiança, pelo enorme apoio e aprendizado que desde a graduação tive o privilégio de receber, e pela oportunidade de chegar até aqui.

À **Profa. Dra. Hermione Elly Melara de Campos Bicudo**, pela sua imensa experiência e dedicação como educadora e pesquisadora, pelo amor com que nos ensina e pelo exemplo de como ser uma pessoa elegante, respeitando a todos independentemente de sua posição.

À **Profa. Dra. LÍlian Castiglioni**, pelo incentivo, pelo exemplo de educadora e pelo carinho que jamais esquecerei.

Aos meus queridos companheiros e amigos do Laboratório de Evolução Molecular, onde trabalhamos muito, aprendemos bastante, mas também demos muita risada. Foram com eles que passei a maior parte dos meus dias nestes últimos anos. Ao **Luis Gustavo, Fabrício, Nathalia, Adriana, Leliane, Elaine, LÍlian e Elias**, meu muito obrigada pelo aprendizado, pelo apoio, pela amizade, carinho e principalmente pela força.

Aos colegas, **Marina, Marcelo, Márcia, LÍlian, Hederson e Maza** que tive o privilégio de conhecer e conviver durante esse período de pós-graduação e, principalmente, a duas pessoas especiais, **Mariana e Leiza**, pelo apoio, força, amizade e pelas risadas.

Ao meu avô **Walter** e a minha avó **Tercília**, por serem exemplo de pessoas honestas e fortes, pelo amor e dedicação, e por serem os responsáveis por nossa grande família, cheia de união, amor, paz e amizade; coisa rara de ser ver nos dias de hoje. É isso que me faz feliz!

A todos **meus tios, tias, primos, primas** (são muitos) e **pelos irmãos na fé**, pelo apoio, força, incentivo e principalmente pelo amor de todos, que me levantou em cada dificuldade.

À minha tia **Cláudia** em especial, pelo exemplo de força, de fé, de pessoa batalhadora e humilde que venceu mais uma etapa muito difícil em sua vida, sempre com muito ânimo, alegria e coragem.

Ao meu amor **Fernando**, pessoa maravilhosa que Deus colocou em minha vida e que tem sido a minha alegria. Agradeço pela paciência, pela força, pelas palavras de apoio e conforto e pelo seu imenso amor, que me ajudou a vencer essa etapa difícil da minha vida.

As minhas irmãs **Bel** e **Meire**, e ao **Leandro** pelo apoio constante, paciência, força, ajuda e pelo amor e respeito em todo esse processo.

Aos meus pais, Arlindo e Sonia, que sempre me ofereceram oportunidades únicas, como a realização desse estudo. Pais unidos, um lar abençoado com muita paz, amor e carinho! Exemplo de pessoas honestas, humildes, e que passaram por muitas batalhas e venceram a todas, sempre juntos. Agradeço pela imensa paciência, pelo apoio e por sempre me darem forças quando tudo parecia não ter saída! Obrigada meus queridos pais por me ensinarem sempre o caminho certo!!! Deus sempre continuará vos abençoando!!!

Amo vocês!!!

A Deus, minha eterna gratidão por ter iluminado meus passos em todos os dias da minha vida. Pelo amor, pela força, pela saúde, pelo ânimo, pela alegria nas horas de angústia, por ser minha condição em cada etapa. Por aprender a retribuir aquele que te afligir sempre com amor e humildade.

*“A mente que se abre a uma nova idéia jamais voltará ao seu tamanho original.
Procure ser um homem de valor, em vez de ser um homem de sucesso”*

(Albert Einstein)

“O mundo pirou, enlouqueceu. Agora amor é cafona, sinceridade é careta, pudor é ridículo e sentimento é bobagem. Roubar pode, envelhecer não. A sociedade só pensa em imagem, estética medidas e beleza. Não importa os sentimentos, a ajuda, a amizade, não importa o outro”

(Herbert Vianna)

*Respeitar as opções do outro em qualquer aspecto, é uma das maiores virtude que um ser humano pode ter. As pessoas são diferentes, agem diferente e pensam diferente. Nunca Julgue, apenas compreenda.
Respeito é bom sempre.*

"Aquilo que não nos mata, torna-nos mais fortes"

(Friedrich Wilhelm Nietzsche)

“Estou convencido das minhas próprias limitações e esta convicção é minha força”

(Mahatma Gandhi)

“Um homem cheio de si é sempre vazio”

(Pégismanet)

RESUMO

A resistência aos inseticidas é um modelo de processo evolutivo onde o inseticida atua como agente seletivo e, como resposta à seleção, ocorre a evolução da resistência nas populações de insetos. As enzimas citocromo P450 monooxigenases (CYP) formam uma família responsável pela resistência aos inseticidas. Tem sido proposto que a inserção de elementos transponíveis (TEs) em regiões reguladoras ou codificadoras dos genes da família *Cyp* pode alterar a expressão gênica e induzir a resistência aos inseticidas. No presente estudo foram realizadas análises *in silico* que permitiram identificar a ocorrência de inserções de fragmentos de TEs em 35 genes *Cyps* com diferentes funções, e em seus genes flanqueadores, em *Drosophila melanogaster* e *D. simulans*, além de 13 genes *Cyps* de seis espécies do grupo *melanogaster* de *Drosophila*. As inserções de TEs ocorreram principalmente nas regiões flanqueadoras 5' dos *Cyps* associados à resistência aos inseticidas e à função monooxigenase geral. Os resultados não indicaram qualquer relação entre a distância em relação ao gene e o número de inserções. As análises mostraram que a maioria das inserções pertence à classe de transposons de DNA, sendo o transposon DNAREP1_DM o que apresentou o maior número de cópias. O fato de essas seqüências apresentarem putativos sítios de ligação de fatores de transcrição sugere que possam desempenhar algum papel na regulação dos genes *Cyps*. Também foi analisada a ocorrência de polimorfismos de inserção de TEs em regiões flanqueadoras de genes da família *Cyp*, em diferentes linhagens geográficas resistentes e suscetíveis, de *D. melanogaster* e *D. simulans*. Análises evidenciaram a presença de polimorfismo interpopulacional de tamanho das regiões flanqueadoras dos genes *Cyp6w1*, *Cyp6a2* e *Cyp12d1*, porém, não indicaram ocorrência de associação entre a presença de TEs e a resistência aos inseticidas. Os resultados também mostraram que a inserção de TEs nas proximidades dos genes *Cyps* associados à resistência é diferencial entre as seis espécies do grupo *melanogaster*, não seguindo a proporção genômica de TEs em cada espécie, o que também sugere que esses TEs possam desempenhar um papel adaptativo nessas espécies.

Palavras-chave: *Cyps*. elementos de transposição. DNAREP1_DM. regulação da expressão gênica. resistência aos inseticidas. grupo *melanogaster*.

ABSTRACT

Insecticide resistance is a model of evolutionary process where the insecticide acts as the selective agent and resistance in the insect populations evolves as an answer to selection. Cytochrome monooxygenases (CYP) is family of enzymes responsible for the insecticide resistance. It has been proposed that insertion of transposable elements (TEs) in regulatory or coding regions of the *Cyp* genes can alter gene expression and induce insecticide resistance. In the present study *in silico* analyses allowed identifying the insertion of TE fragments in 35 *Cyp* genes with different functions, in *Drosophila melanogaster* and *D. simulans*, as well as in 13 *Cyps* of six species of the *melanogaster* group of *Drosophila*. The TE insertions occurred mainly in the 5' flanking regions of *Cyp* genes associated to resistance and to those with a general monooxygenase function. The results did not indicate any relationship between the number of insertions and the distance in relation to the gene. The analyses showed that most of the insertions belong to the DNA transposon class, being DNAREP1_DM the most numerous. Since this element carry putative binding sites of transcription factors it can be suggested they play same role in gene regulation. The polymorphism of TE insertions in the flanking regions of *Cyp6w1*, *Cyp6a2* and *Cyp12d1*, genes associated to resistance, found in resistant and as well as in susceptible geographical strains of *D. melanogaster* and *D. simulans*, does not indicate any relationship between the presence of TEs in those regions and the insecticide resistance. The results also showed that the insertions of TEs in the proximities of the *Cyps* associated to resistance is differential among six species of the *melanogaster* group, not following the genomic proportion of TEs in each species. These results also suggest that TEs inserted in the *Cyp* flanking regions can carry out an adaptive role for those species.

Keywords: *Cyps*. transposable elements. DNAREP1_DM. gene expression regulation. insecticide resistance. *melanogaster* group.

SUMÁRIO

1	INTRODUÇÃO GERAL	13
1.1	<i>Ocorrência e atuação dos elementos transponíveis no</i>	14
1.2	<i>Relação dos elementos transponíveis com resistência a inseticidas</i>	17
2	CAPÍTULO I	21
2.1	INTRODUÇÃO	22
2.2	OBJETIVOS	26
2.3	MATERIAL E MÉTODOS	27
2.3.1	<i>Análises in silico</i>	27
2.3.2	<i>Análises populacionais</i>	28
2.3.2.1	<i>Espécies</i>	28
2.3.2.2	<i>Amostras de DNA e reação de PCR</i>	29
2.3.2.3	<i>Clonagem e sequenciamento</i>	30
2.3.2.4	<i>Análise das seqüências</i>	30
2.4	RESULTADOS	32
2.4.1	<i>Inserções de TEs em genes Cyps com funções distintas</i>	32
2.4.2	<i>Distribuição de TEs em regiões intergênicas e gênicas dos genes Cyps e seus flaqueadores</i>	37
2.4.3	<i>Ocorrência de inserções de TEs em regiões flaqueadoras distantes em até 3kb do gene</i>	39
2.4.4	<i>Distribuição das classes e ordens de TEs nas regiões intergênicas</i>	44
2.4.5	<i>Ocorrência do elemento DNAREP1_DM em genes Cyps e em seus flaqueadores em D. melanogaster e D. simulans</i>	46
2.4.6	<i>Investigação da presença de sítios de ligação de fatores de transcrição (FTs) nas cópias do DNAREP1_DM dos genes Cyps</i>	51
2.4.7	<i>TEs em regiões flaqueadoras de genes Cyps em linhagens resistentes e suscetíveis de D. melanogaster e D. simulans</i>	55
2.5	DISCUSSÃO	60
2.5.1	<i>Inserção de TEs em genes Cyps e seus flaqueadores</i>	61
2.5.2	<i>Inserção do elemento DNAREP1_DM</i>	63
2.5.3	<i>Análises de seqüências de TEs em regiões flaqueadoras de genes Cyps em linhagens de D. melanogaster e D. simulans resistentes e suscetíveis a inseticidas</i>	64
2.6	CONCLUSÃO	65
	REFERÊNCIAS	66
	APÊNDICE	73

SUMÁRIO

3	CAPÍTULO II	86
3.1	INTRODUÇÃO	87
3.2	OBJETIVOS	89
3.3	MATERIAL E MÉTODOS	90
3.3.1	<i>Análises de Bioinformática</i>	90
3.3.2	<i>Análises Evolutivas</i>	92
3.4	RESULTADOS	93
3.4.1	<i>Ocorrência de inserções de TEs em genes Cyps e em suas regiões flanqueadoras 5' e 3' distantes até 3 kb do início da transcrição do gene</i>	93
3.4.2	<i>DNAREP1_DM e Helitron-1_Dyak: os TEs mais freqüentes nos Cyps analisados nas espécies do grupo melanogaster</i>	104
3.5	DISCUSSÃO	108
3.6	CONCLUSÃO	111
	REFERÊNCIAS	112
4	DISCUSSÃO GERAL	117
5	CONCLUSÕES	121
	REFERÊNCIAS	123

1 INTRODUÇÃO GERAL

Elementos de transposição (TEs) são seqüências de DNA medianamente repetidas capazes de se movimentarem nos genomas. Encontrados na maioria dos organismos, foram descobertos na década de 50 por Barbara McClintock (McCLINTOCK, 1957). Observando o padrão de coloração dos grãos, em estudos realizados com milho, McClintock propôs que os TEs agiam como “elementos controladores” da atividade gênica. Durante quase meio século, muitos cientistas, por não conhecerem completamente a função desses elementos, os consideraram como DNA egoísta ou parasita (DOOLITTLE; SAPIENZA, 1980; ORGEL; CRICK, 1980) ou mesmo DNA lixo (OHNO; YOMO, 1991), pois imaginavam que os TEs agiam nos genomas apenas para garantir sua sobrevivência. Contudo, estudos posteriores mostraram que essas seqüências também podem ter efeitos positivos, gerando mutações e rearranjos cromossômicos que contribuem significativamente para a evolução da estrutura e função gênica e do genoma de várias espécies (BROSIUS, 1999; MAKALOWSKI, 2000; 2003; KIDWELL, 2002; LORENC; MAKALOWSKI, 2003; KIDWELL; HOLYOAKE, 2001; FONTANILLAS et al., 2007).

Os TEs foram classificados inicialmente por Finnegan (1989) que, de acordo com seu modo de transposição, agrupou-os em duas classes: retrotransposons (classe I) e transposons de DNA (classe II). Os elementos da classe I são aqueles que se transpõem via um RNA intermediário (o qual é transcrito reversamente) pelo mecanismo “*copy-and-paste*”, gerando assim, múltiplas cópias no genoma. Por outro lado, os elementos da classe II, denominados transposons de DNA, apresentam atividade transposicional mediada pela transposase e se movimentam pelo mecanismo “*cut-and-paste*”. Os elementos da classe I codificam algumas proteínas, dentre elas a proteína do capsídio (GAG), a proteinase aspártico (AP), a integrase (INT), a transcriptase reversa (RT), a RNase (RH), a proteína do envelope (ENV) e a endonuclease apurínica (APE). Os TEs da classe II codificam outras proteínas, tais como a transposase (Tase), a tirosina recombinase (YR), a proteína de replicação A (RPA), a helicase (HEL), a protease cisteína (CYP), a DNA polimerase B (POL) e a integrase C (C-INT). Recentemente, Wicker et al. (2007) propuseram um sistema de classificação unificado para os TEs, também com base em seu mecanismo de transposição, similaridades de seqüência e características estruturais. Na nova classificação, mantiveram-se as duas classes já usuais (retrotransposons e transposons de DNA), mas a classe II foi subdividida em duas subclasses.

Os retroelementos da classe I foram subdivididos em cinco ordens: de acordo com sua estrutura, organização e filogenia da transcriptase reversa (RT). A ordem *LTR* (do Inglês: *Long Terminal Repeat*), que apresenta como característica uma longa repetição terminal, sítio de duplicação e domínios codificantes de proteínas (GAG e POL), foi subdividida em cinco superfamílias: *Copia*, *Gypsy*, *Bel-Pao*, *Retrovirus* e *ERV*. A ordem *DIRS*, cujos componentes não geram sítios de duplicação no local de inserção, é representada pelas superfamílias *DIRS*, *Ngaro* e *VIPER*. A ordem *PLE* (*Penelope-like elements*) contém elementos que codificam apenas a transcriptase reversa (RT) e a endonuclease (EN), contidos em apenas uma superfamília (*Penélope*). As ordens *LINE* (Long Interspersed Nuclear Elements) e *SINE* (Short Interspersed Nuclear Elements), por sua vez, não apresentam *LTRs*. Os elementos *LINE* estão agrupados em cinco superfamílias (*R2*, *RTE*, *Jockey*, *L1* e *I*) e, os *SINEs*, em três (*tRNA*, *7SL RNA* e *5sRNA*).

As duas subclasses da classe II diferem quanto ao número de fitas de DNA cortadas durante a excisão: na subclasse 1, as duas fitas são cortadas e, na subclasse 2, apenas uma fita é cortada. A subclasse 1 é composta pela ordem *TIR* (do Inglês: *Terminal Inverted Repeat*), que apresenta repetições terminais invertidas e sítio de duplicação no local de inserção no genoma, e possui nove superfamílias (*Tc1-Mariner*, *hAT*, *Mutator*, *Merlin*, *Transib*, *P*, *PiggyBac*, *PIF – Harbinger*, *CACTA*) e, pela ordem *Crypton*, com apenas uma superfamília (*Crypton*), cujos transposons não possuem *TIRs*. A subclasse 2 é representada por duas ordens, com apenas uma superfamília cada: *Helitron* (*Helitron*) e *Maverick* (*Maverick*). Os *Helitrons* não geram duplicação no sítio alvo de inserção e apresentam um curto *hairpin* na extremidade 3'; enquanto os elementos *Maverick*, por sua vez, possuem longas *TIRs*.

1.1 Ocorrência e atuação dos elementos transponíveis no genoma

Devido a sua capacidade de se replicarem aumentando o número de cópias, e de se moverem para porções diferentes do genoma, os TEs constituem uma parcela representativa (revisão em KIDWELL; LISCH, 2000; 2001; WICKER et al., 2007) e estão envolvidos na evolução do tamanho dos genomas (PETROV, 2001; KIDWEL, 2002). Representam uma grande parcela do genoma de plantas, alcançando cerca de 80 % em *Zea mays* (MEYERS et al., 2001);

dentre os animais, compreendem cerca de 45 % do genoma humano (VAN DE LAGEMAAT et al., 2003; POLAVARAPU et al., 2008); 12 % do genoma do verme *Caenorhabditis elegans* (GANKO et al., 2003); e, do genoma de *Drosophila melanogaster* 5,5 % (BERGMAN et al., 2006), mas apenas 2,73 % em sua espécie críptica *D. simulans* (CLARCK et al., 2007).

A maior parte das inserções de TEs está localizada em regiões intergênicas, introns e UTRs, possuindo desse modo maior chance de serem mantidas no genoma, pois, estão sob menor pressão seletiva (KIDWELL, 2002; SELA et al., 2007; FONTANILLAS et al., 2007; YANG; BARBASH, 2008). Segundo Jordan et al. (2003), inserções de TEs em regiões promotoras de genes de mamíferos, próximas do sítio de início da transcrição do gene, são possivelmente deletérias sendo, portanto, eliminadas por seleção. Por outro lado, quando inseridos em regiões reguladoras, os TEs também podem afetar positivamente a expressão do gene, como também de genes adjacentes, devido à introdução de elementos reguladores (CONTE et al., 2002; JORDAN et al., 2003; THORNBURG et al., 2006). Em diversos TEs, foram identificadas seqüências que atuam como sítios de ligação de fatores transcrição (*TFBS*), nas quais se ligam motivos reguladores da expressão gênica (THORNBURG et al., 2006; ALMEIDA; CARARETO, 2006; POLAVARAPU et al., 2008).

Os TEs podem também se inserir em regiões codificadoras de proteínas. Tem sido mostrado que estas inserções, inicialmente deletérias, podem também ser mantidas por gerarem mudanças adaptativas, caracterizando o evento de exaptação, ou domesticação (NEKRUTENKO; LI, 2001; LORENC; MAKALOWSKI, 2003; LOPES et al., 2008). Três principais tipos de inserções relacionadas a regiões codificantes são encontrados nos genomas, com diferentes resultados evolutivos: 1) em introns próximos a um éxon, podendo posteriormente se tornar em uma extensão do éxon; 2) em introns, de modo a se tornar um novo éxon, caracterizando um evento de exonização; 3) diretamente nos éxons, podendo alterar a seqüência protéica e, conseqüentemente, induzir a modificações no fenótipo. A Figura 1 ilustra os três tipos.

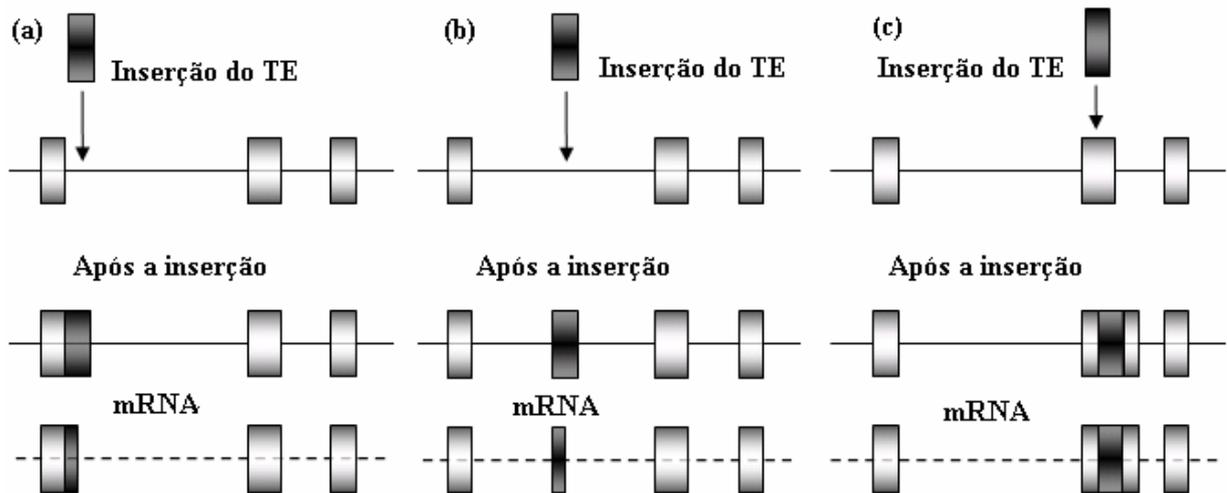


Figura 1. Inserção de TE em regiões codificantes de proteínas. Inserção do TE (a) em introns próximos a um éxon, podendo se tornar extensão do éxon; (b) em introns, de modo a se tornar um novo éxon e (c) diretamente no éxon (elaborada por F. R. Lopes).

As inserções de TEs em regiões codificantes têm sido observadas preferencialmente em nível transcricional, por serem possivelmente deletérias em nível protéico; como reportado em humanos (NEKRUTENKO; LI, 2001; GOTEA; MAKALOWSKI, 2006; SELA et al., 2007), vertebrados (LORENC; MAKALOWSKI, 2003), camundongo (SELA et al., 2007) e café (LOPES et al., 2008). Entretanto, domesticação de TEs, isto é, manutenção de suas seqüências na proteína tem sido também relatada (MILLER et al., 1999, BUNDOCK; HOOYAKAAS, 2005, REISS et al., 2005; GOTEA; MAKALOWSKI, 2006). Desta maneira, os TEs podem ser considerados fontes de variabilidade para o genoma, além de desenvolverem um importante papel na evolução gênica e em sua regulação (revisão em KIDWELL; LISCH, 2000; MAKALOWSKI, 2000; JORDAN et al., 2003; VAN DE LAGEMAAT et al., 2003; GANKO et al., 2003; 2006; THORNBURG et al., 2006).

Dentre as 12 espécies de *Drosophila* que possuem o genoma seqüenciado, a porcentagem de inserções de TEs varia de 2,73 % em *D. simulans* e 24,93 % em *D. ananassae*, respectivamente as espécies com a menor e a maior porcentagem de TEs (CLARCK et al., 2007). Dentre esses elementos, o transposon DNAREP1_DM é o mais abundante nesses 12 genomas, apresentando centenas a milhares de cópias.

D. melanogaster e *D. simulans* diferem quanto à porcentagem genômica de TEs. Na primeira espécie, a porcentagem é praticamente o dobro (5,35 %) do valor observado na segunda (2,73 %). Nas duas espécies, os TEs têm sido associados à evolução da resistência aos inseticidas, por estimularem a expressão de genes responsáveis pela resistência metabólica, ou mesmo, por atuarem na geração de novas funções protéicas devido à alteração da estrutura gênica (DABORN et al., 2002; AMINETZACH et al., 2005).

1.2 Relação dos elementos transponíveis com resistência aos inseticidas

A resistência aos inseticidas pode ser causada por mutações de ponto, amplificação gênica e elevada transcrição. Mutações de ponto podem alterar as proteínas-alvo dos inseticidas, impedindo a ligação do inseticida e, conseqüentemente, sua ação no organismo, gerando a resistência (FFRENCH-CONSTANT et al., 1998; AMICHOT et al., 2004). Por outro lado, a amplificação gênica aumenta o número de cópias dos genes e, conseqüentemente, sua expressão, aumentando o número de proteínas-alvo dos inseticidas. A terceira causa da resistência se deve a alteração da regulação dos genes envolvidos na resistência (revisão em LI et al., 2007). Essa via se caracteriza pela ocorrência de alterações genômicas, como por exemplo, inserções de TEs, que podem elevar a expressão dos genes codificadores das enzimas responsáveis pelo metabolismo dos inseticidas (DABORN et al., 2002).

O metabolismo de xenobióticos, como os inseticidas, é composto por dois tipos de reações enzimáticas (denominadas fases), as quais tornam o substrato mais solúvel em água, facilitando assim sua eliminação. As enzimas de fase I (P450 monooxigenases) são responsáveis pelo metabolismo oxidativo de muitos xenobióticos (biotransformação), enquanto que as de fase II (glutathione-S-transferases - GSTs) agem como inativadoras dos produtos da fase I, tornando-os passíveis de excreção. Nesta fase, ocorre a conjugação com um substrato endógeno por meio das GSTs (conjugação). As enzimas P450 monooxigenases agem principalmente no fígado, contudo, existe expressão por todo o organismo, incluindo intestino, túbulos de Malpighi, pulmões, rins e corpo gorduroso (McCART; FFRENCH-CONSTANT, 2008; CHUNG et al., 2009).

Em geral, os inseticidas possuem uma ação que leva à mortalidade rápida, por serem neurotóxicos. Os insetos podem apresentar resistência aos xenobióticos naturais (como defensivos químicos produzidos pelas plantas), bem como, aos inseticidas sintéticos. A resistência aos inseticidas envolve aumento no nível de metabolização das enzimas que transformam o xenobiótico, reduzindo a sensibilidade ao inseticida para se ligar ao seu sítio-alvo no sistema nervoso. Portanto, a detoxificação só é ativada se o inseticida não alcançar seu sítio alvo, sendo metabolizado e degradado pelas enzimas CYP P450, caracterizando assim, indivíduos resistentes (revisão em HEMINGWAY; RANSON, 2000; RANSON et al., 2002; DABORN et al., 2002; revisão em LI et al., 2007). Dentre os genes codificadores de proteínas relacionadas com a resistência metabólica aos inseticidas, por detoxificação de xenobióticos (drogas e outros compostos), destacam-se aqueles codificadores de três famílias de enzimas (carboxilesterases, glutationa-S-transferases e citocromo P450 monooxigenases).

O genoma de *D. melanogaster* tem aproximadamente 90 seqüências pertencentes à família multigênica citocromo P450 monooxigenase (*Cyps*), sendo 83 genes ativos e sete prováveis pseudogenes (TIJET et al., 2001). A família CYP possui genes envolvidos em uma ampla variedade de processos biológicos, como no desenvolvimento do organismo (biossíntese de hormônios ecdisona, juvenil e ferormônios) e no metabolismo de xenobióticos (inseticidas, drogas, carcinógenos). O nome citocromo P450 é derivado do fato que estas são proteínas celulares (cito) coloridas (cromo) com um pigmento com um comprimento de onda de 450 nm. A reação mais comum catalisada pelo citocromo P450 é uma reação monooxigenase, que se caracteriza pela inserção de um átomo de oxigênio em um substrato orgânico, enquanto o outro átomo oxigênio é reduzido à água.

A nomenclatura usual para os genes da superfamília CYP é grafada em itálico e é composta pelo prefixo *Cyp* seguido pelo número da família a qual ele pertence, uma letra que indica a subfamília e por um número que representa o gene, por exemplo, *Cyp6g1*. Para a designação da proteína, do RNAm ou cDNA, são utilizadas letras maiúsculas (CYP6G1). Membros de uma mesma família compartilham mais de 40% de identidade na seqüência de aminoácidos, enquanto que em subfamílias essa identidade corresponde a mais de 55% (revisão em HEMINGWAY; RANSON, 2000; TIJET et al., 2001).

Dentre os genes *Cyps* associados à resistência aos inseticidas em *D. melanogaster* destacam-se nove: *Cyp4e2* (AMICHOT et al., apud CHEN; LI, 2007), *Cyp6a2* (MAITRA et al.,

2000; PEDRA et al., 2004), *Cyp6a8* (MAITRA et al., 2000), *Cyp6a9* (MAITRA et al., 1996), *Cyp6g1* (DABORN et al., 2002; CATANIA et al., 2004; SCHLENKE; BEGUN, 2004), *Cyp6w1* (PEDRA et al., 2004), e os *Cyp12a4* (BOGWITZ et al., 2005), *Cyp12d1* (BRANDT et al., 2002) e *Cyp308a1* (*Gene Ontology*). Muitos dos genes *Cyps* responsáveis pela detoxificação de inseticidas em *D. melanogaster*, como por exemplo, o *Cyp6g1* (DABORN et al., 2001, BRANDT et al., 2002) e o *Cyp12d1* (BRANDT et al., 2002) estão localizados no loco *Rst(2)DDT*, ou *DDT-R* (DABORN et al., 2001).

Estudos realizados com espécies de *Drosophila* têm associado à resistência aos inseticidas, a superexpressão de genes *Cyps*, em decorrência da inserção de fragmentos de elementos de transposição em suas regiões reguladoras ou mesmo dentro dos genes. A inserção do elemento *Accord* a 291 pb antes do sitio de início de transcrição do gene *Cyp6g1* em *D. melanogaster* (DABORN et al., 2002; CATANIA et al., 2004; CHUNG et al. 2007), do elemento *Doc* no gene ortólogo ao *Cyp6g1* em *D. simulans* (SCHLENKE; BEGUN, 2004) e do transposon de DNA *Bari1* na região 3' do gene *Cyp12a4* (MARSANO et al., 2005; BOGWITZ et al., 2005), são alguns exemplos. Essa associação, contudo, não está completamente caracterizada, visto que tem sido relatado que em algumas linhagens susceptíveis, que apresentam a inserção de *Accord*, a expressão do *Cyp6g1* também é aumentada (SCHLENKE; BEGUN, 2004; FESTUCCI-BUSELLI et al., 2005; KURUGANTI et al., 2007). A relação entre a inserção de elementos transponíveis em genes associados à resistência aos inseticidas e a conseqüente evolução da resistência é um problema atual não solucionado, que merece, portanto, ser investigado. Neste trabalho, duas estratégias foram utilizadas para se investigar essa questão. A primeira foi investigar *in silico*, a ocorrência de inserções preferenciais de TEs em genes *Cyps* associados ao metabolismo de xenobióticos e, assim, inferir sua relação com a origem da resistência aos inseticidas em espécies de *Drosophila* do grupo *melanogaster*, com foco em *D. melanogaster* e *D. simulans*. A outra foi analisar, nessas duas espécies, a ocorrência de polimorfismos de inserção de TEs nas regiões intergênicas ou codificantes de genes *Cyps* em linhagens susceptíveis e resistentes a inseticidas, e avaliar se há correlação positiva entre a inserção dos TEs e a resistência. O conhecimento dos processos associados à evolução da resistência aos inseticidas e aos xenobióticos tem importância científica, como também, econômica e de saúde pública. Esclarecer o papel dos elementos transponíveis nesse processo é, portanto, uma questão que merece ser investigada.

2.1 INTRODUÇÃO

A análise do tema resistência aos inseticidas é de particular interesse, visto que o uso destes compostos, em lavouras e centros urbanos, no controle de pragas, tem aumentado consideravelmente nas últimas décadas. Essa resistência pode ser causada por mutações de ponto, amplificação gênica e elevada transcrição. Mutações de ponto podem alterar as proteínas-alvo dos inseticidas, impedindo a ligação do inseticida e, conseqüentemente, sua ação no organismo, gerando a resistência (FFRENCH-CONSTANT et al., 1998; AMICHOT et al., 2004). Por outro lado, a amplificação gênica aumenta o número de cópias dos genes e, conseqüentemente, sua expressão, aumentando o número de proteínas-alvo dos inseticidas. A terceira causa da resistência se deve à alteração da regulação dos genes envolvidos na resistência. Essa via se caracteriza pela ocorrência de alterações genômicas, as quais elevam a expressão dos genes codificadores das enzimas responsáveis pelo metabolismo dos inseticidas (revisão em LI et al., 2007). Dentre os genes codificadores de proteínas relacionadas com a resistência metabólica aos inseticidas, por detoxificação de xenobióticos (drogas e outros compostos), destacam-se aqueles codificadores de três famílias de enzimas (carboxilesterases, glutathione-S-transferases e citocromo P450 monooxigenases). Neste estudo focamos a análise de genes que compõem a família das citocromo P450 monooxigenases (CYP).

O genoma de *D. melanogaster* tem aproximadamente 90 genes *Cyps*, sendo 83 genes ativos e sete prováveis pseudogenes (TIJET et al. 2001). A família CYP possui genes envolvidos em uma ampla variedade de processos biológicos, como no desenvolvimento do organismo (biossíntese de hormônios ecdisona, juvenil e ferormônios) e no metabolismo de xenobióticos (inseticidas, drogas, carcinógenos). Dentre os genes *Cyps* associados à resistência aos inseticidas em *D. melanogaster* destacam-se nove: *Cyp4e2* (AMICHOT et al., apud CHEN; LI, 2007), *Cyp6a2* (MAITRA et al., 2000; PEDRA et al., 2004), *Cyp6a8* (MAITRA et al., 2000), *Cyp6a9* (MAITRA et al., 1996), *Cyp6g1* (DABORN et al., 2002; CATANIA et al., 2004; SCHLENKE; BEGUN, 2004), *Cyp6w1* (PEDRA et al., 2004), e os *Cyp12a4* (BOGWITZ et al., 2005), *Cyp12d1* (BRANDT et al., 2002) e *Cyp308a1* (Gene acesso FBgn0030949). Em seis deles, a resistência aos inseticidas (a mais do que um composto), de maneira geral, está associada à elevada transcrição gênica, tais como o *Cyp6g1* (DDT, imidaclopride, malathion, nitenpyram,

dicyclanil, acetamipride, regulador de crescimento lufenuron), *Cyp6g2* (diazinon, nitenpyram), *Cyp6a2* (DDT, malathion), o *Cyp6a8* (imidaclopride, ácido láurico), o *Cyp12a4* (lufenuron), e o *Cyp12d1* (DDT, dicyclanil) (DABORN et al., 2002; 2007; LE GOFF et al., 2003; AMICHOT et al., 2004; HELVIG et al., 2004; BOGWITZ et al., 2005).

Diversos estudos têm associado à resistência aos inseticidas em *Drosophila* à inserção de elementos de transposição em genes *Cyps*. Inserções de TEs próximas aos sítios de início da transcrição dos genes são possivelmente deletérias sendo, portanto, eliminadas por seleção (JORDAN et al., 2003); entretanto, quando inseridos em regiões reguladoras, podem afetar a expressão do gene, como também de genes adjacentes, devido à introdução de motivos reguladores da transcrição (CONTE et al., 2002; JORDAN et al., 2003; THORNBURG et al., 2006).

Daborn et al. (2002) relataram que a resistência aos inseticidas, mapeada no loco *DDT-R* de *D. melanogaster*, se deve a superexpressão de um único gene P450, o *Cyp6g1*, e que um alelo relacionado à resistência tem sido distribuído globalmente. Esse alelo, ao contrário dos diversos relacionados à resistência, não é caracterizado por uma mutação de ponto, mas sim, pela inserção de 491 pb do elemento transponível *Accord* a 291 pb antes do sítio de início de transcrição do gene. Essa associação resulta em elevada expressão do *Cyp6g1* (DABORN et al., 2002; CATANIA et al., 2004; CHUNG et al. 2007). Schlenke e Begun (2004) reportaram a associação entre inserção de um TE e a resistência aos inseticidas em *D. simulans*. Nessa espécie, o elemento *Doc* está inserido na região anterior ao ortólogo *Cyp6g1* de *D. melanogaster*, e também está associado a superexpressão desse gene. Também Marsano et al. (2005), e Bogwitz et al. (2005), sugeriram que a presença do transposon de DNA *Bari1* na região 3' do gene *Cyp12a4* de *D. melanogaster* eleva a expressão desse gene. Esses trabalhos associam a inserção dos TEs à superexpressão dos genes *Cyps* e à resistência aos inseticidas. No entanto, estudos posteriores não corroboram essa relação, pois algumas linhagens suscetíveis de *D. melanogaster*, também continham o elemento *Accord* inserido na região flangeadora do gene *Cyp6g1* e apresentavam, da mesma forma, elevada expressão desse gene (SCHLENKE; BEGUN, 2004; FESTUCCIBUSELLI et al., 2005; KURUGANTI et al., 2007). Esses resultados indicam que existe associação entre a inserção do TE e o aumento da expressão gênica, mas não necessariamente a alta expressão desses genes está associada à resistência aos inseticidas e, que o fenótipo para resistência é um fenômeno complexo, que pode ser poligênico e determinado por diversos

fatores, dentre eles, a inserção de um TE, como a do elemento *Accord* na região 5' do gene *Cyp6g1* (KURUGANTI et al., 2007).

A investigação do papel adaptativo da inserção de TEs em genes *Cyps* foi também realizada por meio de análises *in silico*. Recentemente, Chen e Li (2007), pesquisando o genoma de *D. melanogaster*, analisaram a inserção de TEs em oito genes *Cyps* associados à resistência aos inseticidas e em um gene *Gst* (*Gstd1*), responsável pela detoxificação de xenobióticos e, em cinco genes *Cyps* relacionados à biossíntese de ecdisona e à regulação do desenvolvimento (*Cyp302a1*, *Cyp306a1*, *Cyp307a1*, *Cyp314a1*, *Cyp315a1*). Sete dos oito *Cyps* relacionados com a resistência aos inseticidas possuíam inserções de TEs (*Cyp6a9* foi o único gene *Cyp* que não possuía TEs). Por outro lado, todos os *Cyps* relacionados à biossíntese de ecdisona e ao desenvolvimento em *D. melanogaster* não apresentaram TEs inseridos nas regiões intergênicas 5' ou 3'. Chen e Li (2007) explicaram as diferenças observadas pelo fato que genes envolvidos em respostas ambientais como o estresse, e o metabolismo de xenobióticos, tendem a ser mais tolerantes a mudanças genômicas causadas pela inserção de TEs e, assim, apresentarem maior número de inserções (VAN DE LAGEMAAT et al., 2003). Essas mudanças podem alterar a regulação desses genes e, conseqüentemente, aumentar sua expressão, conferindo a resistência aos inseticidas via elevada expressão (revisão em LI et al., 2007), visto que os TEs podem afetar a transcrição de genes por conterem elementos reguladores em suas seqüências (SCHLENKE; BEGUN, 2004). Entretanto, genes com funções básicas essenciais para o desenvolvimento, são altamente conservados, estando sob uma forte seleção negativa contra inserções de TEs (VAN DE LAGEMAAT et al., 2003; LOWE et al., 2007).

Em decorrência da inserção preferencial de TEs em genes *Cyps* relacionados à detoxificação de xenobióticos, Chen e Li (2007) concluíram que os TEs podem ser seletivamente enriquecidos em genes relacionados a respostas ambientais a xenobióticos, mas relativamente excluídos de genes essenciais *housekeeping*, resultando em maior plasticidade genômica nos primeiros, e maior conservação nos últimos. Tais mudanças genômicas seriam vantajosas para organismos, tais como os insetos, que apresentam alto grau de plasticidade genética que os habilita a co-evoluir com seus ambientes altamente variáveis e seletivos.

Drosophila melanogaster e sua espécie críptica *D. simulans* são as principais espécies do gênero *Drosophila* utilizadas em estudos sobre a associação entre inserções de TEs e a resistência aos inseticidas. Embora estreitamente relacionadas, essas espécies apresentam diferenças

relevantes quanto ao grau de diferenciação genética entre as populações (MORIYAMA; POWELL, 1996), variação nas seqüências de DNA (AQUADRO et al., 1988; MARTIN-CAMPOS et al., 1992) e variação protéica (CHOUDHARY; SINGH, 1987). Outros dois aspectos importantes, e altamente relacionados ao presente estudo, dizem respeito à resposta diferencial dessas espécies sob a seleção a inseticidas (WALLACE, 1986; WINDELSPECHT et al., 1995) e o número e tipos de elementos transponíveis presentes em seus genomas. De acordo Vieira et al. (1999), *D. melanogaster* apresenta o triplo de TEs (15 %) que *D. simulans* (5 %), sendo que as proporções de TEs nas regiões eucromáticas correspondem, respectivamente a 5,35 % e 2,73 % (BERGMAN et al., 2006, CLARCK et al., 2007).

Em face das características particulares de *D. melanogaster* e *D. simulans*, acima mencionadas, neste estudo visamos reavaliar *in silico*, e em estudos populacionais, a questão da inserção preferencial de TEs em genes associados à resistência aos inseticidas, ampliando o número de genes já estudados. As análises *in silico* foram realizadas em 35 genes *Cyps*, dentre eles nove genes associados à resistência, 20 com função monooxigenase ainda não especificada e seis relacionados ao desenvolvimento, estando envolvidos na biossíntese de ecdisona. Adicionalmente, foram analisados os genes flanqueadores desses *Cyps*, tanto em *D. melanogaster*, como em *D. simulans*. As análises populacionais foram realizadas em linhagens de *D. melanogaster* e de *D. simulans*, suscetíveis e resistentes a inseticidas, procedentes de diferentes regiões geográficas.

2.2 OBJETIVOS

Para avaliar a existência de possíveis associações entre a presença de TEs e a resistência aos inseticidas em *D. melanogaster* e *D. simulans*, e comparar as duas espécies quanto a essa característica, os objetivos específicos deste estudo foram os seguintes:

1. Verificar se há inserção preferencial de TEs em regiões intergênicas e gênicas de genes *Cyp6* com funções distintas, comparados aos seus genes flanqueadores nas extremidades 5' e 3'.
2. Relacionar o número de inserções de TEs com o tamanho da região intergênica dos genes estudados;
3. Verificar se os TEs se inserem preferencialmente em regiões promotoras;
4. Analisar, em populações naturais de *D. melanogaster* e *D. simulans*, possíveis polimorfismos de inserção de TEs, nas regiões flanqueadoras 5' e 3' dos genes *Cyp6*, selecionados a partir da análise *in silico*, como potenciais candidatos a apresentarem inserções diferenciais entre linhagens resistentes e suscetíveis a inseticidas nas duas espécies.

2.3 MATERIAL E MÉTODOS

2.3.1 Análises *in silico*

Neste trabalho foram estudados 35 genes *Cyps* e seus genes flanqueadores, nos genomas de *D. melanogaster* e *D. simulans*. Estes genes foram selecionados após uma busca na literatura por genes *Cyps* associados à resistência aos inseticidas, e também por outros *Cyps* que se localizam próximos aos primeiros, em *Drosophila* e outros insetos. Para a análise, foram utilizados o genoma seqüenciado de *D. melanogaster* (versão 5.5: ftp://ftp.flybase.net/releases/FB2008_01/dmel_r5.5/fasta/) e o *assembly* mosaico, correspondente a diferentes linhagens geográficas de *D. simulans*, cujas seqüências dos braços cromossomais 2L, 2R, 3L, 3R, 4, X e a parte não agrupada (chamada de U) correspondem a primeira versão do genoma de *D. simulans*, disponível no sítio *ftp* do *Genome Sequencing Center* da Washington University Medical School (ftp://ftp.flybase.net/releases/FB2008_01/dsim_r1.0/fasta/).

Modelos gênicos estão disponíveis tanto no genoma de *D. melanogaster* (13.733 genes) quanto em *D. simulans* (15.983) entretanto, apenas o primeiro apresenta seus genes extensivamente anotados e depositados no *LocusLink* (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi.db=gene>). Assim, a identificação dos genes da família *Cyp* nos dois genomas foi obtida por meio das coordenadas genômicas em *D. melanogaster* e por meio de comparações por BLASTN (ALTSCHUL et al. 1990) entre as seqüências completas dos genes *Cyps* de *D. melanogaster* contra *D. simulans*, com os parâmetros com $E = e-0.5$, $V = 10000$, $B = 10000$. Após levantamento da localização genômica de cada gene *Cyp*, foram extraídas suas seqüências e de seus genes flanqueadores, além das regiões intergênicas 5' e 3'. Essas seqüências foram utilizadas como fonte de dados para avaliação da ocorrência de fragmentos de elementos de transposição nas regiões intergênicas 5' e 3' e para uma análise comparativa dos resultados obtidos para os genes *Cyps*, e seus flanqueadores. A contagem das inserções dos genes flanqueadores foi realizada na região intergênica não comum ao *Cyp*. Como controle, foi realizada uma busca por inserções de TEs em uma amostra de 70 genes não *Cyps* selecionados aleatoriamente nos genomas de ambas as espécies.

A identificação de elementos de transposição nesses genes, e suas respectivas regiões intergênicas 5' e 3', foi realizada pelo programa *RepeatMasker* (RM) versão 3.1.9 (<http://www.repeatmasker.org>), com a ferramenta de alinhamento *Cross_match* versão 0.990329 (pacote Phrap/cross_match/swat: <http://www.phrap.org/phredphrapconsed.html>) contra o banco de seqüências de TEs de referência do gênero *Drosophila* oriundas do *Repbase* (JURKA et al., 2005). Para eliminar resultados espúrios, somente escores de *matches* RM > 225 foram aceitos, condições de procura de maior sensibilidade/menor velocidade (condição “-s”), matriz relativa ao conteúdo GC do *query* (“-gccalc”) e seqüências de DNA de baixa complexidade não foram mascaradas (“-nolow”). Foram considerados todos os alinhamentos registrados pelo *RepeatMasker* para análises posteriores sem imposição de escores adicionais e comprimento mínimo. Sequências do mesmo elemento que estavam separadas em até 200 pb, foram consideradas como sendo apenas uma cópia. A avaliação da significância das diferenças entre número de inserções entre as regiões 5' e 3', foi realizada por meio do teste do χ^2 .

Foi ainda realizada uma predição de putativos sítios de ligação de fatores de transcrição, nas cópias do elemento DNAREP1_DM, utilizando a ferramenta *Alibaba2* (GRABE, 2002). Este programa consiste em alinhar a seqüência de entrada com uma seqüência já conhecida do banco de dados. Para o parâmetro de conservação da matriz foi utilizado o padrão do programa (75 % de conservação) e para o parâmetro de similaridade da seqüência analisada com a matriz foi utilizada a opção 100 %, o que significa que a seqüência de entrada apresenta praticamente os mesmos nucleotídeos da seqüência da matriz.

2.3.2 Análises populacionais

2.3.2.1 Espécies

Foram analisadas nove linhagens de *D. melanogaster* e 10 de *D. simulans*, classificadas como resistentes ou suscetíveis ao DDT por meio de bioensaios realizados por Granzotto 2007 (Tabela 1).

2.3.2.2 Amostras de DNA e reação de PCR

O DNA genômico foi extraído de 30 indivíduos de cada linhagem (DANIELS; STRAUSBAUGH, 1986, com modificações). Foram desenhados oligonucleotídeos iniciadores que amplificam fragmentos das regiões flangeadoras 5' ou 3' dos genes *Cyps* selecionados a partir da análise *in silico* (Tabela 2). As reações de PCR foram ajustadas para cada seqüência a partir das seguintes condições padrões: 200 ng de DNA genômico, 0.4 mM de cada dNTP, 2 mM de MgCl₂, 0.4 µM de cada iniciador, 1 U de Taq Platinum DNA polimerase (Invitrogen) em tampão 1X, completando-se com água ultrapura até o volume final de 25 µL. A ciclagem constou de 3 minutos a 94 °C para desnaturação inicial; e 30 ciclos de 1 minuto a 94 °C, 1 minuto a 65 °C, 3 minutos a 72 °C, finalizando com um ciclo de 10 minutos a 72 °C para extensão final. Como controle negativo foi utilizada água ultrapura. Os fragmentos amplificados foram separados por eletroforese (90 volts durante 35 minutos) em gel de agarose 1 % preparado com tampão TAE 0,5x e, em seguida, corado com brometo de etídeo e visualizado sob luz ultravioleta.

Tabela 1. Espécies e linhagens analisadas, características quanto à resistência (R) ou suscetibilidade (S), procedências geográficas e ano da coleta.

Espécies	Linhagens	R/S	Procedência Geográfica	Coletor/Ano de Coleta
<i>D. melanogaster</i>	SE	S	Aracaju/SE	Taline B. Pínula/2004
	LE	R	Lençóis/BA	Maura H. Manfrin/2005
	MU	R	Mucuri/BA	Maura H. Manfrin/2005
	ITA	R	Itaúnas/ES	Maura H. Manfrin/2005
	RP	R	São José do Rio Preto/SP	Lilian Madi-Ravazzi/2005
	BG	R	Bento Gonçalves/RS	Gladis F. da Cunha/2006
	SM	R	Santa Maria/RS	Paulo R. P. Hoffman /2005
	TAV	R	Tavares/RS	Paulo R. P. Hoffman/2005
	Canton-S	S	Laboratório Padrão	-
<i>D. simulans</i>	LE	S	Lençóis/BA	Maura H. Manfrin/2005
	MU	S	Mucuri/BA	Maura H. Manfrin /2005
	ITA	S	Itaúnas/ES	Maura H. Manfrin /2005
	PF	R	Paulo de Faria/SP	Leiza Penariol /2005
	OV	R	Onda Verde/SP	Adriana Granzotto/2005
	RAT	R	Ratones/SC	Paulo R. P. Hoffman /2005
	BG	R	Bento Gonçalves/RS	Gladis F. da Cunha /2006
	AMI	S	Amieu/Nova Caledonia	Cristina Vieira/2005
	MAK	S	Makindu/África	Cristina Vieira/2005
	MAD	S	Madagascar	Cristina Vieira/2005

Tabela 2 Seqüências dos oligonucleotídeos iniciadores e tamanho esperado dos fragmentos (pb) amplificados de genes da família *Cyp* relacionados à resistência aos inseticidas.

Genes	Primer Forward	Primer Reverse	pb
<i>Cyp6w1_5'_Dm</i>	5' TCGGTGTGCCATCCAGGTTATACT 3'	5' TGCTACATTAGGGACACCCAACCT 3'	670
<i>Cyp6w1_5'_Ds</i>	5' TCGGTGGCACTCCAGGTTATAGTT 3'	5' TTTGGGCGTCTTTGGAATCTGC 3'	1024
<i>Cyp6w1_3'_Ds</i>	5' AGATGTGGCACCGATTTGGTTGTC 3'	5' GCATTCAACATGACACGGCCAAGA 3'	923
<i>Cyp6a2_5'_Dm</i>	5' CATTGAAAGCAATGTCGCGTAGT 3'	5' TCTTCTCGTCGGGCGCTTCATTAT 3'	896
<i>Cyp6a2_5'_Ds</i>	5' CCAAACCTGGTTTGTGCGCTTATCTG 3'	5' CCCGGCACAAATGCACATCCTAAT 3'	1035
<i>Cyp6a2_3'_Ds</i>	5' TCCAATTTACGCGTGACAAGCCAG 3'	5' TTTCGCAAATGTCAATGCCTCCGC 3'	834
<i>Cyp4e2_5'_Dm</i>	5' CGGAAGAACAACGAACAGAATAATGCCG 3'	5' TTCTGTTCCGCTCCGTCACGTTT 3'	427
<i>Cyp12d1_5'_Dm</i>	5' AATCGCCGTGAATTGCCTACCTCA 3'	5' AGATGCTAGAACTCTGGAGATGCC 3'	636
<i>Cyp12a4_3'_Dm</i>	5' GATCAACTTGCCAAACATTCCGCTC 3'	5' GGACTGGCTCAATGGTATCACAAC 3'	1852

2.3.2.3 Clonagem e seqüenciamento

A maioria dos fragmentos com polimorfismo de tamanho entre as linhagens geográficas suscetíveis e resistentes de *D. melanogaster* e *D. simulans*, obtidos a partir da PCR foram diretamente purificados usando o Kit *GFXTM PCR DNA and Gel Band Purification* (GE Healthcare), e seqüenciados. Outros fragmentos foram extraídos do gel de agarose, purificados e, posteriormente, clonados com o kit *TOPO[®] XL Cloning* (Invitrogen), segundo as especificações do fabricante, e seqüenciados.

2.3.2.4 Análise das seqüências

As seqüências obtidas foram analisadas com auxílio dos programas *BioEdit* (HALL, 1999), *BLAST* (ALTSCHUL et al., 1990), *BLAST 2 sequences* (TATUSOVA; MADDEN, 1999) e alinhadas com o *ClustalW* (THOMPSON et al., 1994). O programa *RepeatMasker* foi utilizado para identificar elementos transponíveis nas seqüências obtidas, visando detectarem-se inserções de TEs em regiões regulatórias dos genes.

Para análise do polimorfismo dos TEs inseridos mais frequentemente nas regiões flangeadoras seqüenciadas dos genes *Cyps*, foi estimada a diversidade nucleotídica (π) (NEI; KI, 1979), como implementado pelo programa DnaSP versão 4.10.9. (ROZAS et al., 2003).

2.4 RESULTADOS

A família multigênica Citocromo P450 monooxigenases possui 90 genes no genoma de *D. melanogaster*, os quais desempenham diferentes funções biológicas (DABORN et al., 2007; McCART; FFRENCH-CONSTANT, 2008). Segundo Chen e Li (2007), dentre os *Cyps* com funções já determinadas, seis estão envolvidos na síntese de hormônios relacionados ao desenvolvimento (*Cyp302a1*, *Cyp306a1*, *Cyp307a1*, *Cyp314a1*, *Cyp315a1* e também o *Cyp18a1*) e nove estão associados ao metabolismo de xenobióticos, tais como os inseticidas (*Cyp4e2*, *Cyp6a2*, *Cyp6a8*, *Cyp6a9*, *Cyp6g1*, *Cyp6w1*, *Cyp12a4*, *Cyp12d1* e também o *Cyp308a1*).

2.4.1 Inserções de TEs em genes *Cyps* com funções distintas

Neste trabalho foram analisados *in silico* 35 genes *Cyps* (incluindo os 15 acima citados) e seus genes flanqueadores nas extremidades 5' e 3', para a detecção de inserções de elementos transponíveis (TEs) no genoma de *D. melanogaster* e *D. simulans*. Para verificar se havia inserção preferencial de TEs em genes *Cyps* com diferentes funções e em seus flanqueadores, os genes *Cyps* foram agrupados em três categorias funcionais, de acordo com sua função dada pela nomenclatura *The Gene Ontology* (2001): 1) nove genes relacionados à resistência aos inseticidas; 2) 20 genes relacionados à atividade de monooxigenase (detoxificação) em geral, cujas funções específicas não foram ainda definidas e, por isso, passam aqui a serem chamados simplesmente de atividade monooxigenase e 3) seis genes relacionados ao desenvolvimento. Os genes flanqueadores foram agrupados de acordo com a função do seu *Cyp* vizinho. As inserções presentes em cada gene são apresentadas nos Apêndices A a D, da seguinte maneira: genes *Cyps* (Apêndice A) e genes flanqueadores (Apêndice B) de *D. melanogaster* e genes *Cyps* e flanqueadores de *D. simulans*, respectivamente nos Apêndices C e D. Essas inserções pertencem à diferentes classes e ordens de elementos de transposição.

O número total de inserções, bem como seus tamanhos médios e divergências médias (em relação ao elemento canônico) são apresentados na Tabela 3. Em *D. melanogaster*, dentre os

genes *Cyps*, o maior número médio de inserções por gene ocorreu nos *Cyps* associados à resistência (1,67), enquanto que o menor foi observado nos *Cyps* associados ao desenvolvimento (0,50). Nos genes flanqueadores dos *Cyps*, tanto a maior média (2,0), quanto a menor (0,33), também foram observadas naqueles associados à resistência e ao desenvolvimento, respectivamente. Para o tamanho, a maior média das inserções nos *Cyps* ocorreu nos genes associados à atividade monooxigenase (1.619 pb) e, a menor, nos genes associados ao desenvolvimento (77 pb). Para os flanqueadores dos *Cyps*, a maior média foi encontrada naqueles associados à resistência (417 pb) e, a menor, nos associados à atividade monooxigenase (75 pb). A divergência das inserções encontradas nos genes *Cyps* variou de 22,75 %, em média, nos *Cyps* de desenvolvimento, a 3,50 %, nos *Cyps* associados à atividade monooxigenase. Nos genes flanqueadores, a divergência variou de 19,86 %, também naqueles associados aos *Cyps* de desenvolvimento, a 14,06 %, nos flanqueadores dos *Cyps* associados à resistência.

Para *D. simulans* (Tabela 3), o número médio de inserções nos genes *Cyps* variou de 3,33 a 0,35, respectivamente nos genes associados à resistência e à atividade monooxigenase. Nos genes flanqueadores dos *Cyps*, a maior média foi 1,12 inserções (flanqueadores dos *Cyps* associados à atividade monooxigenase) e, a menor, foi 0,17 inserções (flanqueadores dos *Cyps* associados à resistência). Para o tamanho, a maior média das inserções nos *Cyps* ocorreu nos genes associados à atividade monooxigenase (314 pb) e, a menor, nos genes associados à resistência (159 pb). Para os flanqueadores dos *Cyps*, a maior média foi encontrada naqueles associados à atividade monooxigenase (238 pb) e, a menor, nos associados à resistência (85 pb). A divergência média das inserções encontradas nos genes *Cyps* variou de 16,72 %, nos associados à resistência, a 10,07 %, nos *Cyps* associados à atividade monooxigenase. Da mesma, a maior e menor divergência média ocorreu nos flanqueadores desses *Cyps*, 20,51 % (dos *Cyps*) e 15,60 % (atividade monooxigenase).

As porcentagens de TEs inseridos em genes e regiões intergênicas foram calculadas para os genes de cada categoria funcional, em relação ao total de TEs nos genes analisados em *D. melanogaster* (29 nos *Cyps* e 38 em seus flanqueadores) e *D. simulans* (40 nos *Cyps* e 23 em seus flanqueadores), respectivamente (Figura 1). Os *Cyps* envolvidos na resistência aos inseticidas apresentaram, em *D. melanogaster*, proporcionalmente maior número de inserções de TEs (15 = 51,72 %) que aqueles associados à atividade monooxigenase (11 = 37,93 %) e ao desenvolvimento (3 = 10,34 %), o mesmo ocorrendo para seus respectivos genes flanqueadores

(resistência: 57,89 %; atividade monooxigenase: 34,21 % e desenvolvimento: 7,89 %) (Figura 1 a). A mesma relação foi observada em *D. simulans*; os genes *Cyps* associados à resistência apresentaram maior proporção de TEs (30 = 75,00 %) que os associados à atividade monooxigenase (7 = 17,50 %) e ao desenvolvimento (3 = 7,50 %). Porém, para os genes flanqueadores, os associados aos *Cyps* com atividade monooxigenase apresentaram maior porcentagem de inserções (82,6 %), enquanto para as demais categorias funcionais, a proporção foi a mesma (8,70 %) (Figura 1 b). As diferenças entre as proporções de inserções de TEs nos genes *Cyps* e seus respectivos flanqueadores não foram significantes em ambas as espécies (*D. melanogaster*, χ^2 : 0,77; $p > 0,05$; *D. simulans*, χ^2 : 3,37; $p > 0,05$).

Tabela 3. Número total de genes analisados em cada categoria funcional (N), valores médios, mínimos (<) e máximos (>) do número, do tamanho e da divergência de elementos de transposição em genes *Cyps* e seus genes flanqueadores.

Genes	<i>D. melanogaster</i>				<i>D. simulans</i>			
	N	número (< >)	tamanho (< >)	divergência (%) (< >)	N	número (< >)	tamanho (< >)	divergência (%) (< >)
<i>Cyps</i> de resistência	9	1,67 (1; 7)	724 (31; 7914)	9,72 (0; 26,67)	9	3,33 (2; 22)	159 (36; 392)	16,72 (7,32; 25,86)
Flanqueadores <i>Cyps</i> de resistência	11	2 (1; 23)	417 (33; 5126)	14,06 (0,14; 24,14)	12	0,17 (2)	85 (74; 95)	20,51 (18,92; 22,11)
<i>Cyps</i> monooxigenase	20	0,55 (1; 3)	1619 (31; 7914)	3,50 (0; 13,63)	20	0,35 (1; 3)	314 (34; 1728)	10,07 (0,23; 26,15)
Flanqueadores aos <i>Cyps</i> monooxigenase	16	0,81 (1; 7)	75 (32; 142)	16,98 (0; 26,27)	17	1,12 (1; 10)	238 (35; 1728)	15,60 (0,23; 24,99)
<i>Cyps</i> de desenvolvimento	6	0,50 (3)	77 (80; 84)	22,75 (18,75; 28,92)	6	0,50 (1; 2)	198 (60; 385)	12,43 (2,64; 25,00)
Flanqueadores <i>Cyps</i> de desenvolvimento	9	0,33 (1; 3)	134 (88; 189)	19,86 (12,20; 31,09)	9	0,22 (2)	107 (87; 127)	19,75 (12,79; 26,72)

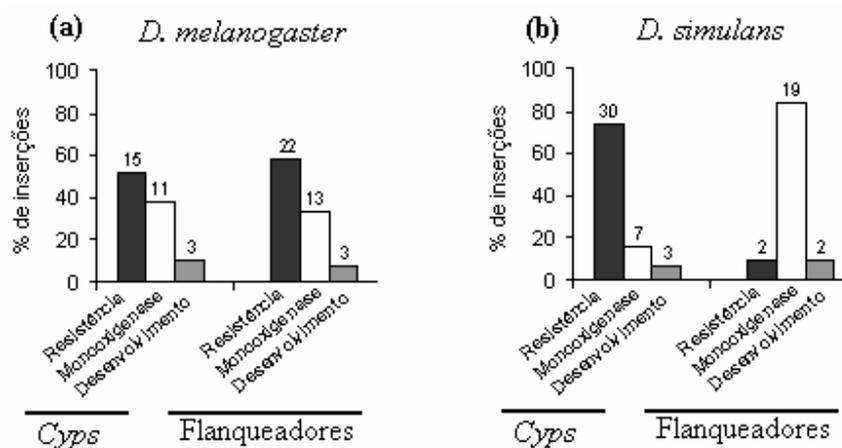


Figura 1. Porcentagem de inserções de TEs (ordenada) em diferentes categorias funcionais de genes *Cyp* e de seus flaqueadores (abscissa), em relação ao total de TEs inseridos em *D. melanogaster* (a) e *D. simulans* (b). Genes *Cyp* e seus flaqueadores associados à resistência aos inseticidas (preto), genes *Cyp* e seus flaqueadores associados à atividade de monooxigenase (branco) e genes *Cyp* e seus flaqueadores associados ao desenvolvimento (cinza). Números de inserções indicados acima de cada coluna.

2.4.2 Distribuição de TEs em regiões intergênicas e gênicas dos genes *Cyps* e seus flaqueadores

Foi realizada uma busca por inserções de TEs tanto nas regiões gênicas dos *Cyps* e seus flaqueadores quanto nas regiões intergênicas inteiras, independente do seu tamanho, para analisar a ocorrência de distribuição diferencial entre essas regiões. As inserções estavam principalmente localizadas em regiões intergênicas quando comparadas às regiões codificantes, ocorrendo em maior número nas extremidades 5' do que nas 3' (Apêndices A a D), tanto para *D. melanogaster* (Figura 2 a), como para *D. simulans* (Figura 2 b). As diferenças observadas entre as duas regiões foram significantes para os genes *Cyps* (χ^2 : 16,60; $p < 0,01$), mas não para seus flaqueadores (χ^2 : 1,80; $p < 0,05$) em *D. melanogaster*, entretanto, foram significantes tanto para os *Cyps* (χ^2 : 7,60; $p < 0,01$) quanto para seus flaqueadores (χ^2 : 13,19; $p < 0,01$) em *D. simulans*. Porém, quando considerados os números de inserções nas regiões 5' e 3' de genes de cada categoria funcional (Apêndices A a D), as diferenças foram significantes para os *Cyps* associados à resistência (χ^2 : 10,42; $p < 0,01$) e para os *Cyps* associados à atividade monooxigenase (χ^2 : 19,60; $p < 0,01$) e seus flaqueadores (χ^2 : 7,01; $p < 0,01$) em *D. melanogaster*. Em *D. simulans*, as diferenças foram significantes para os *Cyps* associados à resistência (χ^2 : 13,69; $p < 0,01$) e à atividade monooxigenase (χ^2 : 8,76; $p < 0,01$). Entre os genes flaqueadores houve diferença significativa apenas para aqueles associados aos *Cyps* com atividade monooxigenase (χ^2 : 22,05; $p < 0,01$).

Todas as inserções detectadas nos 35 genes *Cyps* localizavam-se nas regiões intergênicas de *D. melanogaster* (29) e *D. simulans* (40), como pode ser observado nos Apêndices A e C. Para os genes flaqueadores dos *Cyps* de *D. simulans*, o resultado foi muito semelhante, pois, foram detectadas apenas duas inserções (8,69 % do total de 23) em introns, no gene *GD12614* (FBgn0184341), como mostrado no Apêndice D. Por outro lado, os genes flaqueadores dos *Cyps* de *D. melanogaster* apresentaram inserções de TEs não apenas nas regiões intergênicas (15), mas também, em introns (19), éxons (1) e UTR (3). Quando se compara o número de inserções de TEs entre introns e éxons dos genes flaqueadores dos *Cyps* de *D. melanogaster*, nota-se que os primeiros apresentaram 19 inserções (50 % do total de 38), enquanto que os últimos apresentaram apenas uma inserção (2,63 % do total de 38), como mostrado no Apêndice B.

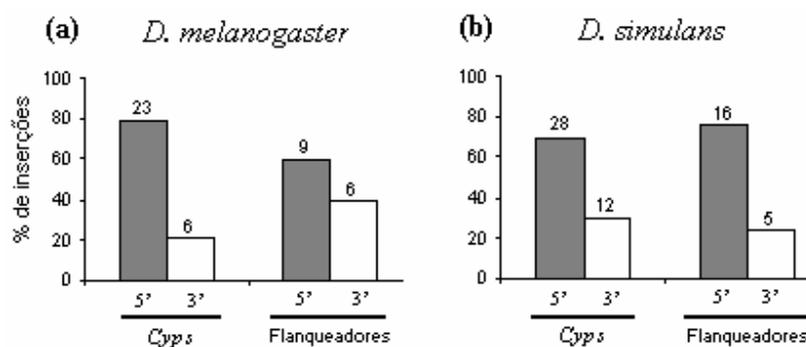


Figura 2. Porcentagem de inserções de TEs (ordenada) em relação ao total de inserções nas regiões intergênicas 5' (cinza escuro) ou 3' (branco) em genes *Cyp5* e em seus flaqueadores (abscissa) em *D. melanogaster* (a) e *D. simulans* (b). Números de inserções indicados acima de cada coluna.

2.4.3 Ocorrência de inserções de TEs em regiões flanqueadoras distantes em até 3 kb do gene

Como a presença de sinais reguladores nas seqüências de TEs pode afetar a transcrição dos genes vizinhos e, conseqüentemente, alterar seu padrão de expressão, neste trabalho, foi realizada uma busca por inserções de TEs em regiões flanqueadoras 5' distantes em até 3 kb do início da seqüência gênica, onde se localizam a maior parte dos promotores (Figura 3 e Apêndices E e F), como também distantes em até 3 kb do término do gene (3').

A Figura 3 mostra que as proporções de inserções de TEs nas regiões flanqueadoras em até 3 kb dos genes analisados em *D. melanogaster*, ocorreram preferencialmente nas extremidades 5' dos genes *Cyps* (71,43%; χ^2 : 8,76; $p < 0,01$), o mesmo não ocorrendo para os genes resultantes da amostragem aleatória do genoma (63,64 %; χ^2 : 3,44; $p > 0,05$). Os genes flanqueadores dos *Cyps* de *D. melanogaster* apresentaram apenas uma inserção na região 3', próxima em até 3 kb do término do gene. Por outro lado, em *D. simulans*, as proporções de TEs nas extremidades 5' e 3' dos genes *Cyps* (53,33 % vs 46,67 %) e dos genes amostrados aleatoriamente (57,41 % vs 42,59 %) não diferiram significativamente (*Cyps*: χ^2 : 0,16; $p > 0,05$; amostra aleatória: χ^2 : 0,95; $p > 0,05$). Os genes flanqueadores dos *Cyps* de *D. simulans*, ao contrário, apresentaram uma inserção na região 5' (25 %) e três na região 3' (75 %), sendo a diferença significativa (χ^2 : 12,00; $p < 0,01$).

Outra abordagem foi a análise das proporções de inserções de TEs em regiões flanqueadoras 5' e 3' em até 3 kb do gene, em intervalos de 0,5 kb (0 a 0,5 kb e 0,5 a 1 kb) e de 1 kb (1 a 2 kb e 2 a 3 kb), de modo a se avaliar a ocorrência de inserções preferenciais de TEs em alguns desses intervalos (Figura 4).

Em *D. melanogaster* (Apêndice E), as regiões 5' dos *Cyps* (Figura 4 a) associados à resistência aos inseticidas apresentaram inserções em todos os intervalos de distância, com proporções variando de 20 a 40 %, sendo que a região de 0,5 a 1 kb foi a que apresentou a maior proporção de inserções (40 %). De modo semelhante, as proporções de inserções presentes nos *Cyps* relacionados à atividade monooxigenase (Figura 4 a) variaram de 20 a 40 %, não havendo inserções no intervalo de maior distância (2 a 3 kb). As regiões distantes de 0 a 0,5 kb e de 1 a 2 kb apresentaram maior proporção de inserção (40 %). Por outro lado, tanto os genes *Cyps* associados ao desenvolvimento (Figura 4 c), quanto os genes flanqueadores aos *Cyps* das três

categorias funcionais (Figura 4 e), não apresentaram inserções de TEs nesta região. Nas regiões 3', os genes *Cyps* que apresentaram inserções foram apenas aqueles associados à resistência (Figura 4 b), sendo estas distribuídas em distâncias de 0 a 2 kb. Quanto aos genes flanqueadores, foi encontrada apenas uma única inserção na região 3' de um gene flanqueador a um *Cyp* associado ao desenvolvimento (Figura 4 d).

Em *D. simulans*, as inserções presentes na região 5' dos genes *Cyps* associados à resistência estavam distribuídas nas distâncias de 0 a 2 kb, sendo o maior valor encontrado no intervalo de 1 a 2 kb (71,42 %), enquanto que os *Cyps* associados à atividade monooxigenase apresentaram inserções apenas no intervalo de 1 a 2 kb de distância (Figura 5 a). Entre os genes flanqueadores, houve apenas uma inserção na região 5' de um gene flanqueador de um *Cyp* associado à atividade monooxigenase, sendo esta localizada na região mais próxima ao gene (0 a 0,5 kb) (Figura 5 d). Quando analisadas as regiões 3', nota-se a ocorrência de inserções apenas em *Cyps* associados à resistência (Figura 5 b), e que as inserções estavam distribuídas em dois intervalos de distância, 0 a 0,5 kb (57,14 %) e 1 a 2 kb (42,85 %). Ainda, para a região 3', foram encontradas inserções apenas nos genes flanqueadores dos genes *Cyps* associados à atividade monooxigenase (Figura 5 e), sendo estas localizadas nos intervalos de menor e maior distâncias, 66,67 % e 33,33 %, respectivamente. A heterogeneidade de distribuição e a variação das proporções de inserções dos TEs nos diferentes intervalos de distância do gene sugerem não haver associação entre distância em relação ao gene e o número de inserções, nas regiões próximas aos genes.

Foram também comparadas as proporções de inserções de TEs nas regiões próximas em até 3 kb dos genes de cada categoria funcional dos genes *Cyps*. Em *D. melanogaster*, o número médio de inserções por *Cyp* em cada categoria funcional foi 2 (resistência: 10/5 genes), 1,33 (atividade monooxigenase: 4/3 genes) e zero (desenvolvimento). Em *D. simulans*, as médias de inserções foram 4,66 (resistência: 14/3 genes), 1 (atividade monooxigenase: 1/1gene) e zero (desenvolvimento). Para as duas espécies, ocorre, portanto, maior número de inserções de TEs próximos a genes *Cyps* que nas demais categorias.

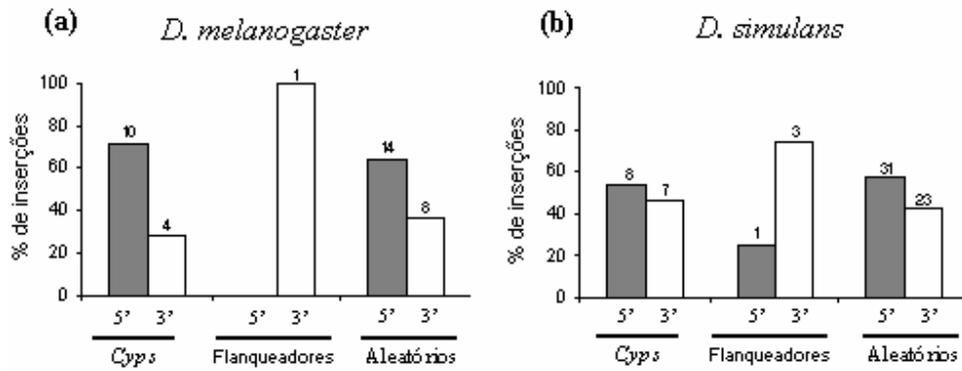


Figura 3. Porcentagem de inserções de TEs (ordenada) nas regiões flanqueadoras 5' (cinza) e região 3' (branco), distantes até 3 kb (abscissa) do gene. (a) *D. melanogaster* (b) *D. simulans*. Números de inserções indicados acima de cada coluna.

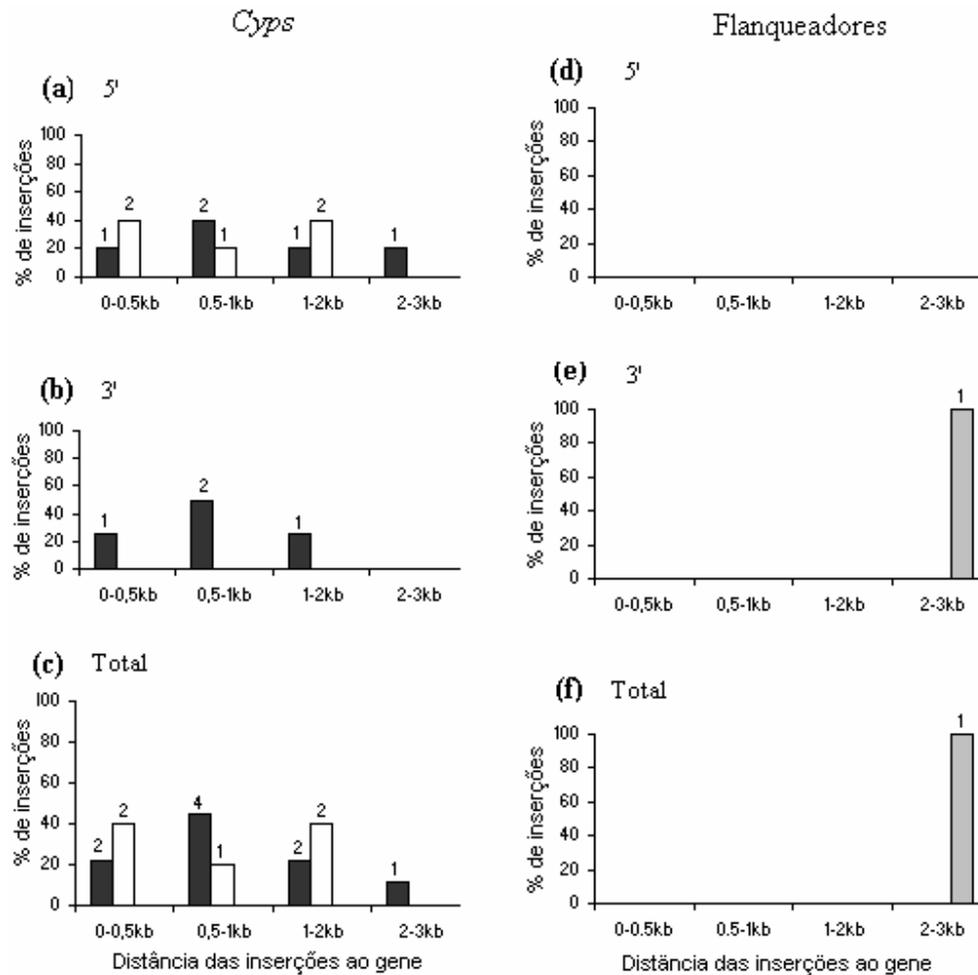


Figura 4. Porcentagem de inserções de TEs em relação a distância em que estão inseridas em regiões intergênicas distantes em até 3 kb do gene alvo em *D. melanogaster*: região 5' dos genes *Cyps* (a) e seus flaqueadores (b); região 3' dos genes *Cyps* (c) e seus flaqueadores (d); Total de inserções dos genes *Cyps* (e) e seus flaqueadores (f). Genes *Cyps* associados à resistência e seus flaqueadores (preto), genes *Cyps* e seus flaqueadores associados à atividade de monooxigenase (branco) e genes *Cyps* e seus flaqueadores associados ao desenvolvimento (cinza). Números de inserções indicados acima de cada coluna.

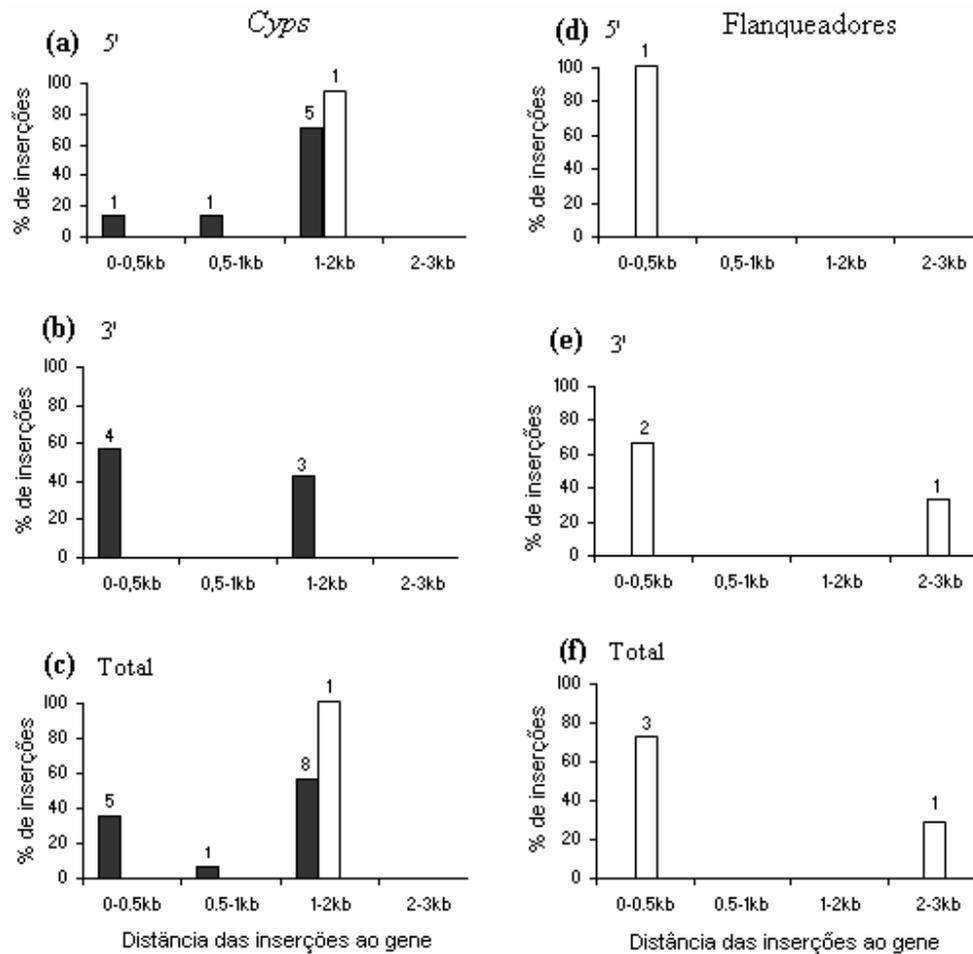


Figura 5. Porcentagem de inserções de TEs em relação a distância em que estão inseridas em regiões intergênicas distantes em até 3 kb do gene alvo em *D. simulans*: região 5' dos genes *Cyps* (a) e seus flanqueadores (b); região 3' dos genes *Cyps* (c) e seus flanqueadores (d); Total de inserções dos genes *Cyps* (e) e seus flanqueadores (f). Genes *Cyps* associados à resistência e seus flanqueadores (preto), genes *Cyps* e seus flanqueadores associados à atividade de monooxigenase (branco) e genes *Cyps* e seus flanqueadores associados ao desenvolvimento (cinza). Números de inserções indicados acima de cada coluna.

2.4.4 Distribuição das classes e ordens de TEs nas regiões intergênicas

Para avaliar se ocorre distribuição diferencial de TEs de diferentes classes e ordens, as inserções foram agrupadas em transposons de DNA, retrotransposons com LTR e LINEs. Do total de inserções nos *Cyps* de *D. melanogaster* (Apêndice G; Figura 6 a, b, c), as pertencentes aos elementos de DNA (DNAREP1_DM, PROTOP, PROTOP_A, PROTOP_B, ISFUN1, POGON1, BARI_DM, hAT-1_DP) foram as mais freqüentes (73,92 %: 17/23), sendo o de DNAREP1_DM o mais freqüente desta classe (41,17 %: 7/17) e do total geral de inserções (30,43 %: 7/23). Tanto os retrotransposons com LTRs (MAX_LTR e MAX_I), quanto os LINEs (TART_DV e LINEJ1_DM), representam 13,05 % do total de 23 inserções. Nos genes flaqueadores, as inserções presentes pertenciam a elementos de DNA (DNAREP1_DM e PROTOP_B), retrotransposons com LTRs (ROO_I) e elementos LINEs (TART_DV). Os genes flaqueadores apresentaram maior porcentagem de inserções de elementos de DNA (60 %: 9/15), seguido pelos elementos LINEs (33,33 %: 5/15), enquanto que os retrotransposons com LTRs apresentaram a menor porcentagem (6,67 %: 1/15). DNAREP1_DM foi o elemento mais freqüente (53,33 %: 8/15), seguido pelo TART_DV (33,33 %: 5/15) e os elementos PROTOP_B e ROO_I (6,67 %: 1/15), com apenas uma inserção cada.

Em *D. simulans* (Apêndice H, Figura 6 d, e, f), as inserções dos genes *Cyps* foram também preponderantemente (95 %: 38/40) de elementos de DNA (DNAREP1_DM, Helitron-1_Dyak, Helitron-1_Dvir, BARI_DM, PROTOP_B, FB4_DM, TransibN1_DP, BARI1 e HOBO), representados principalmente pelo DNAREP1_DM (39,47 %: 15/38) e o Helitron-1_Dyak (34,21 %: 13/38). Apenas duas inserções não se classificaram como transposon de DNA: o retroposon (MINIME_DN) e um elemento não classificado DMRP1. Os genes flaqueadores apresentaram menor número de superfamílias de elementos de DNA, quando comparados aos *Cyps*, pois os TEs Helitron-1_Dyak, PROTOP_B, TransibN1_DP e HOBO não foram encontrados, mas ainda assim foi a classe mais freqüente (47,61 %: 10/21). Por outro lado, apresentaram algumas inserções adicionais de LTRs (Stalker2_LTR e ROO_I) e LINEs (TART_DV). O elemento com o maior número de cópias foi o TART_DV (42,86 %: 9/21), seguido pelo Helitron-1_Dyak (19,05 %: 4/21), DNAREP1_DM e BARI_DM (9,52 %: 2/21) e, por último, os elementos BARI1, FB4_DM, ROO_I e Stalker2_LTR (4,76 %: 1/21).

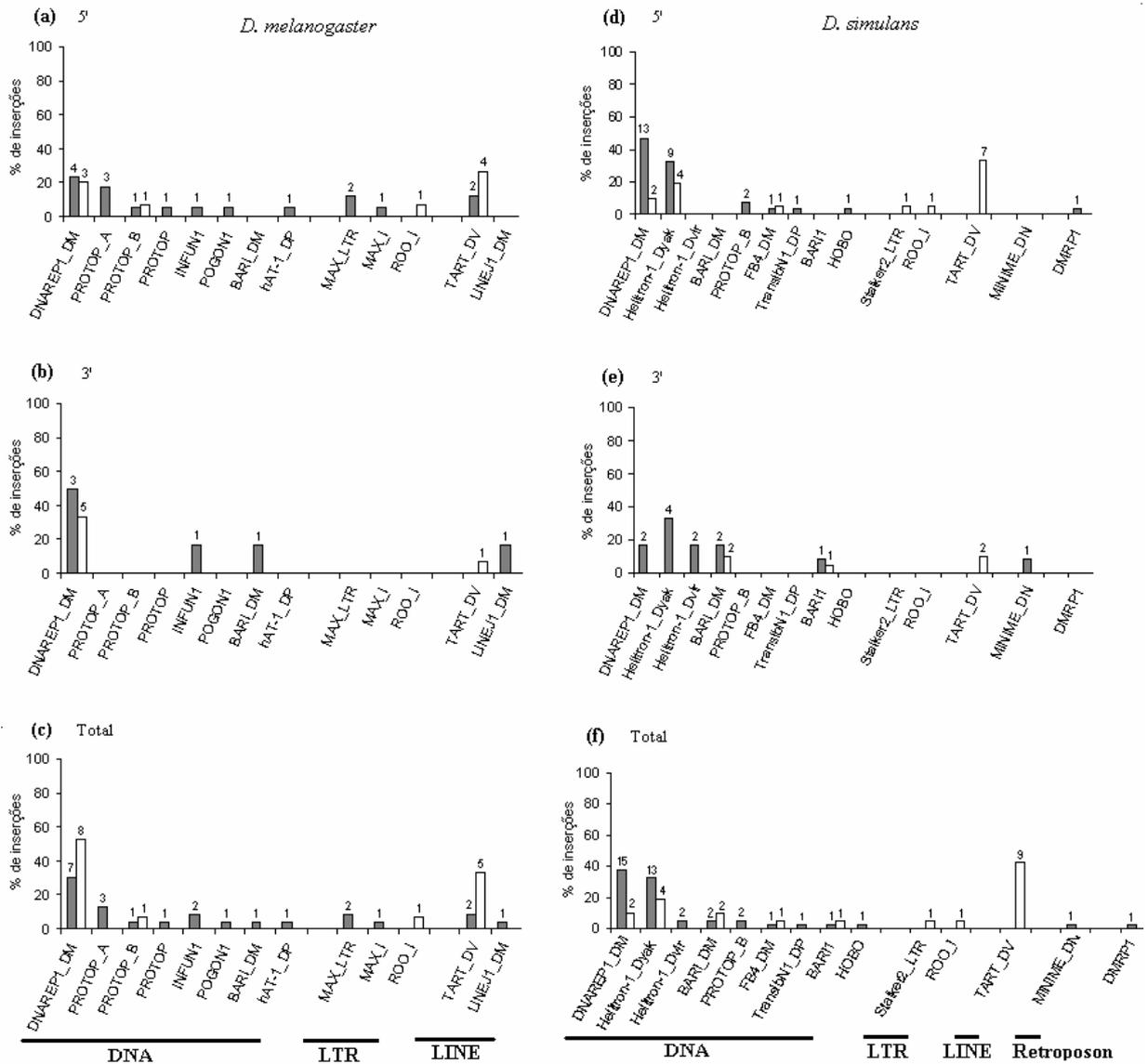


Figura 6. Porcentagem de inserções de TEs em regiões intergênicas. Em *D. melanogaster*: (a) região 5' (b) região 3' (c) Totais de inserções. Em *D. simulans*: (d) região 5' (e) região 3' (f) Totais de inserções. Genes *Cyp*s são representados em cinza e genes flanqueadores em branco. Números de inserções indicados acima de cada coluna; classe ou ordem a que cada TE pertence indicadas abaixo do eixo x.

2.4.5 Ocorrência do elemento DNAREP1_DM em genes *Cyps* e seus flaqueadores em *D. melanogaster* e *D. simulans*

O elemento transponível DNAREP1_DM (*Drosophila interspersed element 1*) ou também chamado DINE-1 ou INE-1, descrito por Kapitonov e Jurka (1999) como um transposon de DNA não autônomo de 594 pb, é o TE mais abundante do genoma de *D. melanogaster*. Segundo Kapitonov e Jurka (2003), o DNAREP1_DM é um elemento ancestral que possivelmente sofreu uma explosão transposicional há mais de três milhões de anos e, a partir daí, permaneceu inativo. Diversos estudos têm mostrado a alta abundância deste elemento no genoma de *D. melanogaster*. Yang e Barbash (2008), analisando os 12 genomas de *Drosophila*, comprovaram que todas as espécies possuem centenas a milhares de cópias de DNAREP1_DM distribuídas no genoma. Em geral, a estrutura do elemento DNAREP1_DM consiste em dois blocos conservados (A e B), separados por uma região de repetições centrais variáveis (Figura 7). O bloco A é subdividido em região A1 e A2, separadas por uma região de repetição microssatélite. O bloco A1 é composto por uma repetição terminal invertida 5' (*subTIR* 5'), outra repetição parcial (*IR*) e uma região conservada (*Core*). O bloco B é formado por uma *subTIR* 3' e por duas *IRs* que formam um *Stem loop* na extremidade 3'.

Cada cópia do DNAREP1_DM inserida nas regiões intergênicas dos genes *Cyps* foi comparada à cópia consenso do DNAREP1_DM (Figura 7), para identificar a qual região do elemento elas correspondiam. Em *D. melanogaster*, de modo geral, tais cópias não correspondem preferencialmente a uma região específica do elemento. Das sete cópias, três (42,85 %) correspondem à extremidade 5' do elemento: uma delas, a cópia *Cyp6w1_5'* inclui a região 5' *subTIR* (*subterminal inverted repeats*), a 5' *IR* (*inverted repeat*) e parte da região *Core*, sendo todas regiões bem conservadas quando comparadas ao elemento consenso; a cópia do *Cyp6a14_3'* também corresponde as regiões 5' *IR* e parte da região *Core*; a cópia do *Cyp6a2_3'* é formada pela região de microssatélite (CCGT)_n, (CTGT)_n e pela região A2. Das quatro cópias restantes, uma cópia do *Cyp6w1_3'* corresponde a uma parte da região A2, a região de repetições centrais (com tamanho e repetições variáveis) e praticamente todo bloco B (3' *subTIR* e 3' *stem-loop*), e outra cópia *Cyp6w1_5'* corresponde à metade da região de repetições centrais até o fim

do bloco B. As outras duas cópias correspondem à parte central das repetições centrais (*Cyp6w1_5'*) e à região 3' *subTIR* e 3' *stem-loop* do bloco B (*Cyp6a22_5'*).

Em *D. simulans*, as 15 cópias do DNAREP1_DM ocupavam as regiões intergênicas de seis genes *Cyps*, três associados à resistência aos inseticidas (*Cyp6w1*, *Cyp6a2* e *Cyp12d1*), dois a atividade de monooxigenase (*Cyp6a14* e *Cyp6a22*) e um ao desenvolvimento (*Cyp315a1*). Nessa espécie, ao contrário do observado em *D. melanogaster*, os resultados mostram que cópias do DNAREP1_DM com diferentes tamanhos correspondem preferencialmente à extremidade 3' do elemento consenso (Figura 8). Oito do total de 15 cópias (53,33 %) de DNAREP1_DM em *D. simulans* estavam associadas ao gene *Cyp6w1*, sete delas na região intergênica 5' e uma na 3'. Este elevado número de cópias na região intergênica 5' do *Cyp6w1* pode ser devido ao grande tamanho desta região, com 21.297 pb. Do total de 15 cópias, apenas três (20 %) correspondem à extremidade 5' do elemento: a cópia do *Cyp6w1_5'*, que inclui a 5' *subTIR*, a 5' *IR*, o *Core*, a região A2 e o início da região de repetições centrais; a cópia do *Cyp6a14_5'*, que se estende da região 5' *IR* até o primeiro nucleotídeo do *Core*; e outra cópia do *Cyp6w1_5'*, representando parte do *Core*, a região de microsatélite total até o início da região A2. Todas as outras 12 cópias (80 %) do DNAREP1_DM correspondem à extremidade 3' do elemento, sendo que oito delas (66,66 %) correspondem exclusivamente à região do bloco B, principalmente das regiões 3' *subTIR* e 3' *stem-loop*. As outras quatro cópias (33,33 %) correspondem à parte do bloco B e parte da região de repetições centrais, sendo que uma delas representa o bloco B todo e quase toda a região de repetições centrais.

Neste estudo, foi possível analisar a porcentagem de identidade e o tamanho das cópias do DNAREP1_DM inseridas nas regiões intergênicas dos genes *Cyps* (Figura 8 a) e de seus genes flaqueadores (Figura 8 b) de *D. melanogaster* e de *D. simulans* (Figura 8 c e 9 d). Dos 35 *Cyps* analisados, apenas quatro (11,42 %) apresentaram inserções de DNAREP1_DM (*Cyp6w1*, *Cyp6a2*, *Cyp6a22* e *Cyp6a14*) em *D. melanogaster* (Figura 8 a). Do total de sete inserções, quatro (57,14 %) estão localizadas no *Cyp6w1* (três na região intergênica 5' e uma na região 3'). As cópias são em geral de pequeno tamanho, com maior distribuição entre 40 e 100 pb, com exceção de duas no *Cyp6w1*, com 220 e 400 pb. A identidade dessas cópias, em relação ao elemento consenso, variou de 73 % a 87 %, com uma média de 82 %. Dentre as sete cópias deste elemento, seis estavam localizadas em até 3 kb do gene.

Por outro lado, os genes flanqueadores desta mesma espécie (Figura 8 b) apresentaram oito cópias de DNAREP1_DM em suas regiões intergênicas. Dentre os 36 genes flanqueadores, três (8,33 %) apresentaram cópias deste TE (CG12780, CG6040 e CG9628). Tais cópias possuíam tamanho médio de 85 pb, quatro delas (50 % do total) inseridas na região 3' do gene CG12780. Essas cópias apresentaram tamanhos que variaram de 37 a 189 pb e a porcentagem de identidade variou de 76 a 92 %, sendo a identidade média de 85 %. Apenas a cópia do DNAREP1_DM inserida na região 3' do gene CG9628 estava localizada na região próxima em até 3 kb do gene.

Em *D. simulans*, dentre os 35 genes *Cyps*, apenas seis (17,14 % do total de 35) possuíam cópias do DNAREP1_DM (Figura 7), as quais variaram de 34 a 332 pb. Em relação ao elemento consenso, a porcentagem de identidade das cópias variou de 74 a 92 %, sendo a identidade média 82 %. Do total de 15 cópias deste elemento presentes nos *Cyps*, sete estavam localizadas até 3 kb de distância dos genes. Os quatro genes *Cyps* de *D. melanogaster* que possuíam cópias do DNAREP1_DM são associados à resistência aos inseticidas (*Cyp6w1* e *Cyp6a2*) ou à atividade monooxigenase (*Cyp6a14* e *Cyp6a22*). No total de sete inserções, quatro (57,14 %) estão localizadas nas regiões intergênicas 5' (três inserções) e 3' (uma inserção) do gene *Cyp6w1*. Os outros três genes apresentaram apenas uma inserção cada um (*Cyp6a2_3'*, *Cyp6a14_3'* e *Cyp6a22_5'*).

No entanto, apenas duas cópias do elemento DNAREP1_DM foram localizadas nas regiões intergênicas dos genes flanqueadores aos *Cyps*. Tais cópias estavam inseridas na região 5' dos genes GD10156 e GD19271 e apresentaram, respectivamente, tamanho de 215 e 95 pb e porcentagem de identidade de 82,48 % e 78,95 %.

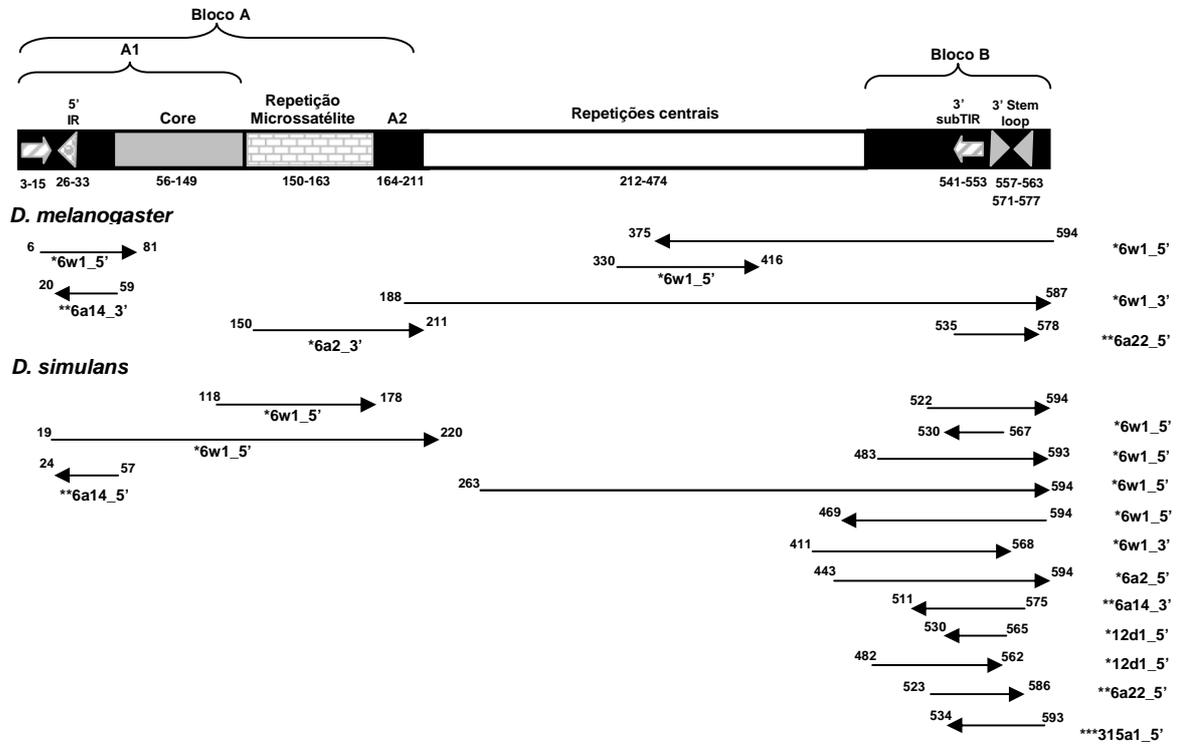


Figura 7. Localização das cópias do elemento DNAREP1_DM inseridas em genes *Cyps* de *D. melanogaster* e *D. simulans* na cópia consenso do elemento. **Cyps* associados à resistência; ** *Cyps* associados à atividade monooxigenase; *** *Cyps* associados ao desenvolvimento.

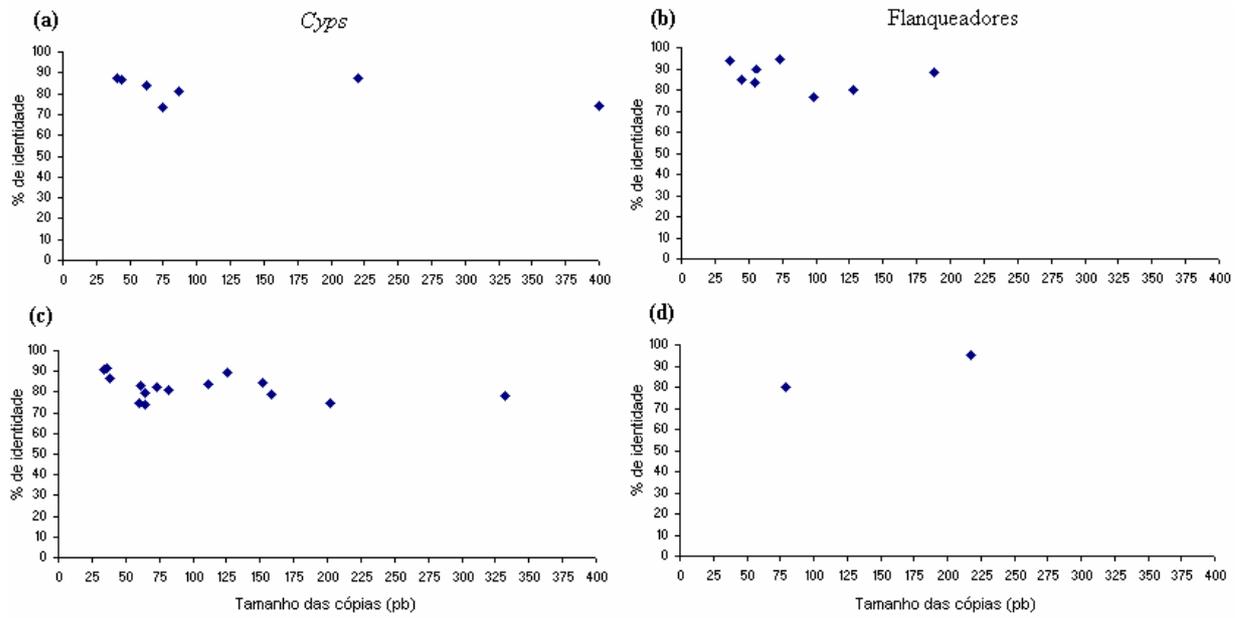


Figura 8. Porcentagens de identidade de cópias de diferentes tamanhos do elemento DNAREP1_DM inseridas em: (a) genes *Cyps* (b) genes flanqueadores de *D. melanogaster*; (c) genes *Cyps* (d) genes flanqueadores de *D. simulans*.

2.4.6 Investigação da presença de sítios de ligação de fatores de transcrição (FTs) nas cópias do DNAREP1_DM dos genes *Cyps*

A ligação de fatores de transcrição (FTs) ao DNA tem um importante papel na regulação da transcrição gênica. Assim, muitos sítios de ligação de fatores de transcrição são adjacentes a genes expressos. Utilizando a ferramenta *Alibaba2* (GRABE, 2002), foi realizada uma busca por possíveis sítios de ligação de fatores de transcrição nas cópias do elemento DNAREP1_DM presentes nas regiões intergênicas dos genes *Cyps* de *D. melanogaster* e *D. simulans* (Tabela 4). Todas as cópias do DNAREP1_DM que apresentam motivos homólogos a sítios de ligação de FTs, são seqüências correspondentes à região 3' (bloco B e parte das repetições centrais) do elemento consenso, com exceção de apenas uma cópia, em *D. simulans* (Figura 9).

Em *D. melanogaster*, os motivos foram encontrados em três cópias de DNAREP1_DM inseridas nas regiões intergênicas 5' (duas inserções, localizadas a aproximadamente 1 kb do início do gene) e 3' (uma inserção, localizada a 1.724 pb do início do gene) do gene *Cyp6w1*, as quais apresentam de três a quatro seqüências similares a FTs. Essas seqüências podem ser incluídas em quatro superclasses de fatores de transcrição. Da superclasse *Helix-turn-helix* foram localizados os sítios de ligação de FTs Oct-1 (fator octâmero ligante) envolvido na regulação de genes relacionados ao ciclo celular, e o HSTF (fator *heat shock*) em genes expressos quando submetidos a situações de estresse. Da superclasse *Zinc finger* foram encontrados os sítios de ligação de FTs Hb (fator *hunchback*), o qual age na fase do desenvolvimento do embrião, determinando sua segmentação, e Sp1 (proteína estimulante 1) inserido geralmente em genes reguladores do ciclo celular. O sítio de ligação de FTs NF-1 (fator nuclear 1) pertence à terceira superclasse, de domínios básicos, o qual também se liga a genes reguladores do ciclo celular. Da última superclasse, *beta-Scaffold*, foi identificado o sítio de ligação de FTs SRF (sinais de resposta externa), geralmente inserido em genes expressos em resposta a estímulos externos. Todos os sítios de ligação de FTs estão localizados na região das repetições centrais do elemento, com exceção do HSTF (*Cyp6w1_5'*) o qual corresponde ao início do bloco B. Entre eles o mais freqüente foi o Sp1, o qual está localizado em todas as cópias do DNAREP1_DM. Todos os sítios de ligação de FTs encontrados nessas cópias de *D. melanogaster* foram também localizados na cópia consenso do elemento, com exceção do Oct-1 e SRF.

Todos os motivos correspondentes aos sítios de ligação de FTs encontrados nas cópias do DNAREP1_DM de *D. simulans* foram localizados também no elemento consenso. Foram encontrados motivos em seis cópias de DNAREP1_DM, localizadas em três genes *Cyps* associados à resistência aos inseticidas: quatro cópias nas regiões intergênicas do gene *Cyp6w1* (três na região 5' e uma na região 3') e uma cópia na região intergênica 5' dos genes *Cyp6a2* e *Cyp12d1*. Apenas uma cópia de DNAREP1_DM no *Cyp6w1_5'* apresentou um motivo (TBP) correspondente à extremidade 5' do elemento consenso. Em outra cópia do *Cyp6w1_5'*, foram localizados motivos correspondentes à região das repetições centrais (Sp1, NF-1 e Hb) e ao bloco B do elemento (HSTF). As outras quatro cópias (*Cyp6w1_5'*, *Cyp6w1_3'*, *Cyp6a2_5'* e *Cyp12d1_5'*) possuem o mesmo motivo (HSTF) correspondente à região do bloco B, do elemento consenso.

Tabela 4. Predição de sítios de ligação de fatores de transcrição (FTs) na sequência consenso do DNAREP1_DM em *D. melanogaster* e *D. simulans*, de acordo com Alibaba2 (GRABE, 2002).

<i>D. melanogaster</i>			<i>D. simulans</i>				
Motivos (Consenso)	Seqüência	Gene	Posição no TE	Motivos (Consenso)	Seqüência	Gene	Posição no TE
Superclasse Helix-turn-helix							
Oct-1 (fator octâmero ligante) wTwnCATATk	ttttcatatt	<i>Cyp6w1_5'</i>	337-346	HSTF (fator heat shock) GmAryTTCKs	gaaatttgcg	<i>Cyp6w1_5'</i> <i>Cyp6w1_3'</i> <i>Cyp6a2_5'</i> <i>Cyp12d1_5'</i>	511-520
HSTF (fator heat shock) GmAryTTCKs	gaaatttgcg	<i>Cyp6w1_3'</i>	511-520	Superclasse Helix-turn-helix			
Superclasse Zinc finger							
Hb (fator hunchback) nmmAAAAAAC	gccaaaaaac	<i>Cyp6w1_5'</i>	377-386	Hb (fator hunchback) nmmAAAAAAC	gccaaaaaac	<i>Cyp6w1_5'</i>	377-386
Sp1 (proteína estimulante 1) CmCrCCCmyn	cacgccact	<i>Cyp6w1_3'</i>	266-275	Sp1 (proteína estimulante 1) CmCrCCCmyn	cacgccact	<i>Cyp6w1_5'</i>	394-403
CGCsCmnnwCT	cgccactct	<i>Cyp6w1_3'</i>	268-277	CGCsCmnnwCT	cgccactct	<i>Cyp6w1_5'</i>	396-405
mymCGCCymy	ccacgccac	<i>Cyp6w1_5'</i>	396-405	mymCGCCymy	ccacgccac	<i>Cyp6w1_5'</i>	303-312
krGGyGksGy	gccacgccca	<i>Cyp6w1_3'</i>	303-312	Superclasse domínios básicos			
GwGGGnGnGG	ccacgccac	<i>Cyp6w1_5'</i>	393-402	NF-1 (fator nuclear 1) wnnTTGGCAA	ttgccaaaa	<i>Cyp6w1_5'</i>	375-384
GTGrGsGkGr	tcacgccac		394-403	Superclasse beta-Scaffold			
nGTGkGGGnG	cgccacact		432-441	SRF (sinais de resposta externa) wwTGGGyrkC	ggcgccaaa	<i>Cyp6w1_5'</i>	418-427
			435-444	Motivos: nome e seqüência consenso do fator identificado no banco de dados TRANSFAC; <i>D. melanogaster</i> : seqüências identificadas como motivos na seqüência do DNAREP1 de <i>D. melanogaster</i> ; posição nt : a posição de cada motivo na seqüência DNAREP1; <i>D. simulans</i> : seqüências identificadas como motivos na seqüência de DNAREP1 de <i>D. simulans</i> .			

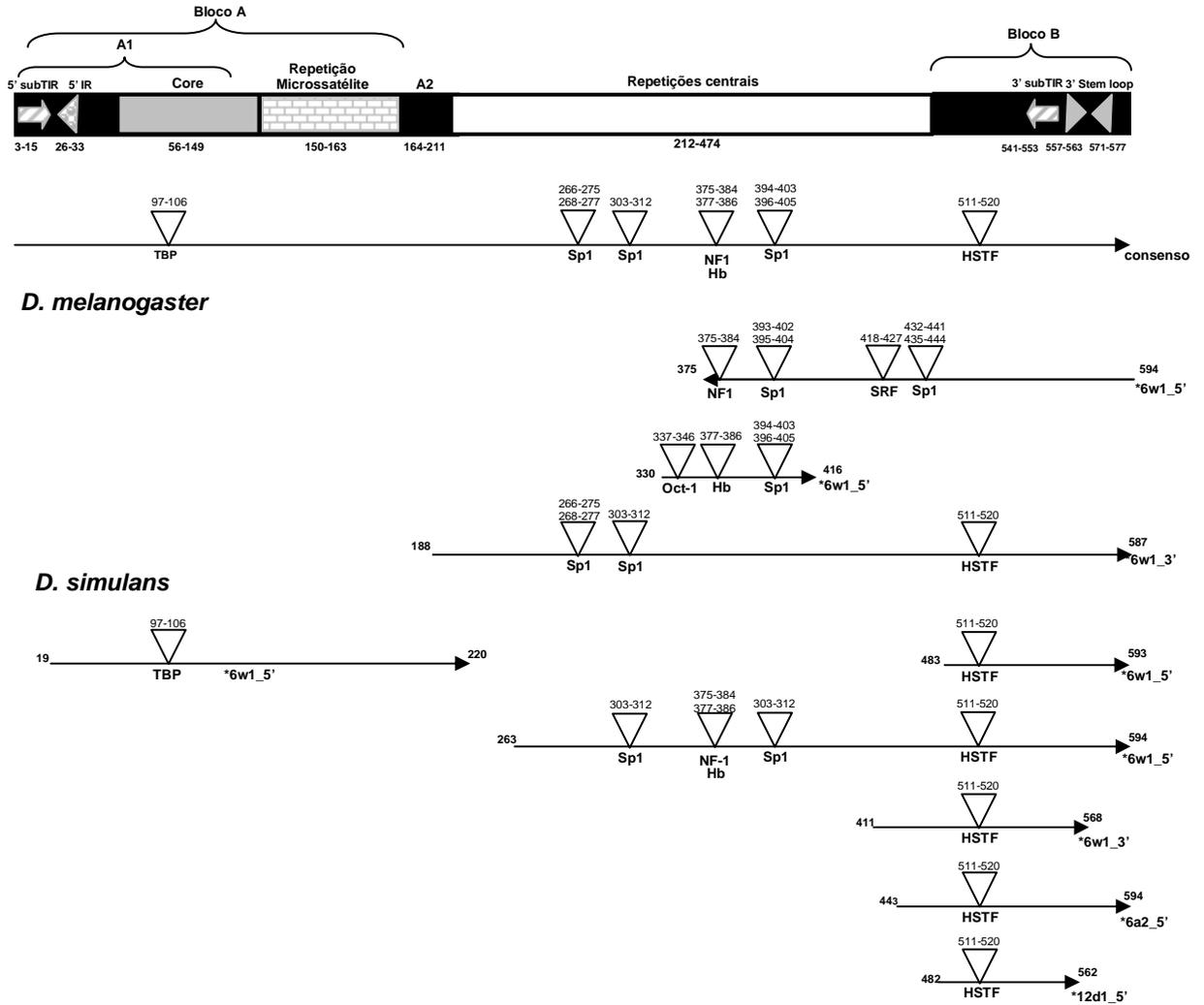


Figura 9. Possíveis sítios de ligação de fatores de transcrição identificados pela ferramenta *Alibaba2* nas cópias do TE DNAREP1_DM inseridas em genes *Cyps* de *D. melanogaster* e *D. simulans*. O símbolo ∇ corresponde à localização do possível sítio de ligação de FTs.

2.4.7 TEs em regiões flaqueadoras de genes *Cyps* em linhagens resistentes e suscetíveis de *D. melanogaster* e *D. simulans*

Os genes *Cyps* associados à resistência foram selecionados para serem investigados quanto à ocorrência inserções de TEs em suas regiões flaqueadoras 5' e 3' em linhagens resistentes e suscetíveis de *D. melanogaster* e *D. simulans*. Esses genes foram selecionados a partir dos resultados obtidos da análise *in silico*, em face de apresentarem inserções diferenciais de TEs em suas regiões flaqueadoras nas linhagens que tiveram seus genomas seqüenciados, sendo, portanto, potenciais candidatos a apresentarem polimorfismos populacionais relacionados com a resistência aos inseticidas. Foram eles os genes *Cyp6w1*, *Cyp6a2*, *Cyp4e2*, *Cyp12d1* e *Cyp12a4*.

O gene *Cyp6w1* foi investigado, em *D. melanogaster* por apresentar três inserções do transposon DNAREP1_DM em sua região flaqueadora 5' e uma inserção na região 3'. *D. simulans*, apresentou em sua região flaqueadora 5' uma inserção do DNAREP1_DM e duas do Helitron-1_Dyak. Para o *Cyp6a2* *D. melanogaster* apresentou na região flaqueadora 3' uma inserção do DNAREP1_DM e uma do ISFUN1. Por outro lado, na região flaqueadora 5' de *D. simulans* ocorreram uma inserção DNAREP1_DM e outra do Helitron-1_Dyak, enquanto que na região 3' foram encontradas uma cópia do TE Helitron-1_Dyak e do MINIME_DN, e duas cópias do TE Helitron-1_Dvir. No gene *Cyp4e2* de *D. melanogaster*, ocorreu apenas uma inserção do elemento ISFUN1 na região flaqueadora 5. O gene *Cyp12d1*, que apesar de não conter inserções nas regiões flaqueadoras da linhagem seqüenciada de *D. melanogaster*, foi analisado nas linhagens de ambas as espécies, pois é um *Cyp* associado à resistência aos inseticidas, e por conter em *D. simulans* duas inserções do DNAREP1_DM na região flaqueadora 5'.

As análises de PCR evidenciam a presença de polimorfismo interpopulacional de tamanho das regiões flaqueadoras dos genes *Cyp6w1*, *Cyp6a2* e *Cyp12d1* (Figura 10). Não parece existir qualquer padrão relacionado à resistência aos inseticidas. O padrão de bandas não difere em relação a essa característica, nem em *D. melanogaster* (linhagens resistentes: LE, UM, ITA, RP, BG, SM, TAV; linhagens suscetíveis: SE, CS), nem em *D. simulans* (linhagens resistentes: ITA, PF, OV, RAT, BG; linhagens suscetíveis: LE, MU, AMI, MAK, MAD). As bandas polimórficas, ou aquelas monomórficas que supostamente apresentavam um TE na região flaqueadora do

respectivo gene foram seqüenciadas em algumas linhagens, de modo a ser realizada uma análise populacional das inserções nas diferentes linhagens e espécies. A Tabela 5 apresenta as informações sobre os fragmentos de TEs encontrados nas regiões flangeadoras desses genes.

Apesar de não ocorrer polimorfismo, a região flangeadora 5' do gene *Cyp6w1* foi seqüenciada em cinco linhagens de *D. melanogaster*, para verificar se as inserções presentes nas populações correspondiam às mesmas da análise *in silico*. Todas, exceto *RP*, apresentaram uma das três inserções do elemento DNAREP1_DM (fragmento 594-375), que estava presente na linhagem genômica. Além disso, as linhagens *ITA*, *BG* e *TAV* apresentaram uma inserção do Helitron-1_Dyak (fragmento 314-415), que não ocorre linhagem genômica. Em *D. simulans*, foi seqüenciada a região flangeadora 5' do gene *Cyp6w1* na linhagem *PF*, que apresentou exatamente as mesmas inserções encontradas na linhagem genômica. Para a análise na região flangeadora 3' do gene *Cyp6w1* de *D. melanogaster* foram seqüenciadas as linhagens *RP* e *BG*. A primeira apresentou uma inserção do DNAREP1_DM (fragmento 419-587), que corresponde à parte da cópia presente na linhagem genômica e segunda não apresentou inserção de TEs nessa região. Para *D. simulans*, essa região foi monomórfica.

A região flangeadora 5' do gene *Cyp6a2* de *D. melanogaster* foi monomórfica, sem inserções de TEs. Em *D. simulans*, foi seqüenciada a linhagem *ITA*, que apresentou a mesma inserção do DNAREP1_DM, mas não a inserção do Helitron-1_Dyak, presentes na linhagem genômica. Na região flangeadora 3' da linhagem *RP* em *D. melanogaster* foi encontrada uma inserção do TE Helitron-1N1_Dvir (fragmento 108-263), que difere do fragmento encontrado na linhagem genômica, mas não as inserções dos elementos ISFUN1 e DNAREP1_DM encontrados nessa linhagem. Em *D. simulans*, essa região foi analisada nas linhagens *AMI* e *MAK*, que apresentaram a mesma inserção presente em na linhagem *RP* *D. melanogaster*, o Helitron-1N1_Dvir. No entanto, *AMI* apresentou mais duas inserções, uma do Helitron-1_Dvir (108-212) e outra do elemento MINIME_DN (fragmento 319-474), que correspondem às cópias presentes na linhagem genômica, mas não a inserção do Helitron-1_Dyak.

A análise *in silico* mostrou que o gene *Cyp12d1* não possuía nenhuma inserção de TEs em suas regiões flangeadoras, mas, por se tratar de um *Cyp* associado à resistência aos inseticidas, essa região foi seqüenciada em linhagens de *D. melanogaster* (*LE*, *ITA*, *RP* e *BG*) para verificar um possível polimorfismo de inserção de TEs em populações naturais. Em concordância com os resultados da análise *in silico*, não foram encontradas inserções nessas regiões.

Foi calculada a diversidade nucleotídica (π) das cópias inseridas nas regiões flangeadoras 5' e 3' de duas ou mais linhagens. O valor de π dos fragmentos do DNAREP1_DM inseridos nas regiões flangeadoras 5' e 3' do gene *Cyp6w1* nas diferentes linhagens de *D. melanogaster* (quatro linhagens geográficas mais a linhagem genômica), foi igual a 0.011. Para *D. simulans* (linhagem genômica e PF), o valor π foi igual a 0. Já, as cópias do Helitron-1_Dyak, que ocorrem nas mesmas linhagens de *D. melanogaster*, têm uma diversidade muito menor (π : 0.006). A diversidade nucleotídica do Helitron-1N1_Dvir presente na região flangeadora 3' do *Cyp6a2* das linhagens de *D. melanogaster* (genômica e RP) foi igual a 0.097 e nas linhagens de *D. simulans* (genômica, AMI e MAK) foi igual a 0.028.

A análise populacional das regiões flangeadoras dos genes *Cyp6w1* e *Cyp6a2* também confirma os dados da análise *in silico* das linhagens genômicas de que as cópias do elemento DNAREP1_DM, presentes nas regiões flangeadoras desses genes, correspondem à extremidade 3' do elemento consenso. Assim, também foi realizada uma busca por possíveis sítios de ligação de fatores de transcrição (FTs) nas cópias seqüenciadas do DNAREP1_DM provenientes das linhagens geográficas. Foram encontrados os mesmos FTs da análise *in silico* (Tabela 4), como também novos sítios (C/EBPalpha, GATA-1, Repressor, Kr, NF-ATc3 e HNF-3B). As cópias do elemento Helitron-1N1_Dvir presentes na região 3' do mesmo gene das linhagens de *D. melanogaster* (RP) e duas linhagens de *D. simulans* (AMI e MAK) não apresentaram possíveis sítios de ligação para FTs.

Tabela 5. Fragmentos de TEs encontrados em genes *Cyps* cujas regiões flanqueadoras 5' e 3' foram seqüenciadas (+: inserção em sentido sense; -: inserção em sentido antisense; FTs: sítios de ligação de fatores de transcrição).

Genes	Elemento	Posição no TE	FTs nos TEs
<i>Cyp6w1_5'_DmGenômica</i>	DNAREP1_DM	(+) 6-81 (-) 594-375 (+) 330-416	- TBP NF-1, C/EBPalha
<i>Cyp6w1_5'_DmLE</i>	DNAREP1_DM	(-) 594-375	TBP
<i>Cyp6w1_5'_DmITA</i>	DNAREP1_DM Helitron-1_Dyak	(-) 594-375 (+) 314-415	TBP, NF-1, C/EBPalha C/EBPalha, NF-1
<i>Cyp6w1_5'_DmRP</i>	-		
<i>Cyp6w1_5'_DmBG</i>	DNAREP1_DM Helitron-1_Dyak	(-) 594-375 (+) 314-415	TBP, NF-1, C/EBPalha, Oct-1 NF-1, C/EBPalha, Oct-1
<i>Cyp6w1_5'_DmTAV</i>	DNAREP1_DM Helitron-1_Dyak	(-) 594-375 (+) 314-415	TBP, NF-1, C/EBPalha NF-1, C/EBPalha, TBP
<i>Cyp6w1_5'_DsGenômica</i>	DNAREP1_DM Helitron-1_Dyak	(-) 594-469 (-) 564-314 (+) 1-415	C/EBPalha Repressor, Sp1, Hb, TBP, Kr C/EBPalha
<i>Cyp6w1_5'_DsPF</i>	DNAREP1_DM Helitron-1_Dyak Helitron-1_Dyak	(-) 594-469 (-) 564-314 (+) 1-415	Sp1, C/EBPalha, SRF, Repressor, Kr Repressor, Sp1, C/EBPalha, SRF, Kr C/EBPalha, Sp1
<i>Cyp6w1_3'_DmGenômica</i>	DNAREP1_DM	(+) 188-587	NF-ATc3, YBP, HNF-3B, Hb, Oct-1, GATA-1
<i>Cyp6w1_3'_DmRP</i>	DNAREP1_DM	(+) 419-587	TBP, Hb, C/EBPalha, GATA-1,
<i>Cyp6w1_3'_DmBG</i>	-		
<i>Cyp6a2_5'_DsGenômica</i>	Helitron-1_Dyak	(+) 1-392	TBP, GATA-1, Sp1, C/EBPalha, HNF-3B
	DNAREP1_DM	(+) 443-594	GATA-1
<i>Cyp6a2_5'_DsITA</i>	DNAREP1_DM	(+) 443-594	GATA-1
<i>Cyp6a2_3'_DmGenômica</i>	ISFUN1 DNAREP1_DM	(+) 87-208 (+) 150-211	- -
<i>Cyp6a2_3'_DmRP</i>	Helitron-1N1_Dvir	(+) 108-263	-
<i>Cyp6a2_3'_DsGenômica</i>	Helitron-1_Dvir MINIME_DN Helitron-1N1_Dvir Helitron-1_Dyak	(+) 108-212 (+) 319-475 (+) 148-263 (+) 135-211	- - - -
<i>Cyp6a2_3'_DsAMI</i>	Helitron-1_Dvir MINIME_DN Helitron-1N1_Dvir	(+) 108-212 (+) 319-474 (+) 148-263	- - -
<i>Cyp6a2_3'_DsMAK</i>	Helitron-1N1_Dvir	(+)108-263	-

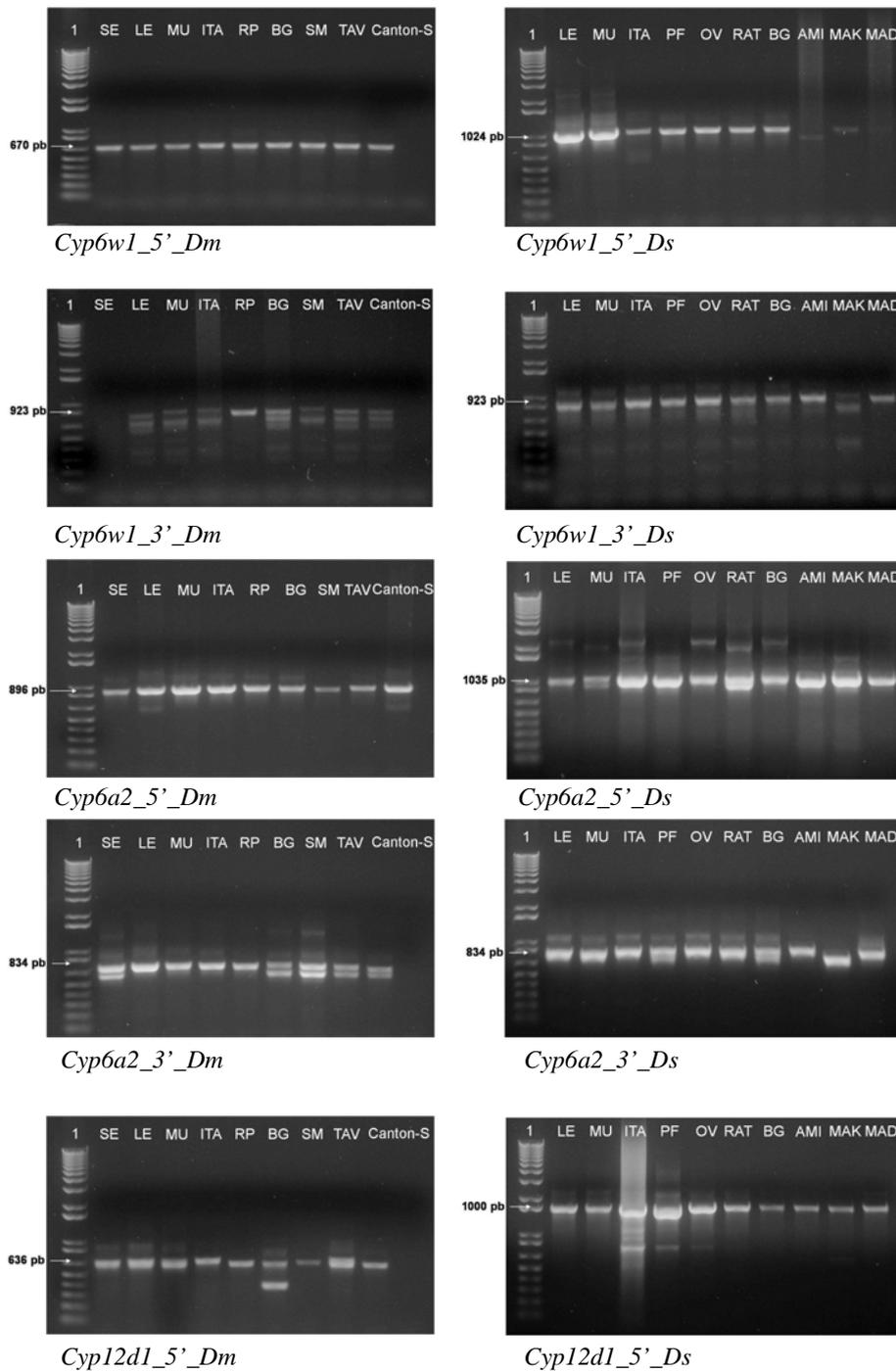


Figura 10. Análises por PCR das regiões flangeadoras 5'e 3'dos genes *Cyp6w1*, *Cyp6a2* e *Cyp12d1* em linhagens de *D. melanogaster* (SE: Aracaju, SE; LE: Lençóis, BA; MU: Mucuri, BA; ITA: Itaúnas, ES; RP: São José do Rio Preto, SP; BG: Bento Gonçalves, RS; SM: Santa Maria, RS; TAV: Tavares, RS; Canton-S: laboratório) e de *D. simulans* (LE: Lençóis, BA; UM: Mucuri, BA; ITA: Itaúnas, ES; PF: Paulo de Faria, SP; OV: Onda Verde, SP; RAT: Ratoões, SC; BG: Bento Gonçalves, RS; AMI : AMIEU, Nova Caledônia; MAK: Makindu, Quênia; MAD: Madagáscar, África).1 (1Kb Plus DNA Ladder).

2.5 DISCUSSÃO

Genes responsáveis por codificarem hormônios, fatores de transcrição e outros produtos envolvidos no desenvolvimento de um organismo, são considerados essenciais para sua sobrevivência. Devido a essa restrição funcional, tais genes tendem a ser altamente conservados, e a ocorrência de mutações deletérias, dentro ou próximas a eles, resulta em efeitos letais (WAGNER et al., 2003; SIMONS et al., 2006). Por outro lado, genes ligados a respostas ambientais e a estímulos externos, tais como aqueles envolvidos no metabolismo de xenobióticos, no sistema imunológico e de resposta ao estresse, apresentam maior plasticidade genômica (VAN DE LAGEMAAT et al., 2003; CHEN; LI, 2007).

Os genes da família Citocromo P450 monooxigenases (*Cyps*) são responsáveis pela detoxificação de xenobióticos, dentre eles, os inseticidas (AMICHOT et al., 2004). A inserção de TEs nas proximidades desses genes pode alterar a regulação e causar um forte impacto na expressão gênica e, conseqüentemente, conduzir à superexpressão das enzimas CYPs, o que pode induzir a resistência aos inseticidas (DABORN et al., 2001; 2002; CATANIA et al., 2004; JOUBEN et al., 2008). Entretanto, a ocorrência de TEs em genes essenciais para o funcionamento do organismo, devido à alta restrição funcional, geralmente é deletéria e estes são eliminados da população.

A relação entre genes envolvidos na imunidade ou a respostas a estímulos externos e maior número de TEs em suas regiões promotoras, que em genes relacionados ao desenvolvimento foi demonstrada, *in silico*, por van de Lagemaat et al. (2003), em regiões transcritas dos genes de humanos e camundongo, e por Chen e Li (2007), em *Cyps* associados a resistência, em *D. melanogaster*. Em face desses resultados, os autores formularam a hipótese de que inserções de TEs nos *Cyps* relacionados à resistência a xenobióticos podem ser adaptativas e serem mantidas nas populações, enquanto aquelas em genes relacionados ao desenvolvimento estariam sob forte pressão seletiva negativa, resultando em maior plasticidade genômica nos primeiros, e maior conservação nos últimos. Tal fonte de variabilidade genética seria vantajosa para organismos, tais como os insetos, que ocupam ambientes altamente variáveis e seletivos.

2.5.1 Inserção de TEs em genes *Cyps* e seus flaqueadores

A maior parte das inserções de TEs nos 35 genes investigados neste trabalho, em *D. melanogaster* e *D. simulans*, ocorreu em seqüências não codificantes (regiões intergênicas e introns), e foram praticamente ausentes em regiões codificantes, como descrito anteriormente em *Drosophila* (MILLER et al., 2000; KAMINKER et al., 2002; FONTANILLAS et al., 2007). Esta distribuição se deve ao fato de inserções em regiões não codificantes serem mais toleradas enquanto que, inserções em regiões codificantes ou próximas a elas podem causar efeitos deletérios, e assim, serem removidas por seleção normalizadora (LIPATOV et al., 2005; SELA et al., 2007; YANG; BARBASH, 2008).

Para testarmos a hipótese de que genes *Cyps* relacionados à resistência aos inseticidas apresentam maior plasticidade genômica, foram comparados os números de inserções de TEs em genes *Cyps* com diferentes funções, em genes flaqueadores dos *Cyps* e em uma amostra aleatória de 35 genes, tanto em *D. melanogaster*, como em *D. simulans*. Neste estudo, as inserções de TEs foram, em média (total de inserções/número de genes), mais freqüentes em regiões intergênicas de genes *Cyps* associados à resistência aos inseticidas, tanto em *D. melanogaster* (três vezes maior) como em *D. simulans* (seis vezes maior) do que nos *Cyps* associados ao desenvolvimento. Os genes flaqueadores aos *Cyps* em *D. simulans* também apresentaram maior número de inserções na região intergênica, porém, os flaqueadores aos *Cyps* em *D. melanogaster* apresentaram maior número de inserções em regiões gênicas, principalmente nos introns. Diferente dos *Cyps*, a maior média de inserções ocorreu nos flaqueadores dos *Cyps* associados à atividade monooxigenase, em ambas as espécies. Esses resultados sugerem que a pressão seletiva contrária à inserção, ou manutenção de TEs, em genes associados ao desenvolvimento se estende aos genes vizinhos.

As inserções de TEs encontradas nas regiões intergênicas não correspondem a elementos completos, ao contrário, são fragmentos pequenos (*D. melanogaster*: 564 nt, *D. simulans*: 196 nt, em média), correspondentes em média a 15,49 % (*D. melanogaster*) e 4,93 % (*D. simulans*) do tamanho das seqüências consensos (3.641 pb e 3.970 pb) dos TEs. Essas inserções, de forma geral, localizam-se preferencialmente na região intergênica 5' dos genes, como também descrito por Yant et al. (2005), em humanos e camundongo. Este resultado suporta a hipótese de que

fragmentos de TEs enriquecem as regiões promotoras, doando sinais reguladores da transcrição e podendo alterar a expressão gênica, como relatado em humanos por Jordan et al. (2003) e Thornburg et al. (2006). A comparação do número de inserções nas regiões flangeadoras distantes em até 3 kb do início dos genes, independente de sua categoria funcional, reforça essa hipótese apenas para os *Cyps* de *D. melanogaster*. Entretanto, considerando as diferentes classes funcionais, os genes *Cyps* associados à resistência apresentaram, em média, duas vezes mais inserções em *D. melanogaster*, e quatro vezes mais em *D. simulans*, do que nas demais categorias. Esses resultados reforçam a proposta já discutida anteriormente de que fragmentos de TEs são seletivamente enriquecidos em genes associados à resistência em *Drosophila* (CHEN; LI, 2007) e indicam que *D. melanogaster* e *D. simulans* apresentam características particulares não apenas quanto à proporção genômica de TEs, como também quanto a proporção de TEs inseridos nas regiões adjacentes aos *Cyps*.

Em *D. melanogaster*, as classes de TEs com maior abundância (BERGMAN et al., 2006) são, em ordem decrescente: DNA (56,6 %: 3.050/5.390) > LTR (24,5 %: 1.321/5.390) > Retroposons (18,9 %: 1.019/5.390), sendo as proporções em parênteses correspondentes ao número de TEs de cada classe em relação ao número total de TEs do genoma. Quando excluído o DNAREP1_DM da classe dos transposons de DNA, em decorrência da dúvida quanto a sua classificação, a proporção de TEs se altera para: LTR > Retroposons > DNA. Nossos resultados mostram que, de modo geral, mesmo excluindo o DNAREP1_DM, os genes *Cyps* apresentam maior quantidade de transposons de DNA, pois, outros elementos de DNA como o PROTOP_A e Helitron-1_Dyak, também apresentaram elevado número de cópias, seguidos pelos Retroposons e LTRs. Tendência à inserção de transposons próximos a regiões reguladoras de genes pode explicar esse resultado, como relatado para o transposon *P* (TOWER et al., 1993; SPRADLING et al., 1995) e *Sleeping Beauty* (YANT et al., 2005). Outra explicação poderia ser que a inserção de diferentes elementos em todas as regiões intergênicas é casual, mas ocorre manutenção preferencial dos elementos de DNA na região 5' dos genes. Feschotte (2008) observou que a manutenção de transposons de DNA no genoma, ou de seqüências codificantes de suas transposases, se deve ao fato de que essas seqüências podem ser fontes para ligação de fatores de transcrição. Assim, os fragmentos de TEs podem representar um importante papel na regulação gênica, sendo mantidos ou até exaptados nas proximidades dos genes.

2.5.2 Inserção do elemento DNAREP1_DM

O elemento DNAREP1_DM ocorre em elevado número de cópias nas doze espécies de *Drosophila* que tiveram seus genomas seqüenciados estando, a maioria delas, inserida em introns e regiões intergênicas (YANG; BARBASH, 2008). De acordo com Yang e Barbash (2008), as cópias desse TE em *D. melanogaster*, *D. simulans*, *D. sechellia*, *D. erecta* e *D. grimshawi* apresentam identidade média de 90 %, mas em *D. yakuba*, *D. ananassae*, *D. persimilis*, *D. pseudoobscura* e *D. willistoni*, a identidade é maior que 95 %; por isso, esses autores sugeriram que o DNAREP1_DM deve ter sofrido uma recente explosão transposicional nas espécies do segundo grupo, por apresentarem cópias jovens e pouco divergentes.

Nosso estudo mostrou que as sete cópias do DNAREP1_DM localizadas nas regiões intergênicas dos genes *Cyps* de *D. melanogaster*, dessas 6 localizadas nas regiões próximas até 3 kb de cada gene, apresentaram identidade média de 82 %. As 15 cópias (seis inseridas em até 3kb do gene) dos *Cyps* de *D. simulans* apresentaram 81 % de identidade média. Esses resultados indicam que as inserções do DNAREP1_DM em regiões intergênicas dos *Cyps* são mais divergentes do que a média das cópias distribuídas ao longo do genoma. Este percentual inferior de identidade, associado ao pequeno tamanho, corroboram o fato de serem cópias antigas nas duas espécies, desse modo já expostas ao crivo da seleção normalizadora que resulta em eliminação de TEs deletérios em regiões próximas a genes. Segundo Gotea e Makalowski (2006), são necessários longos períodos evolutivos para que as cópias de TEs sejam exaptadas e possam adquirir uma função biológica. Assim, cópias jovens, além de sofrerem grande pressão seletiva, não teriam tido tempo necessário para evoluírem e adquirirem novas funções. A presença de cópias pequenas e divergentes do elemento DNAREP1_DM, inseridas preferencialmente nas regiões 5' dos genes e, apresentando putativos sinais reguladores da transcrição gênica (motivos de ligação de fatores de transcrição), sugerem que possam atuar de maneira positiva na expressão de genes *Cyps*.

2.5.3 Análises de sequências de TEs em regiões flanqueadoras de genes *Cyp*s em linhagens de *D. melanogaster* e *D. simulans* suscetíveis e resistentes a inseticidas

A diversidade nucleotídica entre algumas cópias dos elementos mais frequentes nas linhagens geográficas foi utilizada para avaliar o grau de diversidade em diferentes elementos inseridos nas regiões flanqueadoras dos genes *Cyp*s associados à resistência aos inseticidas. Embora os valores de π tenham sido obtidos com um número reduzido de sequências, foi feita uma comparação aproximada com o nível de diversidade nucleotídica de outros TEs. Os valores estimados para os TEs DNAREP1_DM e Helitron-1N1_Dvir foram semelhantes aos encontrados por Subramanian et al. (2008) para TEs presentes em populações de *Anopheles gambiae*. O elemento *Topi*, por exemplo, localizado em regiões codificantes, apresentou um elevado nível de diversidade nucleotídica ($\pi = 0.051$), sugerindo serem residentes antigos nessa espécie. O mesmo pode ser concluído em relação aos elementos DNAREP1_DM ($\pi = 0.011$) e Helitron-1N1_Dvir ($\pi = 0.028$ e 0.097), inseridos nas regiões flanqueadoras dos *Cyp*s de *D. melanogaster* e *D. simulans*. Essas cópias antigas poderiam ser remanescentes do crivo da seleção normalizadora e terem adquirido uma função biológica (GOTEA; MAKALOWSKI, 2006), atuando na expressão dos genes vizinhos, devido a esses elementos transportarem diferentes sítios de ligação para fatores de transcrição.

2.6 CONCLUSÃO

Neste trabalho não foi possível identificar qualquer diferença quanto à frequência de elementos de transposição nas regiões flanqueadoras de genes *Cyp*s associados à resistência aos inseticidas em linhagens de *D. melanogaster* e *D. simulans* resistentes ou suscetíveis. Entretanto, foi possível mostrar que os genes associados à resistência aos inseticidas acumulam fragmentos de TEs em maior proporção que os demais *Cyp*s e que essas proporções variam nas duas espécies crípticas, *D. melanogaster* e *D. simulans*. Um desses elementos foi o DNAREP1_DM, que ocorre nesses genes como fragmentos antigos, correspondentes à extremidade 3' do elemento consenso, que contém seqüências homólogas a sítios de ligação de fatores de transcrição. Essas características corroboram achados anteriores (DABORN et al., 2002; CATANIA et al., 2004; CHUNG et al. 2007) de que TEs interfiram na transcrição, elevando a expressão desses genes.

O que torna os TEs ricos estoques de elementos cis-regulatórios? De acordo com Feschotte (2008) uma explicação pode ser que o acúmulo de TEs cria o material bruto para evolução “de novo” dessas seqüências por mutação, desde que a maior parte dos sítios de ligação de fatores de transcrição é constituída por seqüências curtas e degeneradas (WRAY et al., 2003). Desse modo, mutações de ponto poderiam originar essas seqüências em TEs. Outro cenário, segundo Feschotte (2008), é que os elementos cis-regulatórios pré-existam nos TEs, no momento de sua inserção, e sejam co-optados imediatamente após a inserção, ou após a modificação do ambiente flanqueador. Uma larga variedade de elementos regulatórios tem sido identificada em seqüências ativas ou consensos reconstruídos de TEs ativos (JOHNSON et al., 2006; POLAK & DOMANY, 2006; THORNBURG et al., 2006; WANG, et al. 2007) e muitos estudos empíricos demonstram que esses elementos são incorporados nos aparatos regulatórios dos genes adjacentes (WESSLER et al., 1995; FERRIGNO, et al., 2001; BROSIUS, 2003; MEDSTRAND et al., 2005; ROMANISH et al., 2007). Esse segundo cenário ajusta-se aos resultados obtidos para o elemento DNAREP1_DM neste trabalho, entretanto, estudos empíricos com transfecção e testes de expressão dessas putativas seqüências reguladoras são necessários para corroborar essa hipótese.

REFERÊNCIAS

- ALTSCHUL, S. F. et al. Basic local alignment search tool. *J. Mol. Biol.*, v. 215, p. 403-410, 1990.
- AMICHOT, M. et al. Point mutations associated with insecticide resistance in the *Drosophila* cytochrome P450 *Cyp6a2* enable DDT metabolism. *Eur. J. Biochem.*, v. 271, p. 1250-1257, 2004.
- AQUADRO, C. F.; LADO, K. M.; NOON, W. A. The rosy region of *Drosophila melanogaster* and *Drosophila simulans*. 1. Contrasting levels of naturally-occurring DNA restriction map variation and divergence. *Genetics*, v. 119, p. 875-888, 1988.
- BERGMAN, C. M. et al. Recurrent insertion and duplication generate networks of transposable elements in the *Drosophila melanogaster* genome. *Genome Biol.*, v. 7, R112, 2006.
- BOGWITZ, M. R. et al. CYP12A4 confers lufenuron resistance in a natural population of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, v. 102, p. 12807-12812, 2005.
- BRANDT, A. et al. Differential expression and induction of two *Drosophila* cytochrome P450 genes near the *Rst(2)DDT* locus. *Insect Mol. Biol.*, v. 11, p. 337-341, 2002.
- BROSIUS, J. The contribution of RNAs and retroposition to evolutionary novelties. *Genetica*, v. 118, p. 99-116, 2003.
- CATANIA, F. et al. World-wide survey of an *Accord* insertion and its association with DDT resistance in *Drosophila melanogaster*. *Mol. Ecol.*, v. 13, p. 2491-2504, 2004.
- CHEN, S.; LI, X. Transposable elements are enriched within or in close proximity to xenobiotic-metabolizing cytochrome P450 genes. *BMC Evol. Biol.*, v. 7, 2007.
- CHOUDHARY, M.; SINGH, R. S. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster* .III. variations in genetic-structure and their causes between *Drosophila melanogaster* and its sibling species *Drosophila simulans*. *Genetics*, v. 117, p. 697-710, 1987.

CHUNG, H. et al. Cis-regulatory elements in the *Accord* retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene *Cyp6g1*. *Genetics*, v. 175, p. 1071-1077, 2007.

CLARCK, A. G. et al. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*, v. 450, p. 203-218, 2007.

CONTE, C.; DASTUGUE, B.; VAURY, C. Coupling of enhancer and insulator properties identified in two retrotransposons modulates their mutagenic impact on nearby genes. *Mol. Cell. Biol.*, v. 22, p. 1767-1777, 2002.

DABORN, P. et al. DDT resistance in *Drosophila* correlates with *Cyp6g1* over-expression and confers cross-resistance to the neonicotinoid imidacloprid. *Mol. Genet. Genomics*, v. 266, p. 556-563, 2001.

DABORN, P. J. et al. A single P450 allele associated with insecticide resistance in *Drosophila*. *Science*, v. 297, p. 2253-2256, 2002.

DABORN, P. J. et al. Evaluating the insecticide resistance potential of eight *Drosophila melanogaster* cytochrome P450 genes by transgenic over-expression. *Insect Biochem. Mol. Biol.*, v. 37, p. 512-519, 2007.

DANIELS, S. B.; STRAUSBAUGH, L. D. The distribution of *P* element sequences in *Drosophila*: the *willistoni* and *saltans* species groups. *J. Mol. Evol.*, v. 23 p. 138-148, 1986.

FERRIGNO, O. et al. Transposable B2 SINE elements can provide mobile RNA polymerase II promoters. *Nature Genet.*, v. 28, p. 77-81, 2001.

FESCHOTTE, C. *Review*. Transposable elements and the evolution of regulatory networks. *Nature Rev. Genet.*, v. 9, p. 397-405, 2008.

FESTUCCI-BUSELLI, R. A. et al. Expression of *Cyp6g1* e *Cyp12d1* in DDT resistant and susceptible strains of *Drosophila melanogaster*. *Insect Mol. Biol.*, v. 14, p. 69-77, 2005.

FONTANILLAS, P.; HARTL, D. L.; REUTER, M. Genome Organization and Gene Expression Shape the Transposable Element Distribution in the *Drosophila melanogaster* Euchromatin. *PLoS Genet.*, v. 3, p. 2256-2267, 2007.

FFRENCH-CONSTANT, R. H. et al. Why are there so few resistance-associated mutations in insecticide target genes? *Phil. Trans. R. Soc. Lond.*, v. 353, p. 1685-1693, 1998.

GENE ONTOLOGY CONSORTIUM. Creating the gene ontology resource: design and implementation. *Genome Res.*, v. 11, p. 1425-33, 2001.

GOTEA, V.; MAKALOWSKI, W. Do transposable elements really contribute to proteomes? *Trends Genet.*, v. 22, p. 260-267, 2006.

GRABE, N. Alibaba2: context specific identification of transcription factor binding sites. *In Silico Biol.*, v. 2, p. 1-15, 2002.

GRANZOTTO, A. Suscetibilidade ao DDT e imidaclopride de espécies invasoras e neotropicais da família Drosophilidae com ênfase na expressão dos genes *Cyp6g1* e *Cyp12d1*. 2007. 94f. Dissertação (Mestrado em Genética) - Universidade Estadual Paulista "Júlio de Mesquita Filho", São José do Rio Preto, 2007.

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, England, v. 41, p. 95-98, 1999.

HELVIG, C. et al. *Drosophila melanogaster* CYP6A8, an insect P450 that catalyzes lauric acid (ω -1)-hydroxylation. *Biochem. Biophys. Res. Commun.*, v. 325, p. 1495-1502, 2004.

JOHNSON, R. et al. Identification of the REST regulon reveals extensive transposable element-mediated binding site duplication. *Nucleic Acids Res.*, v. 34, p. 3862-3877, 2006.

JORDAN, I. K. et al. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.*, v. 19, p. 68-73, 2003.

JOUBEN, N. et al. Metabolism of imidacloprid and DDT by P450 CYP6G1 expressed in cell cultures of *Nicotiana tabacum* suggests detoxification of these insecticides in *Cyp6g1*-overexpressing strains of *Drosophila melanogaster*, leading to resistance. *Pest Manag. Sci.*, v. 64, p. 65-73, 2008.

JURKA, J. et al. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.*, v. 110, p. 462-467, 2005.

KAMINKER, J. S. et al. The transposable elements of the *Drosophila melanogaster* euchromatin: a genomic perspective. *Genome Biol.*, v. 3, 2002.

KAPITONOV, V. V.; JURKA, J. DNAREP1_DM. *Rebase Update Release 3.4*, 1999.

KAPITONOV, V. V.; JURKA, J. Molecular paleontology of transposable elements in *Drosophila melanogaster* genome. *Proc. Natl. Acad. Sci. USA*, v. 100, p. 6569-6574, 2003.

KURUGANTI, S. et al. High expression of *Cyp6g1*, a cytochrome P450 gene, does not necessarily confer DDT resistance in *Drosophila melanogaster*. *Gene*, v. 388, p. 43-53, 2007.

LE GOFF, G. et al. Microarray analysis of cytochrome P450 mediated insecticide resistance in *Drosophila*. *Insect Biochem. Mol. Biol.*, v. 33, p. 701-708, 2003.

LI, X.; SCHULER, M. A.; BERENBAUM, M. R. Molecular Mechanisms of Metabolic Resistance to Synthetic and Natural Xenobiotics. *Annu. Rev. Entomol.*, v. 52, p. 231-253, 2007.

LIPATOV, M. et al. Paucity of chimeric gene-transposable elements transcripts in the *Drosophila melanogaster* genome. *BMC Biology*, v. 3, 2005.

LOWE, C. B.; BEJERANO, G.; HAUSSLER, D. Thousands of human mobile elements fragments undergo strong purifying selection near developmental genes. *Proc. Natl. Acad. Sci. USA*, v. 104, p. 8005-8010, 2007.

MAITRA S. et al. Three second chromosome-linked clustered CYP6 genes show differential and barbital-induced expression in DDT-resistant and susceptible strains of *Drosophila melanogaster*. *Gene*, v. 180, p. 165-171, 1996.

MAITRA S. et al. Factors on the third chromosome affect the level of CYP6A2 and CYP6A8 expression in *Drosophila melanogaster*. *Gene*, v. 248, p. 147-156, 2000.

MARSANO, R. M. et al. Evidence for a functional interaction between the *Bari1* transposable element and the cytochrome P450 *cyp12a4* gene in *Drosophila melanogaster*. *Gene*, v. 357, p. 122-128, 2005.

- MARTINS-CAMPOS, J. M. et al. Intraspecific and interspecific variation at the y-ac-sc region of *Drosophila melanogaster* and *Drosophila simulans*. *Genetics*, v. 130, p. 805-16, 1992.
- McCART, C.; FFRENCH-CONSTANT, R. H. Dissecting the insecticide-resistance-associated cytochrome P450 gene *Cyp6g1*. *Pest Manag. Sci.*, v. 64, p. 639-645, 2008.
- MEDSTRAND, P, et al. *Review*. Impact of transposable elements on the evolution of mammalian gene regulation. *Cytogenet. Genome Res.*, v. 110, p. 342–352, 2005.
- MILLER, W. J. et al. Evolutionary Dinamics of the *SGM* transposon Family in the *Drosophila obscura* Species Group. *Mol. Biol. Evol.*, v. 17, p. 1597-1609, 2000.
- MORIYAMA, E. N.; POWELL, J. R. Intraspecific nuclear DNA variation in *Drosophila*. *Mol. Biol. Evol.*, v. 13, p. 261-277, 1996.
- NEI, M.; LI, W. H. Mathematical model for studying genetic variation in terms of restrictions endonucleases. *Proc. Natl. Acad. Sci. USA*, v. 76, p. 5269-5273, 1979.
- PEDRA, J. H. F. et al. Genome-wide transcription profile of field- and laboratory-selected ichlorodiphenyltrichloroethane (DDT)-resistant *Drosophila*. *Proc. Natl. Acad. Sci. USA*, v. 101, p. 7034-7039, 2004.
- POLAK, P; DOMANY, E. *Alu* elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes. *BMC Genomics*, v. 7, p. 133-148, 2006.
- ROMANISH, M. T. et al. Repeated recruitment of LTR retrotransposons as promoters by the anti-apoptotic locus NAIP during mammalian evolution. *PLoS Genet.*, v. 3, e10, 2007.
- ROZAS, J. et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, v. 19, p. 2496-2497, 2003.
- SCHLENKE, T. A.; BEGUN, D. J. Strong selective sweep associated with transposon insertion in *D. simulans*. *Proc. Natl. Acad. Sci. USA*, v. 101, p. 1626-1631, 2004.

SELA, N. et al. Comparative analysis of transposed element insertion within human and mouse genomes reveals *Alu*'s unique role in shaping the human transcriptome. *Genome Biol.*, v. 8, R137.1-19, 2007.

SIMONS, C. et al. Transposon-free regions in mammalian genomes. *Genome Res.*, v. 16, p. 164-172, 2006.

SPRADLING, A. C. et al. Gene disruptions using *P* transposable elements: An integral component of the *Drosophila* genome project. *Proc. Natl. Acad. Sci. USA*, v. 92, p. 10824-10830, 1995.

SUBRAMANIAN, R. A. et al. Topi, an IS630/Tc1/mariner-type transposable element in the African malaria mosquito, *Anopheles gambiae*. *Gene*, v. 423, p. 63-71, 2008.

TATUSOVA, T. A.; MADDEN, T. L. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol Lett.*, v. 174, p. 247-250, 1999.

THOMPSON, J. D. et al. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, Ingleterra, v. 22, p. 4673-4680, 1994.

THORNBURG, B. G.; GOTEA, V.; MAKALOWSKI, W. Transposable elements as a significant source of transcription regulating signals. *Gene*, v. 365, p. 104-110, 2006.

TIJET, N.; HELVIG, C.; FEYEREISEN, R. The cytochrome P450 gene superfamily in *Drosophila melanogaster*: Annotation, intron-exon organization and phylogeny. *Gene*, v. 262, p. 189-198, 2001.

TOWER, J. et al. Preferential Transposition of *P* Elements to Nearby Chromosomal Sites. *Genetics Society of America*, v. 133, p. 347-359, 1993.

VAN DE LAGEMAAT, L. M. et al. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trend Genet.*, v. 19, p. 530-536, 2003.

VIEIRA, C. et al. Wake up of transposable elements following *Drosophila simulans* worldwide colonization. *Mol. Biol. Evol.*, v.16, p. 1251-1255, 1999.

WAGNER, G. P.; AMEMIYA, C.; RUDDLE, F. Hox cluster duplications and the opportunity for evolutionary novelties. *Proc. Nat.l Acad. Sci.*, v. 100, p. 14603-14606, 2003.

WALLACE, M. A survey of wild-type strains of *Drosophila melanogaster* for resistance to malathion and phenulthiocarbamide. *Z. Zool. Syst. Evolutonforsh*, v.24, p. 52-58, 1986.

WANG, T. et al. Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proc. Nat.l Acad. Sci. U S A.*, v. 104, p. 18613-18618, 2007.

WESSLER, S. R.; BUREAU, T. E.; WHITE, S. E. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Curr. Opin. Genet. Dev.*, v. 5, p. 814–821, 1995.

WRAY, G. A. The Evolution of Transcriptional Regulation in Eukaryotes. *Mol. Biol. Evol.*, v. 20, p. 1377-1419, 2003.

WINDELSPECHT M., RICHMOND R. C., COCHRANE B. J. Malathion resistance levels in sympatric populations of *Drosophila simulans* (Diptera, Drosophilidae) and *Drosophila melanogaster* differ by 2 orders of magnitude. *J. Econ. Entomol.*, v. 88, p. 1138-1143, 1995.

YANG, H-P.; BARBASH, D. A. Abundant and species-specific *DINE-1* transposable elements in 12 *Drosophila* genomes. *Genome Biol.*, v. 9, R39, 2008.

YANT, S. R. et al. High-Resolution Genome–Wide Mapping of Transposon Integration in Mammals. *Mol. Cell. Biol.*, v. 25, p. 2085-2094, 2005.

Apêndice A - Número total e porcentagem (%) de inserções de TEs em regiões intergênicas 5' e 3', em relação ao total de inserções, em 35 genes *Cyps* associados com resistência aos inseticidas (N: 9), com atividade monooxigenase (N: 20) e ao desenvolvimento (N: 6) em *D. melanogaster*.

<i>Cyp</i> /função	n° de inserções		
	5'	3'	Total
Resistência aos inseticidas			
<i>Cyp6w1</i>	6	1	7
<i>Cyp6a2</i>		2	2
<i>Cyp4e2</i>	1 ^a		1
<i>Cyp12d1</i>	1		1
<i>Cyp6g1</i>			
<i>Cyp6a9</i>			
<i>Cyp6a8</i>	3 ^a		3
<i>Cyp12a4</i>		1	1
<i>Cyp308a1</i>			
%	73,33	26,6	
χ^2		10,42*	
Atividade monooxigenase			
<i>Cyp4ad1</i>	1 ^a		1
<i>Cyp4e1</i>			
<i>Cyp6a14</i>	2 ^a	1	3
<i>Cyp6a15Psi</i>	2 ^a		2
<i>Cyp6a13</i>			
<i>Cyp6g2</i>			
<i>Cyp6i3</i>			
<i>Cyp301a1</i>			
<i>Cyp9h1</i>		1	1
<i>Cyp6a22</i>	1		1
<i>Cyp6a17</i>			
<i>Cyp6a23</i>			
<i>Cyp6a19</i>			
<i>Cyp6a20</i>			
<i>Cyp6a21</i>			
<i>Cyp317a1</i>	3 ^a		3
<i>Cyp12c1</i>			
<i>Cyp12a5</i>			
<i>Cyp6a18</i>			
<i>Cyp4g1</i>			
%	81,81	18,1	
χ^2		9	
		19,60*	
Desenvolvimento			
<i>Cyp302a1</i>			
<i>Cyp307a1</i>			
<i>Cyp314a1</i>			
<i>Cyp315a1</i>			
<i>Cyp306a1</i>	3		3
<i>Cyp18a1</i>			
%	100		
χ^2			

N: número de genes.

* χ^2 : $p < 0,05$ entre o número de inserções 5' e 3'.

Inserções presentes na região intergênica comum a dois genes foram contadas para ambos: ^ainserção comum entre dois genes *Cyps*.

Apêndice B - Número total e porcentagem (%) de inserções de TEs em regiões intergênicas e gênicas em relação ao total de inserções em 36 genes flanqueadores aos *Cyps* associados com resistência aos inseticidas (N: 11), com atividade monooxigenase (N: 16) e ao desenvolvimento (N: 9) em *D. melanogaster*.

Genes flanqueadores aos <i>Cyps</i>	Função	n° de inserções					Total
		5'	3'	éxon	intron	UTR	
Resistência aos inseticidas							
EcR	receptor de ecdisona				14	3	17
CG8343	ligante da manose						
Epac	fosforilação				2		2
SdhB	succinato						
CG30490	<i>desconhecida</i>		1				1
BBS4	mobilidade do				2		2
DebB	biossíntese de ecdisona						
Chmp 1	proteína de transporte						
CG6893	proteína de transporte						
CG5629	sem <i>hit</i>						
CCKLR-17D1	receptor de membrana						
% χ^2			4,55		81,81	13,64	
Atividade monooxigenase							
Lcp4	proteína da cutícula						
LvpH	proteína visceral larval						
CG12780	ligante de bactéria		4				4
CG14748	sem <i>hit</i>						
CG8858	catabolismo de ferormônio						
CG33775	<i>desconhecida</i>						
CG8816	fosforilação de aminoácido						
CG17580	sem <i>hit</i>						
Or49b	receptor do olfato						
Pcf11	clivagem de RNA						
CG10249	sem <i>hit</i>	1					1
CG6040	sem <i>hit</i>	7					7
CG13978	crescimento celular	1					1
Gr98a	receptor gustativo						
Ase	desenvolvimento do sistema nervoso						
Exp6	proteína de transporte						
% χ^2		69,23	30,77				
			7,01*				
Desenvolvimento							
CG18869	atividade transferase						
CG32259	<i>desconhecida</i>						
CG5146	sem <i>hit</i>						
CG10592	atividade fosfatase						
HGTX	sistema nervoso			1	1		2
CG9628	sem <i>hit</i>		1				1
CG6959	proteína de ligação						
CG6962	sem <i>hit</i>						
CG6696	proteólise						
% χ^2			33,33	33,33	33,33		

N: número de genes.

* χ^2 : $p < 0,05$ entre o número de inserções 5' e 3'.

sem *hit*: não foi encontrado alinhamento significativo com nenhuma proteína depositada no *The Gene Ontology*

Apêndice C - Número total e porcentagem (%) de inserções de TEs em regiões intergênicas 5' e 3', em relação ao total de inserções, em 35 genes *Cyps* associados com resistência aos inseticidas (N: 9), com atividade monooxigenase (N: 20) e ao desenvolvimento (N: 6) em *D. simulans*.

<i>Cyp</i> /função	n° de inserções		Total
	5'	3'	
Resistência aos			
<i>Cyp6w1</i>	19	3	22
<i>Cyp6a2</i>	2	4	6
<i>Cyp4e2</i>			
<i>Cyp12d1</i>	2		2
<i>Cyp6a9</i>			
<i>Cyp6a8</i>			
<i>Cyp6g1</i>			
<i>Cyp12a4</i>			
<i>Cyp308a1</i>			
%	77	23	
χ^2		13,69*	
Atividade			
<i>Cyp4ad1</i>			
<i>Cyp6a14</i>	1	2	3
<i>Cyp6a13</i>			
<i>Cyp301a1</i>			
<i>Cyp9h1</i>		3	3
<i>Cyp6a22</i>	1		1
<i>Cyp6a17</i>			
<i>Cyp6a23</i>			
<i>Cyp6a19</i>			
<i>Cyp6a20</i>			
<i>Cyp6a21</i>			
<i>Cyp317a1</i>			
<i>Cyp4e1</i>			
<i>Cyp6a15Psi</i>			
<i>Cyp6g2</i>			
<i>Cyp6t3</i>			
<i>Cyp12c1</i>			
<i>Cyp12a5</i>			
<i>Cyp6a18</i>			
<i>Cyp4g1</i>			
%	29	71	
χ^2		8,76*	
Desenvolvimento			
<i>Cyp302a1</i>			
<i>Cyp307a1</i>			
<i>Cyp314a1</i>	2		2
<i>Cyp315a1</i>	1		1
<i>Cyp306a1</i>			
<i>Cyp18a1</i>			
%	100		
χ^2			

N: número de genes.

* χ^2 : $p < 0,05$ entre o número de inserções 5' e 3'.

Apêndice D - Número total e porcentagem (%) de inserções de TEs em regiões intergênicas e gênicas em relação ao total de inserções em 38 genes flanqueadores aos *Cyps* associados com resistência aos inseticidas (N: 12), com atividade monooxigenase (N: 17) e ao desenvolvimento (N: 9) em *D. simulans*.

Genes flanqueadores aos <i>Cyps</i>	Função	nº de inserções			Total
		5'	3'	intron	
Resistência aos inseticidas					
GD10333	sem <i>hit</i>				
GD10399	ligante da manose				
GD10301	fosforilação				
GD10299	fosforilação				
GD10562	atividade monooxigenase				
GD17879	<i>desconhecida</i>				
GD25944	metabolismo de resistência				
GD15241	atividade monooxigenase				
GD20134	sem <i>hit</i>				
GD15624	sem <i>hit</i>				
GD15623	metabolismo de resistência		2		2
GD15622	ligante de actina				
%			100		
χ^2					
Atividade monooxigenase					
GD10560	proteína da cutícula larval	1			1
GD10159	sem <i>hit</i>				
GD10574	atividade monooxigenase				
GD10156	sem <i>hit</i>	2			2
GD25831	<i>desconhecida</i>				
GD25830	sem <i>hit</i>				
GD10922	sem <i>hit</i>	1	3		4
GD10923	receptor do olfato				
GD25648	clivagem de RNA				
GD25644	metabolismo de resistência				
GD15274	proteína visceral larval				
GD15377	atividade monooxigenase				
GD12332	sem <i>hit</i>				
GD14765	transporte transmembrana				
GD19271	sem <i>hit</i>	10			10
GD21335	sem <i>hit</i>	2			2
GD21337	receptor gustativo				
%		84,21	15,79		
χ^2			22,72*		
Desenvolvimento					
GD13807	glucuronosiltransferase				
GD13298	sem <i>hit</i>				
GD13188	biossíntese de ecdisona				
GD13186	atividade fosfatase alcalina				
GD12614	desenvolvimento do sistema nervoso			2	2
GD14494	sem <i>hit</i>				
GD20627	sem <i>hit</i>				
GD20628	receptor ligante				
GD24857	biossíntese de ecdisona				
%				100	
χ^2					

N: número de genes.

* χ^2 : $p < 0,05$ entre o número de inserções 5' e 3'.

sem *hit*: não foi encontrado alinhamento significativo com nenhuma proteína depositada no *The Gene Ontology*

Apêndice E - Número e orientação sense (s) e antisense (as) de inserções de TEs em intervalos de 500pb até 3 kb das regiões flanqueadoras de *D. melanogaster*.

Distância das inserções	genes <i>Cyp5</i>						genes flanqueadores								
	5'			3'			5'			3'			Total		
	s	as	%	s	as	%	s	as	%	s	as	%	n	%	
Resistência															
0 - 500 pb	1	1	20	1	1	25	2	2	22,22						
500 - 1000 pb	1	1	40	2	2	50	4	4	44,44						
1000 - 2000 pb	1	1	20	1	1	25	2	2	22,22						
2000 - 3000 pb	1	1	20				1	1	11,11						
Subtotal	3	2	5	4	4	10	9								
Monooxigenase															
0 - 500 pb	1	1	40				2	2	40						
500 - 1000 pb	1	1	20				1	1	20						
1000 - 2000 pb	1	1	40				2	2	40						
2000 - 3000 pb	3	2	5				5	5	100						
Subtotal	3	2	5				5								
Desenvolvimento															
0 - 500 pb															
500 - 1000 pb															
1000 - 2000 pb															
2000 - 3000 pb															
Subtotal	6	4	10	4	4		14	14	100						
Total															

Apêndice F - Número e orientação sense (s) e antisense (as) de inserções de TEs em intervalos de 500pb até 3 kb das regiões flanqueadoras de *D. simulans*.

Distância das inserções	genes <i>Cyps</i>						genes flanqueadores									
	5'			3'			5'			3'			Total			
	s	as	n	%	s	as	n	%	s	as	n	%	s	as	n	%
Resistência																
0 - 500 pb	1	1	1	14,28	4	4	4	57,14	5	5	5	35,71				
500 - 1000 pb	1	1	1	14,28			1	7,14								
1000 - 2000 pb	3	2	5	71,42	2	1	3	42,85	8	8	8	57,14				
2000 - 3000 pb																
Subtotal	4	3	7	100	6	1	7	100	14	14	14	100				
Monooxigenase																
0 - 500 pb													1	1	1	100
500 - 1000 pb																
1000 - 2000 pb	1	1	1	100			1	100								
2000 - 3000 pb																
Subtotal	1	1	1	100			1	100	1	1	1	100	1	2	3	100
Desenvolvimento																
0 - 500 pb																
500 - 1000 pb																
1000 - 2000 pb																
2000 - 3000 pb																
Subtotal																
Total	4	4	8		6	1	7		15	15	15	100	1	1	2	100
Subtotal																
Total	4	4	8		6	1	7		15	15	15	100	1	2	3	100

Apêndice G - Classes e sentido de orientação sense (s) e antisense (as) do número de inserções de TEs (n) em regiões intergênicas de genes *Cyp* e de seus flaqueadores em *D. melanogaster*.

Classes	Inserções	5'				3'				Subtotal		Total	
		s	as	n	%	s	as	n	%	s	as	n	%
<i>Cyps</i>													
DNA	DNAREP1_DM	1	3	4	23,50	1	2	3	50	2	5	7	30,43
	PROTOP_A	2	1	3	17,65					2	1	3	13,04
	ISFUN1	1		1	5,88	1		1	16,67	2		2	8,70
	PROTOP_B		1	1	5,88						1	1	4,35
	PROTOP	1		1	5,88					1		1	4,35
	POGON1	1		1	5,88					1		1	4,35
	BARI_DM					1		1	16,67	1		1	4,35
	hAT-1_DP	1		1	5,88					1		1	4,35
LTR	MAX_LTR		2	2	11,76						2	8,70	
	MAX_I		1	1	5,88					1	1	4,35	
LINE	TART_DV		2	2	11,76						2	8,70	
	LINEJ1_DM					1		1	16,67	1		1	4,35
Subtotal		7	10	17	100	4	2	6		11	12	23	100
Genes flaqueadores													
DNA	DNAREP1_DM	1	2	3	33,33	5		5	83,33	6	2	8	53,33
	PROTOP_B		1	1	11,11						1	1	6,67
LTR	ROO_I		1	1	11,11						1	1	6,67
LINE	TART_DV	2	2	4	44,44	1		1	16,67	3	2	5	33,33
Subtotal		3	6	9	100	6		6	100	9	6	15	100

Apêndice H - Classes e sentido de orientação sense (s) e antisense (as) do número de inserções de TEs (n) em regiões intergênicas de genes *Cyp* e de seus flaqueadores em *D. simulans*.

Classes	Inserções	5'				3'				Subtotal		Total	
		s	as	n	%	s	as	n	%	s	as	n	%
<i>Cyps</i>													
DNA	DNAREP1_DM	8	5	13	46,43	1	1	2	16,67	9	6	15	37,50
	Helitron-1_Dyak	6	3	9	32,14	2	2	4	33,33	8	5	13	32,50
	Helitron-1_Dvir					2		2	16,67	2		2	5
	BARI_DM						2	2	16,67		2	2	5
	PROTOP_B	1	1	2	7,14					1	1	2	5
	FB4_DM		1	1	3,57						1	1	2,50
	TransibN1_DP		1	1	3,57						1	1	2,50
	BARI1					1		1	8,33	1		1	2,50
	HOBO		1	1	3,57						1	1	2,50
	Retroposon	MINIME_DN					1		1	8,33	1		1
Não classificado	DMRP1	1		1	3,57					1		1	2,50
Subtotal		1	12	28	100	7	5	12	100	23	17	40	100
Genes flaqueadores													
DNA	Helitron-1_Dyak	2	2	4	25,00					2	2	4	19,04
	DNAREP1_DM		2	2	12,50						2	2	9,52
	BARI_DM						2	2	40,00		2	2	9,52
	FB4_DM		1	1	6,25						1	1	4,76
	BARI1					1		1	20,00	1		1	4,76
LTR	Stalker2_LTR	1		1	6,25					1		1	4,76
	ROO_I	1		1	6,25					1		1	4,76
LINE	TART_DV	4	3	7	43,75		2	2	40,00	4	5	9	42,86
Subtotal		8	8	16	100	1	4	5	100	9	12		100

Apêndice I. Sequências das regiões flangeadoras dos genes *Cyps* amplificadas em linhagens resistentes e suscetíveis de *D. melanogaster* (Dm = LE: Lençóis/BA; ITA: Itaúnas/ES; RP: São José do Rio Preto/SP; BG: Bento Gonçalves/RS; TAV: Tavares/RS) e *D. simulans* (Ds = ITA: Itaúnas/ES; PF: Paulo de Faria/SP; AMI : AMIEU/Nova Caledônia; MAK: Makindu/Quênia;). Letras minúsculas representam a seqüência do TE inserido na região flangeadora do gene *Cyp*.

>6w1_5_DmLE

TACATACATATATACATACATATTTACAAAAGCAGCTTAAACATATGTAT
ATAATTTTGATATTAACATTTGTGACGACATATGATTTTCCATTTAACTA
TGTAATACAAAATaacaagaggtaatgctatagtcgagttcgccgaata
tcagatacccgttacccaactagtgtgaatgcaaacaggaaattttataa
ttttaatgggatattgatagatattgggaataaactgagaaaaaaat
taaattggtccaagtgaggggtgaccggtttagcggcttaagggcgta
gagcgtggcaaacgtttttaggaaaAATATCAAACATTTAACAT

>6w1_5_DmITA

TAACCAGAAAAAGCTTTTAAGCTGTATCTTCTTAGGTGATTNATCCNTAC
ATACATATATACATACATATTTACAAAAGCAGCTTAAACATATGTATATA
ATTTTGATATTAACATTTGTGACGACATATGATTTTCCATTTAACTATGT
AATACAAAATaacaagaggtaatgctatagtcgagttcgccgaatatca
gataccggttaccatctagtgtgaatgcaaacagaaattttataat
tgatgggatattgatagatattgggaataaactgagaaaaaaat
attggtcctaagtgaggggtgaccggtttagcggcttaagggcgtagag
cgtggcaaacgtttttaggaaaaatcaaacatttaacattttctta
tttatgatttgcttttagaatttctatttggtatgccaaaaatgttttgc
cacgcgcactctaacgctcTAACTA

>6w1_3'_DmRP

GNATTGGTTCATGATGAGACACCCTRGWAAA WYATWTYCGCSCCCAATA
TATAKGKATATAAACTTWAWTKTTTCAWTAWAWAWGTYTTGAMTAGAGG
AGTYTGCTAGAGCCAMYGGCTTTGAAAATTGTTACGATTTTAAATCGTT
TTATCAWTTTTCTTCTCATTTATATGGACACGCCMAAATTATTTTGATG
TTTCCGACTCTAACCTCATAGAATTGTCGTTTTAAAAAGTTGTTTATTTT
TTATTTTAAATAATTGAATTTAAATTTATAATTTTATTAATAAAAAAAAAAAT
TATTAATAATTTATAAATTTTATCAGGAGTGTATTAATAATATAGAATAGKG
AAAGATTAGGCGAGAARGAACCCCTTAGARRRACGTTARRACGTTARRAGC
AAATTTATTTGAGAAAGCTGGCAATCCCACAATTctgccccaaaactgca
taccacaaaaatggtctgatttcttttttatttwaactatttttttat
aatttaataacccaattggcaatwaaaaatggttgcaatttcgtgcatgc
actccaaactagctcagtaaccgggtgctctgcatagtcaggaatagc
gttttttATAACCTAATARATTAGTTWACGGCTCGGCTGTTTTACACTGA
CCAGGAWTAAATATACATCTTAAATKGTGCGAAAATGTTTCTCTAACTGG
AAAACATCTCGCCAATCGATTGATCCGYGTTCTTTGACTATTTCGTCATGC
TACAWTARRT

>6w1_5_DmBG

ACTTACATATCTAACTCACATTSKYMAYCYKYTKYACTMAMAAGACTTA
 GGAACACTATTAATCTTATAGTACATTAATATAACCAGAAAAAGCTTTT
 AAGCTGTATCTTCTTAGGTGCATACATACATACATATATACATACA
 TATTTACAAAAGCAGCTTAAACATATGTATATAATTTTGATATTAACAT
 TGTGACGACATATGATTTTCCATTTAACTATGTAATACAAAATaaacaag
 aggtaatgctatagtcgagttcgccgaatatcagataccggttaccacaac
 tagtgtgaatgcgaacaggaaatTTTataatTTTaaatgggatattgata
 gatattggggaataaaactgagaaaaaattTTTaaattggttctaagtgagg
 gggtagccgggttagcggcttaagggcgtagagcgtggcaaacgTTTT
 ttaggaaaaatatcaaacatttaacatattcttattttatgatttgcttt
 tagaatttctattggtatgccaaaaatgTTTTgccacgcgactctaacg
 ctCTAACTACGTCATATAAGCACTAAAGCATTAATAATTKGRRGCGTGCCA
 CRKRMGKTWGATAGATTGTCTTAGGTGGGTGCTAT

>6w1_5_DmTAV

ACGTAGTTAgagcgttagagtgcgcgtaggcaaacattTTTTggcatacca
 atagaaattctaaaagcaatcataaaataagaatatgTTaaatgTTTga
 tTTTTcttaaaaaacgTTTTgccacgctctaaccgcttaagccgcta
 aaccggttaccctcacttagaacaattTaaaattTTTTctcagTTta
 ttcccaatatctatcaatatccattaaaattataaaatttctgTTcg
 cattcacactagttgggtaacgggtatctgatattcggcggaactcgacta
 tagcattacctctgTTtATTTGTATTACATAGTTAAATGGAAAATCATA
 TGTCGTCACAA

> 6a2_3_DmRP

GTACACTAAACGAATAAAGTTCTGGAAATGAAAACAGTTTCCAGCTAAAA
 TAAGAAATATACTACTACGTTTGCCATTGAGAAATAAATAAAAACAAAAGTC
 AAGTGCTGACAACtatagttcttgttcgtgctaactagcccagacgatct
 agccatgcccgtctcgatgtctgtctgtccggtatgtccgtagtccggtt
 tgtccgtttgtccggtcgttcgtgtgaaagtcgagatatccgggaacat
 aaaagctagaaagttgagacttGACAACAGGTATTTAATAGAATCTAATA
 GCTGCCCTCAAGTATATTACTATCTTGTTTTGAATACACGAACAAAAGCA
 GTGTTATTTAATAAATTCAGCTACATTATCCGATATCACCGA

>12d1_5_DmBG (não houve inserção de TEs)

AGCGGCGCAGTGTGCATGGATATCTGCAGAAATTCGCCCTTAATCGCCGTG
 AATTGCCTACCTCACTCAACTGAAAAATAAACAAATGAACGTGACTGGC
 TAGAAAAACAAATTGAGAGCGTTGGAACCTTACAATAAATACACCTCAGAG
 AATTTGATACTACTATTACGACTATTACATACCCGTTACTCATTTAGTGG
 GATCGCTAGAAAAAATCTAAAATATTTAAATATTTTTTAAAAATTTATA
 AGGATTCAGTTCAACGGTAGACAATTTCAAGATAAATAACCAAATTTAAA
 GAAAAATAAAAACAATTCCTAGAGTTGATCCTTGTTCTCTTTGCATACTG
 ATATTTTATTTATTTTTTATTTTATTTTACTGGCGCCGCTTTTCCAC
 AGAGTCCCTCGGTGTCTTTATCATTTTCTGCGAGAACATTAAGATGTCAT
 CGAGAAGACGCATCATCATCCGAAAGTTTGGAGTGGAGATAAATCTCCAC
 ATTGATGGCTGTATATTCAATTGAAAGGCATATCTGAACATCGTGCTTCG
 TGTTTTGGAACACAGTCAATGCATCTGGATCGTCTCGATTTGACTGATC
 AAACCCATTGGGCGATCGAAAGCCACCAAGCCAAGGGATTGAAAACCG
 GCGGCATCTCCAGAGTTCTAGCATCTAAGGGCGAATTCAGCACACTGCG

>6a2_5_DsITA

GGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTCTTCGTTT
 TTATGTTTTGTTTTGCTAATTTCTATCGGTATAACAAAACACTTTGGCCCT
 cctttgaaaatttttaaaatttttttattatattcgccaaaatctaata
 tctatctatatacctagaaaatgacaaaaattttaagttcgcaatcctact
 agctaagtagtaacgggtatctgaaactcgaatattgtattctcttttggt
 tGTTATCAAATCCTTGACATGATTTTAAATATCTACCCATGTTTTGAGA
 TGAGGGAAAGATATACACAATTAAGTATTCACATTAATCGGGTTTTTAA
 AGAAGCCCTTATAAAGCCCAATAGTTAATACGGGTACATTATTGTAATGA
 TATCTCATCTGAATTACCAGCCCCACAATTGTGTGTGCTTAAAGCCCCGA
 TTAAGCTTACGGTTAGCCGTTCCCCAA

>6w1_5_DsPF

GACTAGACTATTAATCCATTATTAATCTATAATAAACTATTATTCTTATA
 ATACATTAATATAATCAGAAAAAGCTTTTAACTGCTCCTTATTGTGTC
 CTTAGGAAAAGAAGCAACATCAACATATATAATTTTGATATTAACTTTTG
 TGATGACATATGATTTAATATTTAACTAAGAAATACAAAATaaacaagag
 gtaatgctatagtggtttcctcgactatcagataccagttactctacta
 ttgtgaatgcaaatttcataaattgtatggatcgcgatagataytggga
 aataaaatgacggttttggcggttttagggcgttagagtgggcgtggcaa
 aaagtttattgaaaaaattttttgaaaaatctaaacatttttcaaaa
 atgtgggcgtggcagttttggcggtatagggcgttagagtggtcggtgca
 aaaagtatttggcaaatcgatagaaatttacaagactaattagattatg
 aaaaaatacccaaaaatggttaaaaactgttttaaaaatgttttaacatt
 ttcttattttattatttcccttagaatttctattgggtatgcccacatg
 ttttccctcaactcaacgttcTAACTACATATAAGCACAAAGCATTCAA
 ATTGGAGGCGCGTCACTGACGTTAGATAAGATAGGTCTAAGTTTGGGTGT
 CCCTAAACAGGTACCAAAATCCATCAGAGCATATttacacccgagactcg
 tagagtaaaaaggtattccagagtcggtgaaaagtaactggtagaaggaa
 gggttcccgcactatataaagtgtatataatttgaataggatcaatagccg
 agttgatctggccatgtccgt

>6a2_3_Ds AMI

ATGCTACATATATACAAGAAATTTTCAATTGATTAAATTTCTCGAATTGAT
 TCGTTGTTTGCCCTCAACTCAATAAGTTCGATAGTACTCACATTGCTAA
 ATATAAATTATTTCTCAAAATGAACTAGAGAGACTGCTTAGCATCATTGA
 CCGGAAACTTAAATCACTCAATCGTACACCAAACGAATAAAGTTCTGGAA
 ATGAAAACAGTTTCCAGCTAATAATAAGAAATAAACAACTACGTTTTGCC
 ATTGAGAATCATAACTTGTTCCTGAATAWAACAAAAGTCAAGTTCTGACT
 AACtatagttctgttctactaacaagcccaaacgatttagccatgccc
 gccttgatggtcgtctgtccgttatgtccgtatgtccgtatgtacgtatg
 tacgtatctccgtatatccgtatgtccgtatgtccgtatgTCCGTatgtc
 cgtatgtccttatgtccgtatgtccgtatgtccgtatgtccgtatgtcct
 tatgtccttatgtccatttgtccgtccgttcgtatgaaagtcgagatctc
 gggatccataaaagctagaaagttgagacttGACAAAAGGTATCTAATAA
 TCGCTGCACTCAAGTATGGTTATATCTTGTTTTTAAATTTAATAATTCAG
 CTACTTTATATGATATCACAAAAATATCTATGAT

>6a2_3_DsMAK

GTGACAGCCAGCCGAAAAAGTAGAATATAGATAAAAATGTATTTAAAGTTT
GAAATTTAWKSMTACATATATACAAGAAATTTTCAATTGATTAAATTCTC
GAATTGATTCGTTGTTTGCCTTCAACTTCAATAAGTTGATAGTACTCAT
ATTGCTAAATATAAATTATCTTCAGAATGAACTAGAGAAACTGCTTAGC
ATCATTGACCGGAAACTCAAATCACTCAATCGTACACCAAACGAATAAAG
TTCTGGAAATGAAAACAGTTTCCAGCTAAAATAAGAAATATTTGCCATAA
CTTTTGCTGAATAAAAACAAAAGTCAAGTTCTGACAACtatagttcttggt
catactaacaagcccaaacgatttaaccgtgcccgtctcgatgtccgtat
gtccgtatgtccgtatgtccgtatgtccgtatgtccatttgccgtccgt
tcgtatgaaagtcgagatctcgggaaccataaaaagctagaaagttgaaac
ttGACAAAAGGTATCTAATAGTCGCTGCACTAAAGTATCAATTATATCTT
GTTTTGAAATTTAATAATTCAGCTACTTTATATGATATCACAAAAAATAT
CTATGATAAGTGGGCGAACTCATGAGCGATAAGCCGTAAAAATTTCTCAT
ATCTGGATGTCTGGCGGAGGCATTGAC

3.1 INTRODUÇÃO

Classicamente, os elementos transponíveis, seqüências medianamente repetidas do genoma, foram considerados como parasitas genômicos (DOOLITTLE; SAPIENZA, 1980; ORGEL; CRICK, 1980; OHNO; YOMO, 1991) em decorrência de sua capacidade intrínseca de mobilização. Um corpo crescente de evidências mostra, entretanto, que os TEs têm exercido forte influência na trajetória evolutiva dos organismos, por desempenharem um papel chave na evolução da estrutura, função e regulação gênica dos eucariotos (BROSIUS, 1999; MAKALOWSKI, 2000; 2003; KIDWELL, 2002; LORENC; MAKALOWSKI, 2003; KIDWELL; HOLYOAKE, 2001; FONTANILLAS et al., 2007; FESCHOTTE, 2008), como também para a evolução do tamanho dos genomas (PETROV 2001; KIDWELL, 2002).

Dentre os estudos que sugerem a participação dos TEs na regulação dos genes (e.g. FESCHOTTE, 2008), destacam-se alguns relacionados com a evolução da resistência aos inseticidas. Tem sido demonstrado que a presença de inserções de TEs em genes *Cyp*s, codificadores da família multigênica citocromo P450 monooxigenase, estimula a expressão de genes responsáveis pela resistência metabólica em *Drosophila* (DABORN et al., 2002; CATANIA et al., 2004; SCHLENKE; BEGUN, 2004; BOGWITZ et al., 2005; MARSANO et al., 2005; CHUNG et al. 2007). A elevada expressão dos genes *Cyp*s associados a detoxificação de xenobióticos devido à inserção desses TEs, pode ser vantajosa por caracterizar a resistência aos inseticidas ao hospedeiro. Análises *in silico* também têm sugerido uma vantagem seletiva da inserção de TEs nas proximidades dos genes *Cyp*s. Chen e Li (2007) mostraram que genes *Cyp*s de *Drosophila melanogaster* associados à resistência aos inseticidas possuem inserções de TEs ao contrário daqueles associados ao desenvolvimento dos insetos, como por exemplo, os *Cyp*s associados à biossíntese de ecdisona. Esses resultados sugerem que genes envolvidos em respostas ambientais tendem a ser mais tolerantes a mudanças genômicas causadas pela inserção de TEs, e assim, apresentarem maior número de inserções, enquanto que genes envolvidos em processos biológicos, como o desenvolvimento do organismo, são altamente conservados e, portanto, as inserções são seletivamente eliminadas.

As espécies de *Drosophila* apresentam proporções bastante variáveis de TEs, como ilustrado na Tabela 1. O grupo *melanogaster*, particularmente, apresenta as variações mais

extremas: *D. ananassae*, por exemplo, apresenta a maior proporção (24,03 %), enquanto *D. simulans* apresenta a menor (2,73 %). Essa variação tende a estar diretamente relacionada ao tamanho dos genomas: espécies com os maiores genomas são as que apresentam a maior porcentagem de TEs e vice e versa. Com exceção de *D. erecta*, essa relação ocorre nas espécies do grupo *melanogaster* (retângulo).

Tabela 1. Proporções de elementos de transposição em regiões eucromáticas de espécies de *Drosophila*.

Grupos de <i>Drosophila</i>	Espécies	Tamanho do Genoma Mb*	TEs (%)
<i>melanogaster</i>	<i>D. melanogaster</i>	118	5,35
	<i>D. simulans</i>	111	2,73
	<i>D. sechellia</i>	115	3,67
	<i>D. yakuba</i>	127	12,04
	<i>D. erecta</i>	134	6,97
	<i>D.annanassae</i>	176	24,93
<i>obscura</i>	<i>D. pseudoobscura</i>	127	2,76
	<i>D. persimilis</i>	138	8,47
<i>saltans</i>	<i>D. willistoni</i>	187	15,57
<i>repleta</i>	<i>D. virilis</i>	172	13,96
	<i>D. mojavensis</i>	161	8,92
<i>hawaiianas</i>	<i>D. grimshawi</i>	138	2,84

* Clark et al., 2007; em negrito as maiores e as menores proporções para o grupo *melanogaster*.

Duas questões podem ser formuladas em relação à inserção preferencial de TEs em genes *Cyps* associados à resistência aos inseticidas. A primeira é se essa é uma particularidade da espécie *D. melanogaster*, ou uma propriedade que se manteria para outras espécies do gênero *Drosophila*. A segunda é se as variações acentuadas do número de TEs que se observa nas diferentes espécies estariam refletidas na inserção preferencial de TEs nesses *Cyps*. Para responder a essas questões, este trabalho teve por objetivo analisar nas espécies do grupo *melanogaster* (*D. melanogaster*, *D. simulans*, *D. sechellia*, *D. yakuba*, *D. erecta* e *D. ananassae*), os 13 genes *Cyps* estudados em *D. melanogaster* (CHEN; LI, 2007), com ênfase nas suas regiões flangeadoras 5' e 3', próximas em até 3 kb do sítio de início da transcrição de cada gene.

3.2 OBJETIVOS

Os objetivos específicos deste estudo foram:

1. Verificar se há inserção preferencial de TEs em genes *Cyps* associados à resistência aos inseticidas nas espécies do grupo *melanogaster* que possuem o genoma seqüenciado;
2. Analisar se há inserção preferencial de TEs nas regiões intergênicas distantes em até 3 kb do sítio de início da transcrição dos genes *Cyps* (região onde se localizam preferencialmente os promotores) do que nas regiões gênicas;
3. Verificar se as espécies que possuem maior proporção de TEs possuem proporcionalmente maior número de inserções nas regiões flangeadoras dos *Cyps*.

3.3 MATERIAL E MÉTODOS

3.3.1 Análises de Bioinformática

Neste estudo foram analisados os genomas seqüenciados das seis espécies do grupo *melanogaster* de *Drosophila*. O genoma de *D. melanogaster* está organizado em cromossomos e o de *D. simulans* foi montado como um *assembly* mosaico correspondente a diferentes linhagens geográficas. As seqüências dos braços cromossomais 2L, 2R, 3L, 3R, 4, X e a parte não agrupada (chamada de U) correspondem à versão do genoma de *D. simulans* disponível no sítio *ftp* do *Genome Sequencing Center* da Washington University Medical School (ftp://genome.wustl.edu/pub/organism/Invertebrates/Drosophila_simulans_mosaic/assembly/Drosophila_simulans_mosaic-1.0). Os genomas das outras quatro espécies (*D. sechellia*, *D. yakuba*, *D. erecta* e *D. ananassae*) foram seqüenciados, mas, ainda não estão organizados em cromossomos. As versões dos genomas utilizados e o número de genes de cada espécie estão indicados na Tabela 2.

Tabela 2. Número de genes e genomas utilizados para a análise de inserções de TEs nas espécies do grupo *melanogaster*.

Espécies	Número de genes ¹	Genomas ²
<i>D. melanogaster</i>	13.733	ftp://ftp.flybase.net/releases/FB2008_01/dmel_r5.5/fasta
<i>D. simulans</i>	15.983	ftp://ftp.flybase.net/releases/FB2008_01/dsim_r1.0/fasta/
<i>D. sechellia</i>	16.884	ftp://ftp.flybase.net/releases/FB2008_01/dsec_r1.0/fasta/
<i>D. yakuba</i>	16.423	ftp://ftp.flybase.net/releases/FB2008_01/dyak_r1.0/fasta/
<i>D. erecta</i>	15.324	ftp://ftp.flybase.net/releases/FB2008_01/dere_r1.0/fasta/
<i>D. ananassae</i>	15.276	ftp://ftp.flybase.net/releases/FB2008_01/dana_r1.0/fasta/

¹Clarck et al. (2007).

²ftp://ftp.flybase.net/releases/FB2008_01/

As seqüências disponibilizadas nestes *releases* foram analisadas quanto à ocorrência de fragmentos de elementos de transposição nas regiões intergênicas 5' e 3', distante em até 3 kb do sítio de início da transcrição do gene, bem como nas seqüências de 13 genes da família *Cyp* nestas espécies. Dentre os 90 genes *Cyps* identificados em *D. melanogaster* (TIJET et al., 2001), 13 estão bem caracterizados quanto a sua função (CHEN; LI, 2007), oito deles estando envolvidos no metabolismo de inseticidas e cinco na biossíntese de ecdisona (Tabela 3).

Tabela 3. Genes *Cyps* analisados.

Genes	Função	
<i>Cyp6w1</i>	Resistência	aos
<i>Cyp6a2</i>	Resistência	aos
<i>Cyp4e2</i>	Resistência	aos
<i>Cyp12d1</i>	Resistência	aos
<i>Cyp6g1</i>	Resistência	aos
<i>Cyp6a9</i>	Resistência	aos
<i>Cyp6a8</i>	Resistência	aos
<i>Cyp12a4</i>	Resistência	aos
<i>Cyp302a1</i>	Biossíntese de ecdisona	
<i>Cyp307a1</i>	Biossíntese de ecdisona	
<i>Cyp314a1</i>	Biossíntese de ecdisona	
<i>Cyp315a1</i>	Biossíntese de ecdisona	
<i>Cyp306a1</i>	Biossíntese de ecdisona	

Modelos gênicos estão disponíveis para as espécies do grupo *melanogaster*, entretanto, apenas o genoma de *D. melanogaster* apresenta seus genes extensivamente anotados e depositados no *LocusLink* (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi.db=gene>). Assim, a identificação dos genes da família *Cyp* nos genomas de *D. simulans*, *D. sechellia*, *D. yakuba*, *D. erecta* e *D. ananassae* foi obtida por meio das coordenadas genômicas em *D. melanogaster* e por meio de comparações por meio de *BLASTN* (ALTSCHUL et al. 1990) entre as seqüências completas dos genes *Cyps* de *D. melanogaster* contra os genes preditos dos demais genomas. Em *D. ananassae*, uma espécie mais divergente de *D. melanogaster* foi necessário realizar um *BLASTX* para obtenção de um resultado mais fidedigno. Após levantamento da localização

genômica, foram extraídas as seqüências de cada gene *Cyp*, bem como suas regiões flanqueadoras 5' e 3' distantes em até 3 kb do sítio de iniciação da transcrição.

A identificação de elementos de transposição foi realizada nas seqüências extraídas por meio do programa *RepeatMasker* (RM) versão 3.1.9 (<http://www.repeatmasker.org>), com a ferramenta de alinhamento *Cross_match* versão 0.990329 (pacote Phrap/cross_match/swat: http://www.phrap.org/_phredphrapconsed.html) contra o banco de seqüências de TEs de referência do gênero *Drosophila* oriundas do *Repbase* (JURKA et al., 2005). Para eliminar resultados espúrios, somente escores de *matches* RM ≥ 225 foram aceitos, condições de procura de maior sensibilidade/menor velocidade (condição “-s”), matriz relativa ao conteúdo GC do *query* (“-gccalc”) e seqüências de DNA de baixa complexidade não foram mascaradas (“-nolow”). Foram considerados todos os alinhamentos registrados pelo *RepeatMasker* para análises posteriores sem imposição de escores adicionais e comprimento mínimo. O teste estatístico aplicado para saber se houve diferença no número de inserções entre as regiões 5' e 3', foi o qui-quadrado (χ^2).

3.3.2 Análises Evolutivas

Foram realizadas análises de distância (p), como implementado no programa DnaSP4 (ROZAS et al., 2003) entre as seqüências dos elementos DNAREP1_DM e das relações evolutivas entre as cópias de diferentes genes e espécies por meio de *neighbor-joining* (NJ), como implementado no programa MEGA 4 (TAMURA et al., 2007).

3.4 RESULTADOS

Como as seqüências de TEs podem possuir sinais reguladores da transcrição, os quais podem alterar a expressão gênica, foi realizada uma busca por inserções de TEs em genes *Cyps* associados à resistência aos inseticidas e *Cyps* associados ao desenvolvimento, para verificar se havia inserção diferencial entre *Cyps* com funções distintas.

3.4.1 Ocorrência de inserções de TEs em genes *Cyps* e em suas regiões flanqueadoras 5' e 3' distantes até 3 kb do início da transcrição do gene

Dentre os *Cyps* associados à resistência, o número médio de inserções por gene, bem como, seus tamanhos médios e taxa de divergência média são apresentados na Tabela 4. As maiores médias de inserções por gene ocorreram em *D. yakuba* (4,75) e *D. simulans* (4,67), seguidas por *D. sechellia* (3,33), *D. erecta* (2) e *D. melanogaster* (1,80). A espécie com menor média de inserções foi *D. ananassae*, com apenas um TE por gene, sendo a espécie com o maior genoma dentre as do grupo *melanogaster*. Por outro lado, a maioria dos *Cyps* associados ao desenvolvimento não apresentou inserções, apenas *D. yakuba* e *D. erecta* tiveram uma inserção respectivamente nos genes *Cyp306a1* e *Cyp302a1*.

O tamanho médio dos fragmentos de TEs inseridos nos *Cyps* associados à resistência, variou 97 pb (*D. ananassae*) a 1.093 pb (*D. melanogaster*) e, nas demais espécies, entre 110 e 222 pb. Os *Cyps* associados ao desenvolvimento que continham inserções apresentaram fragmentos muito pequenos, com 62 pb (*Cyp302a1* de *D. erecta*) e 90 pb (*Cyp306a1* de *D. yakuba*). Os fragmentos de TEs presentes nos *Cyps* associados à resistência apresentaram a maior porcentagem de divergência em *D. ananassae* (21,04%), seguido por *D. sechellia* (19,38%), *D. erecta* (18,07%), *D. simulans* (16,35%) e *D. yakuba* (13,76%). Quanto à cópia inserida no *Cyp* associado ao desenvolvimento de *D. yakuba* apresentou 24,71% de divergência e a cópia presente em *D. erecta*, 16,13%.

Com exceção de *D. erecta*, que apresentou duas inserções na região gênica do *Cyp12a4* (intron) e uma inserção no *Cyp302a1* (intron), as outras cinco espécies apresentaram inserções apenas nas regiões flanqueadoras (Tabelas 5 a 10). *D. yakuba* foi a espécie com maior número de inserções (20; Tabela 8), enquanto que *D. ananassae* foi a espécie com o menor número (2; Tabela 10). As Tabelas 5 a 10 e a Figura 1 mostram que as inserções de TEs nas regiões flanqueadoras estão principalmente localizadas na região 5' dos genes de *D. melanogaster* (seis), de *D. yakuba* (12) e de *D. erecta* (cinco). Por outro lado, em *D. sechellia*, foi observado o maior número de inserções na região 3' (seis) do que na região 5' (quatro) dos *Cyps*. Em *D. simulans* e *D. ananassae*, o número de inserções foi o mesmo para ambas regiões, sendo sete na primeira espécie e uma na segunda. Porém, essas diferenças entre o número de inserções nas regiões 5' e 3' não foram significativas: *D. melanogaster*, *D. sechellia* e *D. yakuba* (χ^2 : 1,80; $p > 0,05$) e *D. erecta* (χ^2 : 0,51; $p > 0,05$).

A Figura 1 mostra a distribuição dos elementos nas regiões flanqueadoras 5' e 3' dos genes *Cyps* em cada espécie do grupo *melanogaster*. *D. melanogaster* apresentou dez inserções, das quais, cinco, foram do elemento DNAREP1_DM, três cópias localizadas na região 5' e duas na 3'. O elemento ISFUN1 apresentou duas cópias inseridas na região 3' dos genes *Cyps*. Os elementos MAX_LTR e MAX_I tiveram apenas uma inserção na região flanqueadora 5', enquanto o BARI_DM também apresentou uma inserção, porém, na região 3'.

Do total de 14 inserções de TEs encontradas nos *Cyps* de *D. simulans*, seis foram do Helitron-1_Dyak, estando três na região 5' e três na região 3'. Das cinco cópias do elemento DNAREP1_DM, quatro se localizavam na região 5' e uma na região 3'. Apenas uma cópia dos elementos Helitron-1_Dvir, Helitron-1N1_Dvir e MINIME_DN foram encontradas na região flanqueadora 3' dos *Cyps*.

D. sechellia apresentou dez inserções, sendo seis do elemento Helitron-1_Dyak, estando três cópias em cada região flanqueadora. Apenas duas cópias do TE DNAREP1_DM foram encontradas nessa espécie, estando uma em cada região flanqueadora. Uma cópia do elemento Helitron-1_Dvir e do Helitron-2_Dvir foram localizadas na região 3'.

Em *D. yakuba*, a espécie com maior número de inserções de TEs (20), dentre as espécies analisadas, metade das inserções (10) foram do DNAREP1_DM, distribuídas igualmente nas duas regiões flanqueadoras. Dentre os elementos restantes, sete foram inserções do TE Helitron-

1_Dyak, localizando-se quatro na região 5' e três na região 3'. Outros três elementos (PROTOP_B, FB4_DM e hAT-1_DP) apresentaram apenas uma cópia na região flanqueadora 5'.

D. erecta apresentou 12 inserções de TEs nas regiões flanqueadoras dos genes *Cyps*. Dessas inserções, cinco foram do DNAREP1_DM, estando quatro localizadas na região 5' e uma na região intrônica. O TE Helitron-1N1_Dvir apresentou duas inserções, estando cada uma em uma das regiões flanqueadoras, e o Helitron-1_Dyak apenas uma inserção na região 3' e duas em região intrônica. Foram ainda encontrados os elementos MINIME_DN e TART_DV, com uma cópia de cada na região flanqueadora 3'.

A espécie com o menor número de TEs nos genes *Cyps*, dentre as analisadas neste estudo, foi *D. ananassae*, com apenas duas inserções, um fragmento do Helitron-2_Dvi, presente na região flanqueadora 5', e um do elemento TART_DV, localizado na região 3'.

Tabela 4. Valores médios (mínimos e máximos) do número, tamanho e divergência dos TEs inseridos em genes *Cyps*.

Espécies	<i>Cyps</i> de resistência (N = 8)			<i>Cyps</i> de desenvolvimento (N = 5)		
	número (< >)	tamanho (< >)	divergência (%) (< >)	número (< >)	tamanho (< >)	divergência (%) (< >)
<i>D. melanogaster</i>	1,80 (1; 3)	1093 (31; 7914)	10,79 (0; 26,67)	0	0	0
<i>D. simulans</i>	4,67 (2; 6)	156 (36; 392)	16,35 (0; 25,86)	0	0	0
<i>D. sechellia</i>	3,33 (1; 7)	160 (31; 317)	19,38 (9,81; 26,85)	0	0	0
<i>D. yakuba</i>	4,75 (2; 6)	222 (38; 623)	13,76 (5,02; 24,13)	1 (1)	90 (90)	24,71 (24,71)
<i>D. erecta</i>	2 (1; 4)	110 (46; 194)	18,07 (10,76; 23,81)	1 (1)	62 (62)	16,13 (16,13)
<i>D. ananassae</i>	1 (1)	97 (77; 116)	21,04 (18,18; 23,90)	0	0	0

Tabela 5. Inserções de TEs em genes *Cyps* de *D. melanogaster* e em suas regiões flaqueadoras 5' e 3'.

Genes tamanho (orientação)	Região	Posição na região flaqueadora		Família	score RM	tamanho (pb/ % no TE consenso)	TE				Tamanho do TE consenso (pb)
		(pb)	orientação (+/-)				divergência (%)	deleção (%)	inserção (%)	Posição no TE consenso (pb)	
Cyp6w1 1747 (+)	5'	1593 - 1667 (+)		DNA/Helitron/DNAREP1_DM	243	75 (12,63 %)	26.67	1.33	0	6 - 81	594
		3293 - 3506 (-)		DNA/Helitron/DNAREP1_DM	1154	220 (37,03 %)	12.62	2.80	0	594 - 375	
	gene	3518 - 3604 (+)		DNA/Helitron/DNAREP1_DM	254	87 (14,65 %)	18.82	2.30	2.30	330 - 416	
Cyp6a2 1741 (-)	3'	1724 - 1981 (+)		DNA/Helitron/DNAREP1_DM	361	400 (67,34 %)	26.10	11.40	5.7	188 - 587	594
	5'										
Cyp4e2 2460 (+)	gene	17052 - 17216 (+)		LINE/Penelope/ISFUN1	228/314	163 (17,56 %)	21.47	2.45	4.29	87 - 208	928
	3'	17163 - 17225 (+)		DNA/Helitron/DNAREP1_DM	370	62 (10,44 %)	16.13	0	1.59	150 - 211	594
	gene	1196 - 1226 (+)		LINE/Penelope/ISFUN1	239	31 (3,34 %)	0	0	0	76 - 106	928
Cyp6a8 1842 (-)	5'	255 - 575 (-)		LTR/Pao/MAX_LTR	2718	321 (100 %)	1.25	0	0	321 - 1	321
	gene	576 - 8483 (-)		LTR/Pao/MAX_I	68921	7914 (100 %)	0.18	0.19	0.11	7914 - 1	7914
	3'	1 - 960 (+)		DNA/Tc1/BARI_DM	8946	960 (55,55 %)	0	0	0	17 - 976	1728

+: sentido sense; -: sentido antisense.

Tabela 6. Inserções de TEs em genes *Cyps* de *D. simulans* e em suas regiões flanqueadoras 5' e 3'.

Genes tamanho (orientação)	Região	Posição na região flanqueadora		Família	score RM	tamanho (pb)/% no TE consenso)	TE					Tamanho do TE consenso (pb)
		(pb)	orientação (+/-)				divergência (%)	deleção (%)	inserção (%)	Posição no TE consenso (pb)		
Cyp6w1 1660 (+)	5'	19610 - 19729 (-)		DNA/Helitron/DNAREPI_DM	746	126 (21,21 %)	10.83	5.00	0	594 - 469	594	
		19727 - 19967 (-)		DNA/Helitron/Helitron-1_Dyak	1254	251 (2,31 %)	13.75	4.56	0.41	564 - 314	10846	
		19967 - 20357 (+)		DNA/Helitron/Helitron-1_Dyak	23/1089	292 (2,70 %)	18.15	4.79	0	1 - 415		
gene 3'	1348 - 1412 (+)		DNA/Helitron/Helitron-1_Dyak	355	65 (0,6 %)	16.92	0	0	92 - 156	10846		
	1800 - 1949 (+)		DNA/Helitron/DNAREPI_DM	405	158 (26,6 %)	21.24	8.00	2.67	411 - 568	594		
	1996 - 2088 (-)		DNA/Helitron/Helitron-1_Dyak	281	90 (0,83 %)	25.56	0	3.23	157 - 68	10846		
Cyp6a2 1590 (-)	5'	1351 - 1810 (+)		DNA/Helitron/Helitron-1_Dyak	1803	392 (3,61 %)	12.89	2.61	17.39	1 - 392	10846	
		1817 - 1967 (+)		DNA/Helitron/DNAREPI_DM	527	152 (25,6 %)	15.71	7.95	7.28	443 - 594	594	
gene 3'	1273 - 1379 (+)		DNA/Helitron/Helitron-1_Dvir	339	105 (1,19 %)	24.03	0.93	2.80	108 - 212	8816		
	1293 - 1551 (+)		MINIME_DN	398/400	248 (20,36 %)	23.39	0.81	3.23	319 - 475	1218		
	1431 - 1550 (+)		DNA/Helitron/Helitron-1N1_Dvir	452	116 (1,32 %)	25.86	0	3.33	148 - 263			
	1471 - 1547 (+)		DNA/Helitron/Helitron-1_Dyak	456	77 (0,71 %)	16.88	0	0	135 - 211	10846		
Cyp12d1 1741 (-)	5'	373 - 408 (-)		DNA/Helitron/DNAREPI_DM	266	36 (6,06 %)	8.33	0	0	565 - 530	594	
		740 - 821 (+)		DNA/Helitron/DNAREPI_DM	316	82 (13,80 %)	19.23	3.66	4.88	482 - 562		
gene 3'												

+: sentido sense; -: sentido antisense.

Tabela 7. Inserções de TEs em genes *Cyps* de *D. sechellia* e em suas regiões flanqueadoras 5' e 3'.

Genes tamanho (orientação)	Posição na		TE									
	Região flanqueadora (pb)	região flanqueadora (pb)	Família	score RM	tamanho (pb/ % no TE consenso)	divergência (%)	deleção (%)	inserção (%)	Posição no TE consenso (pb)	Tamanho do TE consenso (pb)		
Cyp6w1 1660 (+)	5'	1239-1555 (-)	DNA/Helitron/DNAREP1_DM	1958	317 (53,36%)	9.81	5.05	0.32	594 - 263	594		
		1318-1634 (-)	DNA/Helitron/Helitron-1_Dyak	1425	317 (2,92%)	16.72	5.36	0	647 - 314	10846		
		1839-2024 (-)	DNA/Helitron/Helitron-1_Dyak	1164	186 (1,71%)	13.44	4.84	0	1 - 195	10846		
gene 3'	1344-1409 (+)	DNA/Helitron/Helitron-1_Dyak	357	66 (0,60%)	18.18	0	0	0	91 - 156	10846		
	1790-1920 (+)	DNA/Helitron/DNAREP1_DM	306	131 (22,05%)	19.35	19.08	5.34	442 - 590	594			
	1914-2025 (-)	DNA/Helitron/Helitron-2_Dvir	271	112 (1,22%)	26.85	4.46	3.57	169 - 57	9141			
	1940-2025 (-)	DNA/Helitron/Helitron-1_Dyak	288	86 (0,79%)	22.89	5.81	3.49	155 - 68	10846			
Cyp6a2 1590 (-)	5' gene	2147-2281 (+)	DNA/Helitron/Helitron-1_Dvir	524	135 (1,53%)	24.06	3.70	1.48	108 - 245	8816		
		2183-2278 (+)	DNA/Helitron/Helitron-1_Dyak	531	96 (0,88%)	19.79	2.08	0	116 - 213	10846		
Cyp4e2 2238 (+)	5' gene	1978-2130 (+)	DNA/Helitron/Helitron-1_Dyak	437	153 (1,41%)	22.66	25.49	1.96	122 - 310	10846		
		3'										

+: sentido sense; -: sentido antisense.

Tabela 8. Inserções de TEs em genes *Cyps* de *D. yakuba* e em suas regiões flanqueadoras 5' e 3'.

Genes tamanho (orientação)	Posição na		TE									
	Região	região flanqueadora (pb)	Família	score RM	tamanho pb/ % no TE consenso)	divergência (%)	deleção (%)	inserção (%)	Posição no TE consenso (pb)	Tamanho do TE consenso (pb)		
Cyp6w1 1659 (-)	5'	729-766 (-)	DNA/Helitron/DNAREP1_DM	269	38 (6,39 %)	7.89	2.63	0	594 - 556	594		
		858-911 (-)	DNA/Helitron/Helitron-1_Dyak	310	54 (0,49 %)	9.62	0	3.70	559 - 508	10846		
		888-941 (-)	DNA/P/PROTOP_B	329	54 (4,68 %)	9.62	0	3.70	52 - 1	1153		
		2227-2849 (-)	DNA/Tc1/FB4_DM	1924	623 (15,23 %)	21.12	14.13	1.93	2755 - 2057	4089		
gene	3'	381-591 (-)	DNA/Helitron/DNAREP1_DM	561	211 (35,52 %)	20.59	16.59	3.32	362 - 124	594		
		1626-1899 (-)	DNA/Helitron/Helitron-1_Dyak	921	274 (2,52 %)	15.65	17.88	4.38	591 - 281	10846		
Cyp6a2 1592 (-)	5'	957-1092 (-)	DNA/Helitron/DNAREP1_DM	512	136 (22,89 %)	18.65	2.21	1.47	590 - 454	594		
		1089-1404 (-)	DNA/Helitron/DNAREP1_DM	663	316 (53,19 %)	24.13	5.38	9.49	594 - 292	10846		
		1173-1404 (-)	DNA/Helitron/Helitron-1_Dyak	1168	232 (2,13 %)	11.74	17.67	0.86	564 - 294	594		
		1738-1882 (+)	DNA/Helitron/DNAREP1_DM	800	145 (24,41 %)	15.17	5.52	0	113 - 265	10846		
		2356-2558 (-)	DNA/Helitron/DNAREP1_DM	481-876	200 (33,67 %)	9.17	7	1.41	594-525/309-158	594		
		2578-2670 (-)	DNA/Helitron/Helitron-1_Dyak	686	93 (0,85 %)	8.60	0	0	94 - 2	10846		
Cyp4e2 2228 (+)	5'	1094-1310 (-)	DNA/Helitron/DNAREP1_DM	618	217 (36,53 %)	21.53	3.23	10.14	593 - 392	594		
		1306-1401 (+)	DNA/Helitron/Helitron-1_Dyak	493	96 (0,88 %)	12.63	1.04	1.04	304 - 399	10846		
		1445-1862 (+)	DNA/Helitron/Helitron-1_Dyak	3435	418 (3,85 %)	5.02	3.35	0	1 - 432	10846		
		1882-1953 (+)	DNA/Helitron/DNAREP1_DM	562	72 (12,12 %)	8.33	0	0	523 - 594	594		
		1968-2102 (+)	DNA/Helitron/DNAREP1_DM	578	135 (22,72 %)	15.32	0.74	8.15	470 - 594	10846		
gene	3'	1850-2288 (+)	DNA/Helitron/Helitron-1_Dyak	398-1625	439 (4,04 %)	9.31	10.48	1.66	1-49/105-564	10846		
		1898-2355 (+)	DNA/Helitron/DNAREP1_DM	1407	458 (77,10 %)	17.35	13.32	9.39	119 - 594	594		
Cyp306a1 2244 (+)	5'	2779-2868 (+)	DNA/hAT/hAT-1_DP	238	90 (16,15 %)	24.71	8.89	1.11	2478 - 2574	557		
		gene	3'									

+: sentido sense; -: sentido antisense.

Tabela 9. Inserções de TEs em genes *Cyps* de *D. erecta* e em suas regiões flaqueadoras 5' e 3'.

Genes tamanho (orientação)	Posição na		Região flaqueadora (pb) orientação (+/-)	Família	score RM	tamanho (pb/ % no TE consenso)	divergência (%)	deleção (%)	inserção (%)	Posição no TE consenso (pb)	Tamanho do TE consenso (pb)
	região	TE									
Cyp6a2 1589 (+)	5'										
	gene										
	3'	757-931 (-)	DNA/Helitron/Helitron-IN1_Dvir	457	175 (33,71 %)	22.82	5.14	14.86		265 - 108	519
		760-835 (-)	DNA/Helitron/Helitron-1_Dyak	466	76 (0,70 %)	19.74	0	0		213 - 138	10846
Cyp4e2 2212 (-)		770-909 (-)	MINIME_DN	529	140 (11,49 %)	22.96	0	3.57		464 - 330	1218
		2021-2083 (+)	LINE/TART_DV	242	63 (0,41 %)	23.81	0	0		6982 - 7044	15099
	5'	1190-1383 (+)	DNA/Helitron/DNAREPI_DM	1072	194 (32,65 %)	10.76	5.13	4.62		396 - 591	594
	gene										
Cyp12d1 1743 (+)	5'	2528-2573 (-)	DNA/Helitron/DNAREPI_DM	236	46 (7,74 %)	17.39	2.17	0		575 - 529	594
	gene										
	3'										
		2184-2295 (+)	DNA/Helitron/DNAREPI_DM	703	112 (18,85 %)	11.61	0.89	0		1 - 113	594
Cyp6a9 1566 (+)	5'	2311-2387 (+)	DNA/Helitron/DNAREPI_DM	392	77 (12,96 %)	15.50	0	7.79		523 - 593	
	gene										
	3'										
		1683-1767 (-)	DNA/Helitron/DNAREPI_DM	393	85 (14,30 %)	17.65	5.88	0		590 - 501	594
Cyp12a4 2042 (-)	5'	1799-1851 (-)	DNA/Helitron/Helitron-1_Dyak	449	53 (0,48 %)	0	0	0		53 - 1	10846
	gene										
	3'										
		732-785 (+)	DNA/Helitron/Helitron-1_Dyak	400	54 (9,09 %)	5.56	0	0		2 - 55	10846
Cyp302a1 1887 (+)	5'	72-133 (+)	DNA/Helitron/Helitron-IN1_Dvir	288	62 (11,94 %)	16.13	1.61	0		146 - 208	519
	gene										
	3'										

+: sentido sense; -: sentido antisense.

Tabela 10. Inserções de TEs em genes *Cyps* de *D. ananassae* e em suas regiões flanqueadoras 5' e 3'.

Genes tamanho (orientação)	Posição na		TE							
	Região	região flanqueadora (pb) orientação (+/-)	Família	score RM	tamanho (pb/ % no TE consenso)	divergência (%)	deleção (%)	inserção (%)	Posição no TE consenso (pb)	Tamanho do TE consenso (pb)
Cyp6w1 2131 (+)	5' gene 3'	1927-2042 (-)	DNA/Helitron/Helitron-2_Dvir	311	116 (1,26 %)	23.90	4.31	2.59	170 - 53	9141
Cyp12a4 1954 (-)	5' gene 3'	147-223 (+)	LINE/TART_DV	266	77 (0,50 %)	18.18	3.90	0	11459 - 11538	15099

+: sentido sense; -: sentido antisense.

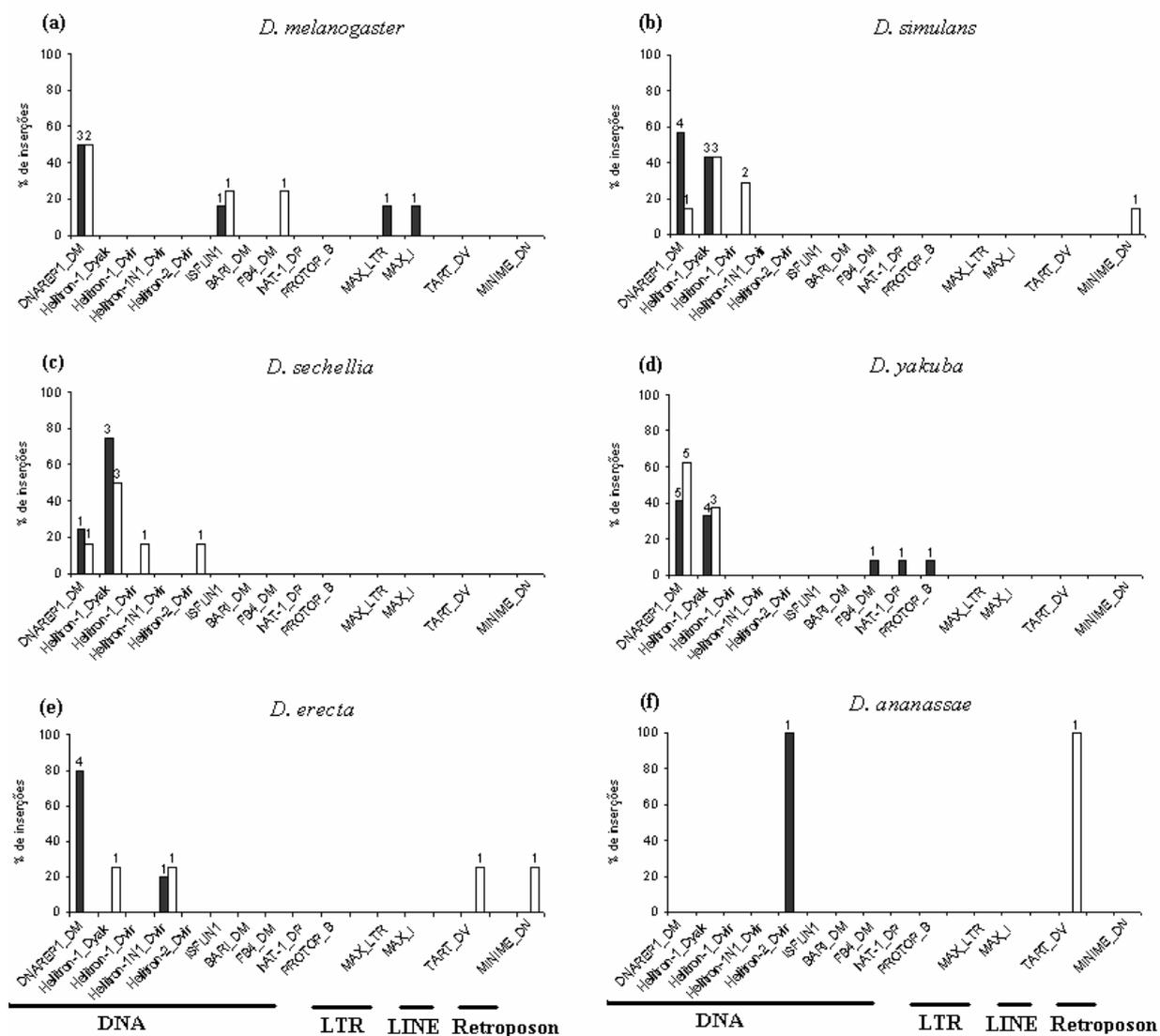


Figura 1. Porcentagem de inserções de TEs em regiões flangeadoras. As colunas em preto representam a porcentagem de inserções na região 5' e as colunas em branco as inserções nas regiões 3'. (a) *D. melanogaster*, (b) *D. simulans*, (c) *D. sechellia*, (d) *D. yakuba*, (e) *D. erecta* e (f) *D. ananassae*. Números de inserções indicados acima de cada coluna; classe ou ordem a que cada TE pertence indicadas abaixo do eixo x.

3.4.2 DNAREP1_DM e Helitron-1_Dyak: os TEs mais freqüentes nos Cyps analisados nas espécies do grupo *melanogaster*

Consideramos como um elemento freqüente nos genes *Cyps* aqueles que estavam presentes no mesmo gene em, no mínimo, três das seis espécies analisadas. Com este critério, dois elementos foram freqüentes dentre os 13 genes *Cyps*, o DNAREP1_DM e o Helitron-1_Dyak, cujas seqüências consenso apresentam, respectivamente, 594 pb e 10.846 pb (Tabela 11). A maioria das inserções do DNAREP1_DM encontradas correspondem à extremidade 3' do elemento consenso, ou à região da repetição central do elemento, com tamanho médio de 170 pb. As inserções do elemento Helitron-1_Dyak correspondem à extremidade 5' do elemento consenso e apresentaram tamanho médio de 181 pb.

Dentre as cópias dos elementos DNAREP1_DM (20), foi possível realizar um alinhamento entre 11 seqüências de modo a se fazer uma análise evolutiva objetivando avaliar se essas eram inserções ancestrais entre as espécies do grupo *melanogaster* ou decorrentes de eventos recentes de transposição. Para o Helitron-1_Dyak (16 cópias) não foi possível realizar tal análise, visto as seqüências apresentarem pequeno tamanho e corresponderem a regiões distintas do elemento consenso.

As análises das seqüências do DNAREP1_DM (Tabela 12 e Figura 2) mostraram que as cópias presentes na região flangeadora 5' do gene *Cyp6w1* de *D. simulans* (594 - 469), *D. sechellia* (594 - 263) e *D. melanogaster* (594 - 375) são seqüências derivadas de uma inserção presente no ancestral das três espécies, inseridas em sentido anti-sense em relação ao *Cyp* (*p* entre *D. simulans* vs *D. sechellia*: 0.000 e, entre as duas espécies e *D. melanogaster*: 0.143). O mesmo cenário se repete com as cópias inseridas em sentido sense na região 3' desse gene; porém, nesse caso, as cópias de *D. simulans* (411 - 568) e *D. sechellia* (442 - 590) são mais divergentes (*p* entre *D. simulans* vs *D. sechellia*: 0.071) do que aquelas na região flangeadora 5'. Do mesmo modo, as cópias presentes no gene *Cyp6a2* não evidenciam terem sido transmitidas verticalmente entre as espécies a partir do ancestral comum. Elas se agrupam com aquelas inseridas na região flangeadora 5' ou 3' do *Cyp6w1*, o que sugere serem derivadas de eventos de transposição ancestrais entre essas posições genômicas.

Tabela 11. Número dos elementos DNAREP1_DM e Helitron-1_Dyak nos genes *Cyp6w1*, *Cyp6a2* e *Cyp12d1*.

Gene	TE	Espécie	5'	3'	Total
<i>Cyp6w1</i>	DNAREP1_DM	<i>D. melanogaster</i>	3	1	4
		<i>D. simulans</i>	1	1	2
		<i>D. sechellia</i>	1	1	2
		<i>D. yakuba</i>	1	1	2
	Helitron-1_Dyak	<i>D. simulans</i>	2	2	4
		<i>D. sechellia</i>	2	2	4
		<i>D. yakuba</i>	1	1	2
<i>Cyp6a2</i>	DNAREP1_DM	<i>D. melanogaster</i>	-	1	1
		<i>D. simulans</i>	1	-	1
		<i>D. yakuba</i>	-	4	4
	Helitron-1_Dyak	<i>D. simulans</i>	1	1	2
		<i>D. sechellia</i>		1	1
		<i>D. yakuba</i>		2	2
		<i>D. erecta</i>		1	1
<i>Cyp12d1</i>	DNAREP1_DM	<i>D. simulans</i>	2		2
		<i>D. yakuba</i>	1		1
		<i>D. erecta</i>	1		1

Tabela 12. Distâncias p entre cópias do elemento *DNAREPI_DM* inseridas em regiões flangeadoras 5' (5) e 3' (3) de genes *Cyps* (*Cyp6w1*, *Cyp6a2* e *Cyp12d1*) em *D. melanogaster* (Dmel), *D. simulans* (Dsim), *D. sechellia* (Dsec) e *D. yakuba* (Dyak). Os números após o símbolo de cada espécie indicam a região que a seqüência da cópia representa no elemento consenso e o sentido sense (+) ou antisense (-) da inserção no TE.

Cópias do <i>DNAREPI_DM</i>	1	2	3	4	5	6	7	8	9	10
1. 6w1_5_Dmel_594-375 (-)										
2. 6w1_5_Dsim_594-469 (-)	0.143									
3. 6w1_5_Dsec_594-263 (-)	0.143	0.000								
4. 6w1_3_Dmel_188-587 (+)	0.357	0.314	0.314							
5. 6w1_3_Dsim_411-568 (+)	0.371	0.329	0.329	0.100						
6. 6w1_3_Dsec_442-590 (+)	0.357	0.314	0.314	0.071	0.071					
7. 6a2_5_Dsim_443-594 (+)	0.314	0.329	0.329	0.400	0.414	0.386	0.586			
8. 6a2_3_Dyak_590-454 (-)	0.371	0.314	0.314	0.329	0.329	0.314	0.600	0.357		
9. 6a2_3_Dyak_594-292 (-)	0.357	0.343	0.343	0.329	0.314	0.329	0.557	0.443	0.371	
10. 12d1_5_Dyak_119-594 (+)	0.557	0.557	0.557	0.557	0.557	0.557	0.571	0.557	0.543	0.457

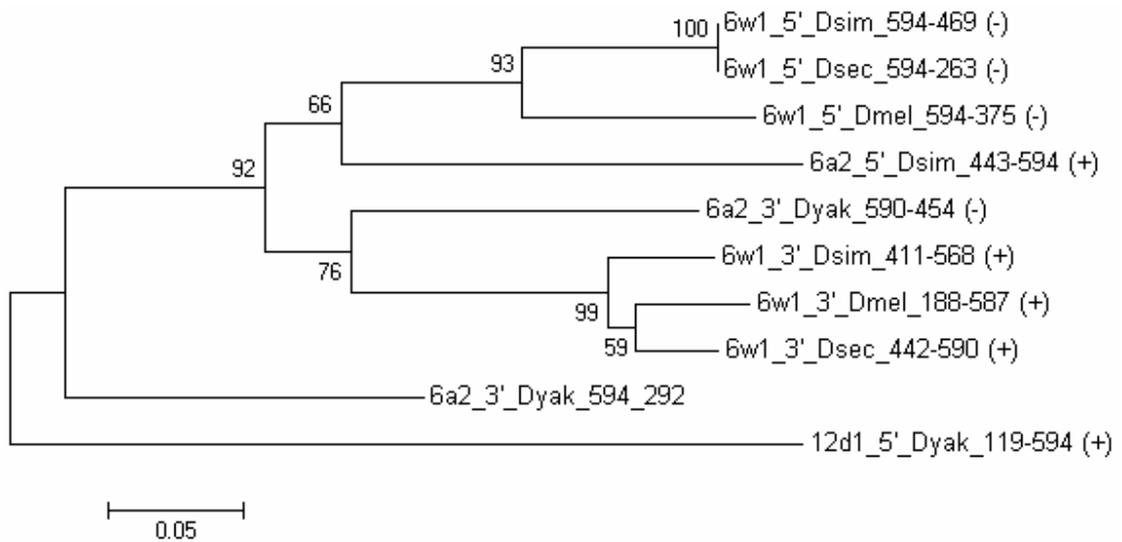


Figura 2. Árvore de neighbor-joining para cópias do elemento *DNAREPI_DM* inseridas em regiões flanqueadoras 5' e 3' de genes *Cyps* (*Cyp6w1*, *Cyp6a2* e *Cyp12d1*) em *D. melanogaster* (Dmel), *D. simulans* (Dsim), *D. sechellia* (Dsec) e *D. yakuba* (Dyak). Os números após o símbolo de cada espécie indicam a região que a seqüência da cópia representa no elemento consenso e o sentido sense (+) ou antisense (-) da inserção no TE.

3.5 DISCUSSÃO

De modo geral, as inserções de TEs tendem a se localizarem em seqüências não codificantes (regiões intergênicas e introns), estando menos presentes em regiões codificantes do genoma de *Drosophila* e outros organismos (e.g. MILLER et al., 2000; KAMINKER et al., 2002; FONTANILLAS et al., 2007) em decorrência de inserções em regiões não codificantes serem mais toleradas, enquanto que, inserções em regiões codificantes ou próximas a elas podem causar efeitos deletérios, e assim, serem removidas por seleção normalizadora (LIPATOV et al., 2005; SELA et al., 2007; YANG; BARBASH, 2008). Os resultados obtidos para os genes *Cyps* foram concordantes com essa expectativa, pois, praticamente todas as inserções estavam presentes nas regiões flanqueadoras dos genes.

Dentre os genes *Cyps* analisados neste trabalho, praticamente todas as inserções ocorreram em *Cyps* associados à resistência aos inseticidas, com exceção de uma inserção na região flanqueadora 5' do gene *Cyp306a1* de *D. yakuba* e de *D. erecta*, e no intron do *Cyp302a1* de *D. erecta*. Esses resultados são concordantes com aqueles obtidos por Chen e Li (2007) em *D. melanogaster*, analisando os mesmo genes estudados neste trabalho, e por Ricci e colaboradores (dados não publicados), em 35 genes *Cyps* de *D. melanogaster* e de *D. simulans*. Comparando os três trabalhos notam-se algumas diferenças no número e ordem das inserções encontradas, as quais podem ser explicadas devido ao tamanho da região flanqueadora analisada em cada trabalho, pois Ricci e colaboradores analisaram as regiões intergênicas 5' e 3' até encontrarem o próximo gene, independente do seu tamanho, e Chen e Li (2007), analisaram essas mesmas regiões até encontrarem o próximo gene, ou até a uma região intergênica de no máximo 10 kb.

Apesar de a diferença não ser significativa, verifica-se que os TEs inserem-se preferencialmente na região 5' dos genes *Cyps* de *D. melanogaster*, *D. yakuba* e *D. erecta*, resultado esse concordante com aquele observado por Ricci e colaboradores (dados não publicados) em *D. melanogaster* e *D. simulans* utilizando uma amostragem maior de genes. Esse resultado reforça a hipótese de que inserções de TEs na região 5' podem afetar a regulação e, conseqüentemente, a expressão dos genes, de maneira geral (JORDAN et al., 2003; van de LAGEMAAT et al., 2003; THORNBURG et al., 2006), dentre eles os *Cyps* associados à

resistência aos inseticidas (DABORN et al., 2002; CATANIA et al., 2004; SCHLENKE; BEGUN, 2004; CHUNG et al. 2007).

Além de ter como objetivo avaliar se as demais espécies de *Drosophila* tinham os genes *Cyps* associados à resistência aos inseticidas enriquecidos por TEs, como demonstrado por Chen e Li (2007), para 13 genes de *D. melanogaster*, e por Ricci e colaboradores (dados não publicados), para 35 genes de *D. melanogaster* e *D. simulans*, um dos objetivos deste trabalho foi também avaliar se existia alguma relação entre a proporção genômica de TEs, que é variável entre espécies do grupo *melanogaster*, e a proporção de TEs inseridos nos *Cyps* associados à resistência. Em outras palavras, nosso objetivo foi avaliar se a distribuição de TEs nas regiões flangeadoras dos *Cyps* associados à resistência é casual, decorrente da proporção genômica de TEs característica de cada espécie, ou não.

Neste trabalho, mostramos que essa relação não se mantém para inserções nas regiões flangeadoras dos genes *Cyps* (distantes em até 3 kb do gene). Dentre as espécies do grupo *melanogaster*, *D. ananassae*, a espécie com a maior porcentagem genômica de TEs (24,93%), apresentou o menor número médio de inserções por gene *Cyp* (1). Por outro, *D. simulans*, a espécie com a menor porcentagem (2,73 %) apresentou o segundo maior número de TEs por gene *Cyp* (4,67), porcentagem essa semelhante àquela observada em *D. yakuba*, que apresenta 4,75 inserções/*Cyp*, mas tem uma proporção genômica de TEs cinco vezes maior (12,04 %). Já, *D. melanogaster*, com praticamente o dobro de TEs que *D. simulans*, apresentou apenas 1,8 inserção por *Cyp* associado à resistência. Esses resultados mostram que a inserção de TEs nas proximidades dos genes *Cyps* associados à resistência é diferencial entre as espécies e não segue a proporção genômica, o que sugere que possam desempenhar um papel adaptativo para essas espécies.

Essa hipótese também se aplica ao TE mais abundante no genoma de *Drosophila*, o DNAREP1_DM, também denominado DINE-1 (*Drosophila* interspersed element 1). Yang et al., (2006) analisando o elemento DNAREP1_DM em *D. melanogaster* e *D. yakuba*, sugerem que esse elemento em *D. melanogaster* seja um resquício de uma transposição ancestral ocorrida aproximadamente de cinco a dez milhões de anos, e que suas cópias sejam inativas. Yang e Barbash (2008), construindo uma filogenia a partir de blocos conservados do DNAREP1_DM mostraram que *D. simulans* e *D. sechellia* se agruparam, seguidas por *D. melanogaster* e depois por *D. yakuba*, da mesma forma que a conhecida filogenia das espécies de *Drosophila*,

confirmando que cópias ancestrais desse elemento foram transmitidas verticalmente de um ancestral comum de *Drosophila*. Entretanto, o número de cópias e a divergência entre essas cópias, variam entre espécies do grupo *melanogaster*. *D. melanogaster* (355), *D. simulans* (478) e *D. sechellia* (502) têm números semelhantes, com identidade entre as cópias de cada espécie entre 90-91 %. Os valores de *D. erecta* (1013) aproximam-se mais àqueles dessas três espécies do que aos de *D. ananassae* (5027) e *D. yakuba* (5424), que apresentam 10 vezes mais cópias do DNAREP1_DM que as demais espécies do grupo *melanogaster*, como também uma porcentagem de identidade entre as cópias maior que 95 %. Essas diferenças levaram Yang e Barbash (2008) a proporem que o DNAREP1_DM sofreu uma explosão recente de transposição no genoma dessas duas espécies.

Como observado para os TEs inseridos nos *Cyps* associados à resistência, a distribuição do DNAREP1_DM também não segue a proporção genômica esperada, se a distribuição nas regiões flangeadoras dos *Cyps* fosse casual. A Figura 1 fornece o número de cópias desse elemento nas regiões distantes em até 3 Kb do início de cada gene. Embora *D. yakuba* apresente mais do que o dobro de inserções (10), como esperado pelo número de cópias genômicas (5424), *D. ananassae*, com número semelhante (5027) não tem qualquer inserção de DNAREP1_DM nesses genes. Já, *D. melanogaster*, *D. simulans* e *D. sechellia*, com números genômicos aproximados, diferem quanto ao número de inserções nesses *Cyps*, a última espécie tem menos que a metade (2) do que as duas primeiras (5). *D. erecta*, por sua vez, com mais que o dobro de seqüências DNAREP1_DM em seu genoma, apresenta número de inserção nas regiões flangeadoras semelhantes aos de *D. melanogaster* e *D. simulans*. Essas distribuições não seguem aquela esperada em decorrência do número de cópias total desse elemento em cada espécie, como também não são concordantes com a distribuição com base nas relações de parentesco entre as espécies (*D. ananassae*(*D. erecta*(*D. yakuba*(*D. melanogaster*,*D. sechellia*,*D. simulans*))))), esperada se essas cópias fossem todas transmitidas verticalmente, inativas e neutras. Um papel seletivo para as inserções do DNAREP1_DM encontradas nas regiões flangeadoras do *Cyps* também descortina a partir dessas comparações.

3.6 CONCLUSÃO

Um possível papel seletivo para as inserções de fragmento de TEs, em geral, e do DNAREP1_DM, em particular, vem dos estudos de Ricci e colaboradores (dados não publicados). Esses autores demonstraram a presença de motivos de ligação de fatores de transcrição nas cópias desse elemento inseridas nas regiões flanqueadoras dos genes *Cyps* em *D. melanogaster* e *D. simulans*, como também na seqüência consenso desse elemento. Uma larga variedade de elementos regulatórios tem sido identificada em seqüências ativas ou consensos reconstruídos de TEs ativos (JOHNSON et al., 2006; POLAK; DOMANY, 2006; THORNBURG et al., 2006; WANG, et al. 2007) e muitos estudos empíricos demonstram que esses elementos são incorporados nos aparatos regulatórios dos genes adjacentes (WESSLER et al., 1995; FERRIGNO, et al., 2001; BROSIUS, 2003; MEDSTRAND et al., 2005; ROMANISH et al., 2007). A presença de motivos regulatórios diferenciais estar relacionada à distribuição diferencial de TEs e do DNAREP1_DM em genes *Cyps* das espécies do grupo *melanogaster*. Contudo, estudos empíricos detalhados de transfecção dessas putativas seqüências reguladoras e testes de expressão são necessários para corroborar essa hipótese.

REFERÊNCIAS

ALTSCHUL, S. F. et al. Basic local alignment search tool. *J. Mol. Biol.*, v. 215, p. 403-410, 1990.

BOGWITZ, M. R. et al. CYP12A4 confers lufenuron resistance in a natural population of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, v. 102, p. 12807-12812, 2005.

BROSIUS, J. Genomes were forged by massive bombardments with retroelements and retrosequences. *Genetica*, v. 107, p. 209-238, 1999.

BROSIUS, J. The contribution of RNAs and retroposition to evolutionary novelties. *Genetica*, v. 118, p. 99-116, 2003.

CATANIA, F. et al. World-wide survey of an *Accord* insertion and its association with DDT resistance in *Drosophila melanogaster*. *Mol. Ecol.*, v. 13, p. 2491-2504, 2004.

CHEN, S.; LI, X. Transposable elements are enriched within or in close proximity to xenobiotic-metabolizing cytochrome P450 genes. *BMC Evol. Biol.*, v. 7, 2007.

CHUNG, H. et al. Cis-regulatory elements in the *Accord* retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene *Cyp6g1*. *Genetics*, v. 175, p. 1071-1077, 2007.

CLARCK, A. G. et al. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*, v. 450, p. 203-218, 2007.

DABORN, P. J. et al. A single P450 allele associated with insecticide resistance in *Drosophila*. *Science*, v. 297, p. 2253-2256, 2002.

DOOLITTLE, W. F.; SAPIENZA, C. Selfish genes, the phenotype paradigm and genome evolution. *Nature*, v. 284, p. 601-603, 1980.

FERRIGNO, O. et al. Transposable B2 SINE elements can provide mobile RNA polymerase II promoters. *Nature Genet.*, v. 28, p. 77-81, 2001.

FESCHOTTE, C. Transposable elements and the evolution of regulatory networks. *Nature Rev. Genet.*, v. 9, p. 397-405, 2008.

FONTANILLAS, P.; HARTL, D. L.; REUTER, M. Genome Organization and Gene Expression Shape the Transposable Element Distribution in the *Drosophila melanogaster* Euchromatin. *PLoS Genet.*, v. 3, p. 2256-2267, 2007.

JOHNSON, R. et al. Identification of the REST regulon reveals extensive transposable element-mediated binding site duplication. *Nucleic Acids Res.*, v. 34, p. 3862–3877, 2006.

JORDAN, I. K. et al. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.*, v. 19, p. 68-73, 2003.

JURKA, J. et al. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.*, v. 110, p. 462-467, 2005.

KAMINKER, J. S. et al. The transposable elements of the *Drosophila melanogaster* euchromatin: a genomic perspective. *Genome Biol.*, v. 3, 2002.

KIDWELL, M. G. Transposable elements and the evolution of genome size in eukaryotes. *Genetica*, v. 115, p. 49-63, 2002.

KIDWELL, M. G.; HOLYOAKE, A. J. Transposons-Induced Hotspots for Genomic Instability. *Genome Res.*, v. 11, p. 1321-1322, 2001.

LIPATOV, M. et al. Paucity of chimeric gene-transposable elements transcripts in the *Drosophila melanogaster* genome. *BMC Biology*, v. 3, p. 1-18, 2005.

LORENC, A.; MAKALOWSKI, W. Transposable elements and vertebrate protein diversity. *Genetica*. v. 118, p. 183-191, 2003.

MAKALOWSKI, W. Genomic scrap yard: how genomes utilize all that junk. *Gene*, v. 259, p. 61-67, 2000.

MAKALOWSKI, W. Not junk after all. *Science*, v. 23, p. 1246-1247, 2003.

MARSANO, R. M. et al. Evidence for a functional interaction between the *Bari1* transposable element and the cytochrome P450 *cyp12a4* gene in *Drosophila melanogaster*. *Gene*, v. 357, p. 122-128, 2005.

MEDSTRAND, P, et al. *Review*. Impact of transposable elements on the evolution of mammalian gene regulation. *Cytogenet. Genome Res.*, v. 110, p. 342–352, 2005.

MILLER, W. J. et al. Evolutionary Dynamics of the *SGM* transposon Family in the *Drosophila obscura* Species Group. *Mol. Biol. Evol.*, v. 17, p. 1597-1609, 2000.

OHNO, S.; YOMO, T. The grammatical rule for all DNA: junk and coding sequences. *Electrophoresis*, v. 12, p. 103-108, 1991.

ORGEL, L. E.; CRICK, F. H. Selfish DNA: the ultimate parasite. *Nature*, v. 284, p. 604-607, 1980.

PETROV, D. A. Evolution of genome size: new approaches to an old problem. *Trends Genet.*, v. 17, p. 23-28, 2001.

POLAK, P; DOMANY, E. Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes. *BMC Genomics*, v. 7, p. 133-148, 2006.

ROZAS, J. et al. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, v. 19, p. 2496-2497, 2003.

ROMANISH, M. T. Et al. Repeated recruitment of LTR retrotransposons as promoters by the anti-apoptotic locus NAIP during mammalian evolution. *PLoS Genet.*, v. 3, 2007.

SCHLENKE, T. A.; BEGUN, D. J. Strong selective sweep associated with transposon insertion in *D. simulans*. *Proc. Natl. Acad. Sci. USA*, v. 101, p. 1626-1631, 2004.

SELA, N. et al. Comparative analysis of transposed element insertion within human and mouse genomes reveals *Alu*'s unique role in shaping the human transcriptome. *Genome Biol.*, v. 8, R137, 2007.

TAMURA K, DUDLEY J, NEI M, KUMAR S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, v. 24, p.1596–1599, 2007.

THORNBURG, B. G.; GOTEA, V.; MAKALOWSKI, W. Transposable elements as a significant source of transcription regulating signals. *Gene*, v. 365, p. 104-110, 2006.

TIJET, N.; HELVIG, C.; FEYEREISEN, R. The cytochrome P450 gene superfamily in *Drosophila melanogaster*: Annotation, intron-exon organization and phylogeny. *Gene*, v. 262, p. 189-198, 2001.

VAN DE LAGEMAAT, L. M. et al. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.*, v. 19, p. 530-536, 2003.

WANG, T. et al. Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proc. Natl. Acad. Sci. U S A.*, v. 104, p. 18613–8, 2007.

WESSLER, S. R.; BUREAU, T. E.; WHITE, S. E. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Curr. Opin. Genet. Dev.*, v. 5, p. 814–821, 1995.

YANG, H-P.; BARBASH, D. A. Abundant and species-specific *DINE-1* transposable elements in 12 *Drosophila* genomes. *Genome Biology*, v. 9, R39, 2008.

YANG, H-P. et al. Genomewide Comparative Analysis of the High Abundant Transposable Element *DINE-1* Suggest a Recent Transpositional Burst in *Drosophila yakuba*. *Genetics*, v. 173, p. 189-196, 2006.

4 DISCUSSÃO GERAL

O movimento e o acúmulo dos TEs nos genomas têm exercido uma forte influência na trajetória evolutiva de seus hospedeiros desde que podem gerar mutações e rearranjos cromossômicos que contribuem para a evolução da estrutura e função do genoma onde estão inseridos (BROSIUS, 1999; MAKALOWSKI, 2003; KIDWELL, 2002; GOTEA; MAKALOWSKI, 2006; FESCHOTTE, 2008). Quando inseridos em regiões reguladoras, também podem afetar positivamente a expressão do gene hospedeiro, como também de genes adjacentes, devido à introdução de elementos reguladores nessas seqüências (CONTE et al., 2002; JORDAN et al., 2003; THORNBURG et al., 2006). Em diversos TEs, foram identificadas seqüências que atuam como sítios de ligação de fatores transcrição (*TFBS*), nas quais se ligam motivos reguladores da expressão gênica (THORNBURG et al., 2006; ALMEIDA; CARARETO, 2006; POLAVARAPU et al., 2008). Feschotte (2008) observou que a manutenção de transposons de DNA ou de suas transposases, no genoma hospedeiro deve-se ao fato de que essas seqüências podem ser fontes para ligação de fatores de transcrição. Assim, os fragmentos de TEs podem representar um importante papel na regulação gênica, sendo mantidos ou até exaptados nas proximidades dos genes hospedeiros.

A maior parte das inserções de TEs está localizada em regiões intergênicas, introns e UTRs tendo, desse modo, maior chance de serem mantidas no genoma, pois, estão sob menor pressão seletiva (KIDWELL, 2002; SELA et al., 2007; FONTANILLAS et al., 2007; YANG; BARBASH, 2008). Neste trabalho, essa relação também fica explícita para os genes *Cyps* e seus genes flanqueadores, nos quais a maior parte das inserções de TEs ocorreu nas regiões intergênicas.

No genoma de *D. melanogaster*, as classes de TEs com maior abundância (BERGMAN et al., 2006) são, em ordem decrescente, DNA > LTRs > Retroposons, quando incluído o elemento DNAREP1_DM, mas LTR > Retroposons > DNA, quando o DNAREP1_DM não é computado na classe dos transposons de DNA, em decorrência da dúvida quanto a sua classificação. Os resultados relatados no Capítulo 1 mostram que, dentre as inserções presentes nos genes *Cyps*, a maior parte corresponde a transposons de DNA (mesmo excluindo o DNAREP1_DM), seguidos pelos Retroposons e LTRs. A mesma tendência à inserção de transposons próximos a regiões

reguladoras de genes foi demonstrada para o transposon *P* (TOWER et al., 1993; SPRADLING et al., 1995) e o *Sleeping Beauty* (YANT et al., 2005). Uma explicação para essa predominância poderia ser que a inserção de diferentes elementos em todas as regiões intergênicas é casual, mas ocorreria uma manutenção preferencial dos elementos de DNA na região 5'. Essa predominância de inserção nas regiões flanqueadoras 5', onde se acumula a maior parte das seqüências reguladoras da expressão gênica, suporta a hipótese de que fragmentos de TEs enriquecem as regiões promotoras, doando sinais reguladores da transcrição e podendo alterar a expressão gênica, como relatado em humanos por Jordan et al. (2003) e Thornburg et al. (2006). Outra hipótese seria que as seqüências codificadoras das transposases podem ser fontes para ligação de fatores de transcrição e, assim, seriam mantidas ou até exaptadas nas proximidades dos genes hospedeiros, desempenhando um importante papel na regulação gênica (FESCHOTTE, 2008).

Dentre os elementos freqüentemente inseridos em regiões flanqueadoras dos genes *Cyp*s destacou-se o elemento DNAREP1_DM, ocorrendo em maior número nas seis espécies estudadas do grupo *melanogaster* de *Drosophila*. As cópias desse TE nos genomas de *D. melanogaster* e de *D. simulans* apresentam identidade média de 90 % (YANG; BARBASH, 2008). Mostramos no capítulo I deste trabalho que as cópias do DNAREP1_DM localizadas nas regiões intergênicas dos genes *Cyp*s dessas duas espécies são mais divergentes, e mais antigas, em relação às cópias distribuídas ao longo do genoma. Este percentual inferior de identidade, associado ao pequeno tamanho, corroboram o fato de serem cópias antigas nas duas espécies, desse modo, já expostas ao crivo da seleção normalizadora que resulta em eliminação de TEs deletérios em regiões próximas a genes. Esses fragmentos do elemento DNAREP1_DM correspondem à extremidade 3' do elemento consenso, que mostramos conter seqüências homólogas a sítios de ligação de fatores de transcrição. A análise da inserção de TEs em linhagens de diferentes regiões geográficas também confirma os achados da análise *in silico* de que as cópias do elemento DNAREP1_DM presentes nas regiões flanqueadoras desses genes correspondem à extremidade 3' do elemento consenso. Embora tendo tamanhos variáveis, essas seqüências preservam possíveis sítios de ligação de fatores de transcrição e essa pode ser a causa de sua manutenção nas regiões flanqueadoras dos genes *Cyp*s.

Os elementos de transposição têm sido associados à evolução da resistência aos inseticidas, por estimularem a expressão de genes responsáveis pela resistência metabólica em insetos, ou mesmo, por atuarem na geração de novas funções protéicas devido à alteração da

estrutura do gene (DABORN et al., 2002; AMINETZACH et al., 2005). Porém, essa associação, não está completamente caracterizada, visto ter sido relatada também a presença de inserções de TEs e elevada expressão de genes *Cyp6* em algumas linhagens suscetíveis (SCHLENKE; BEGUN, 2004; FESTUCCI-BUSELLI et al., 2005; KURUGANTI et al., 2007). Os resultados relatados nos Capítulos 1 e 2 indicaram que as regiões flanqueadoras distantes em até 3kb dos genes *Cyp6* associados à resistência acumulam fragmentos de TEs em maior proporção que os demais *Cyp6* e, que essas proporções, variam entre as espécies do grupo *melanogaster* de *Drosophila*. Mostramos, neste trabalho, que a ocorrência de TEs preferencialmente em regiões flanqueadoras de *Cyp6* associados à resistência não se restringe a *D. melanogaster*, como inicialmente sugerido por Chen e Li (2007).

Entretanto, a relação entre a presença dos fragmentos de TEs em genes *Cyp6* associados à resistência e a manifestação da resistência não ficou evidenciada pelo estudo de linhagens de *D. melanogaster* e *D. simulans* resistentes e suscetíveis ao DDT (Capítulo 1). Não encontramos qualquer evidência de que esses fragmentos desempenhem alguma função diretamente relacionada à resistência aos inseticidas. Entretanto, o fato de que a inserção de TEs nas proximidades dos genes *Cyp6* associados à resistência é diferencial entre as espécies, e não segue a proporção genômica de TEs em cada uma, como mostrado no Capítulo 2, é mais um indício de que essas sequências possam desempenhar algum papel adaptativo nessas espécies.

5 CONCLUSÕES

O presente trabalho permitiu estabelecer as seguintes conclusões:

1. Genes *Cyps* associados à resistência aos inseticidas e à função monooxigenase geral, bem como, seus flaqueadores (exceto em *D. simulans*) apresentaram maior acúmulo de elementos de transposição, ao contrário dos *Cyps* associados ao desenvolvimento nos quais as inserções foram praticamente ausentes. Esses resultados indicam que os primeiros apresentam maior plasticidade genômica enquanto os últimos estão sob pressão seletiva contrária à inserção ou à manutenção de elementos de transposição.
2. A maior parte das inserções de TEs nos 35 genes investigados ocorreu em seqüências não codificantes (regiões intergênicas e introns), e foram praticamente ausentes em regiões codificantes.
3. As inserções de elementos de transposição se localizaram preferencialmente nas regiões flaqueadoras 5' dos genes, onde se encontram a maior parte das seqüências reguladoras, podendo assim, influenciar a expressão gênica.
4. As análises populacionais de seqüências de elementos de transposição inseridas nas regiões flaqueadoras dos genes *Cyps* associados à resistência evidenciaram a presença de polimorfismo interpopulacional, porém, não permitiram esclarecer a associação entre inserção desses elementos em genes *Cyps* e resistência aos inseticidas em linhagens resistentes de *D. melanogaster* e *D. simulans*.
5. A análise *in silico* nas seis espécies do grupo *melanogaster* mostrou que a distribuição de elementos de transposição nas regiões flaqueadoras dos *Cyps* associados à resistência é diferencial e não segue a proporção genômica de TEs das espécies, o que sugere que esses elementos possam desempenhar um papel adaptativo nessas espécies.
6. O elemento DNAREP1_DM, o elemento de transposição mais abundante nas regiões flaqueadoras dos genes *Cyps*, pode estar sendo mantido nessas regiões pelo fato de apresentarem putativos sítios de ligação de fatores de transcrição, e desse modo, desempenhar algum papel na regulação dos genes *Cyps* nas espécies de *Drosophila*.

REFERÊNCIAS

ALMEIDA, L. M.; CARARETO, C. M. A. Sequence heterogeneity and phylogenetic relationships between the *copia* retrotransposon in *Drosophila* species of the *repleta* and *melanogaster* groups. *Genet. Sel. Evol.*, v. 38, p. 535-550, 2006.

AMICHOT, M. et al. Point mutations associated with insecticide resistance in the *Drosophila* cytochrome P450 *Cyp6a2* enable DDT metabolism. *Eur. J. Biochem.*, v. 271, p. 1250-1257, 2004.

AMINETZACH, Y. T.; MACPHERSON, J. M.; PETROV, D. A. Pesticide Resistance via Transposition-Mediates Adaptive Gene Truncation in *Drosophila*. *Science*, v. 309, p. 764-767, 2005.

BERGMAN, C. M. et al. Recurrent insertion and duplication generate networks of transposable elements in the *Drosophila melanogaster* genome. *Genome Biol.*, v. 7, R112, 2006.

BOGWITZ, M. R. et al. CYP12A4 confers lufenuron resistance in a natural population of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, v. 102, p. 12807-12812, 2005.

BRANDT, A. et al. Differential expression and induction of two *Drosophila* cytochrome P450 genes near the *Rst(2)DDT* locus. *Insect Mol. Biol.*, v. 11, p. 337-341, 2002.

BROSIUS, J. Genomes were forged by massive bombardments with retroelements and retrosequences. *Genetica*, v. 107, p. 209-238, 1999.

BUNDOCK, P.; HOOYAKAAS, P. An *Arabidopsis hAT*- like transposase is essential for plant development. *Nature*, v. 436, p. 282-284, 2005.

CATANIA, F. et al. World-wide survey of an *Accord* insertion and its association with DDT resistance in *Drosophila melanogaster*. *Mol. Ecol.*, v. 13, p. 2491-2504, 2004.

CHEN, S.; LI, X. Transposable elements are enriched within or in close proximity to xenobiotic-metabolizing cytochrome P450 genes. *BMC Evol. Biol.*, v. 7, p. 1-13, 2007.

CHUNG, H. et al. Cis-regulatory elements in the *Accord* retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene *Cyp6g1*. *Genetics*, v. 175, p. 1071-1077, 2007.

CHUNG, H. et al. Characterization of *Drosophila melanogaster* cytochrome P450 genes. *Proc. Natl. Acad. Sci. USA*, doi. 10.1073/pnas.0812141106, 2009.

CLARCK, A. G. et al. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*, v. 450, p. 203-218, 2007.

CONTE, C.; DASTUGUE, B.; VAURY, C. Coupling of enhancer and insulator properties identified in two retrotransposons modulates their mutagenic impact on nearby genes. *Mol. Cell. Biol.*, v. 22, p. 1767-1777, 2002.

DABORN, P. et al. DDT resistance in *Drosophila* correlates with *Cyp6g1* over-expression and confers cross-resistance to the neonicotinoid imidacloprid. *Mol. Genet. Genomics*, v. 266, p. 556-563, 2001.

DABORN, P. J. et al. A single P450 allele associated with insecticide resistance in *Drosophila*. *Science*, v. 297, p. 2253-2256, 2002.

DOOLITTLE, W. F.; SAPIENZA, C. Selfish genes, the phenotype paradigm and genome evolution. *Nature*, v. 284, p. 601-603, 1980.

FESTUCCI-BUSELLI, R. A. et al. Expression of *Cyp6g1* e *Cyp12d1* in DDT resistant and susceptible strains of *Drosophila melanogaster*. *Insect Mol. Biol.*, v. 14, p. 69-77, 2005.

FINNEGAN, D. J. Eukaryotic transposable elements and genome evolution. *Trends Genet.*, v. 5, p. 103-107, 1989.

FONTANILLAS, P.; HARTL, D. L.; REUTER, M. Genome Organization and Gene Expression Shape the Transposable Element Distribution in the *Drosophila melanogaster* Euchromatin. *PLoS Genet.*, v. 3, p. 2256-2267, 2007.

FFRENCH-CONSTANT, R. H. et al. Why are there so few resistance-associated mutations in insecticide target genes? *Phil. Trans. R. Soc. Lond.*, v. 353, p. 1685-1693, 1998.

GANKO, E. W. et al. Evidence for the contribution of LTR retrotransposons to *C. elegans* gene evolution. *Mol. Biol. Evol.*, v. 20, p. 1925-1931, 2003.

GANKO, E. W. et al. LTR Retrotransposon-Gene Associations in *Drosophila melanogaster*. *J. Mol. Evol.*, v. 62, p. 111-120, 2006.

GOTEA, V.; MAKALOWSKI, W. Do transposable elements really contribute to proteomes? *Trends Genet.*, v. 22, p. 260-267, 2006.

HEMINGWAY, J.; RANSON, H. Review Insecticide Resistance in Insect Vectors of Human Disease. *Annu. Rev. Entomol.*, v. 45, p. 371-391, 2000.

JORDAN, I. K. et al. Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.*, v. 19, p. 68-73, 2003.

KIDWELL, M. G. Transposable elements and the evolution of genome size in eukaryotes. *Genetica*, v. 115, p. 49-63, 2002.

KIDWELL, M. G.; LISCH, D. R. Transposable elements and host genome evolution. *Trends Ecol. Evol.*, v. 15, p. 95-99, 2000.

KIDWELL, M. G.; LISCH, D. R. Perspective: transposable elements, parasitic DNA, and genome evolution. *Evolution Int. J. Org. Evolution*, v. 55, p. 1-24, 2001.

KIDWELL, M. G.; HOLYOAKE, A. J. Transposons-Induced Hotspots for Genomic Instability. *Genome Res.*, v. 11, p. 1321-1322, 2001.

KURUGANTI, S. et al. High expression of *Cyp6g1*, a cytochrome P450 gene, does not necessarily confer DDT resistance in *Drosophila melanogaster*. *Gene*, v. 388, p. 43-53, 2007.

LI, X.; SCHULER, M. A.; BERENBAUM, M. R. Molecular Mechanisms of Metabolic Resistance to Synthetic and Natural Xenobiotics. *Annu. Rev. Entomol.*, v. 52, p. 231-253, 2007.

LOPES, F. R. et al., Transposable elements in *Coffea* (Gentianales: Rubiaceae) transcripts and their role in the origin of protein diversity in flowering plants. *Mol. Genet. Genomics*, v. 279, p. 385-401, 2008.

LORENC, A.; MAKALOWSKI, W. Transposable elements and vertebrate protein diversity. *Genetica*, v. 118, p. 183-191, 2003.

MAITRA S. et al. Three second chromosome-linked clustered CYP6 genes show differential and barbital-induced expression in DDT-resistant and susceptible strains of *Drosophila melanogaster*. *Gene*, v. 180, p. 165-171, 1996.

MAITRA S. et al. Factors on the third chromosome affect the level of CYP6A2 and CYP6A8 expression in *Drosophila melanogaster*. *Gene*, v. 248, p. 147-156, 2000.

MAKALOWSKI, W. Genomic scrap yard: how genomes utilize all that junk. *Gene*, v. 259, p. 61-67, 2000.

MAKALOWSKI, W. Not junk after all. *Science*, v. 23, p. 1246-1247, 2003.

MARSANO, R. M. et al. Evidence for a functional interaction between the *Bari1* transposable element and the cytochrome P450 *cyp12a4* gene in *Drosophila melanogaster*. *Gene*, v. 357, p. 122-128, 2005.

McCLINTOCK, B. Controlling elements and the gene. *Cold Spring Harbor Symposia on Quantitative Biology*, EUA, v. 21, p. 197-216, 1957.

MEYERS, B. C.; TINGEY, S. V.; MORGANTE, M. Abundance, Distribution, and Transcriptional Activity of Repetitive Elements in the Maize Genome. *Genome Res.*, v. 11, p. 1660-1676, 2001.

MILLER, W. J. et al. Molecular domestication – more than a sporadic episode in evolution. *Genetica*, v.107, p. 197-207, 1999.

NEKRUTENKO, A.; LI, W. H. Transposable elements are found in a large number of human protein-coding genes. *Trends Genet.*, v. 17, p. 619-621, 2001.

OHNO, S.; YOMO, T. The grammatical rule for all DNA: junk and coding sequences. *Electrophoresis*, v. 12, p. 103-108, 1991.

ORGEL, L. E.; CRICK, F. H. Selfish DNA: the ultimate parasite. *Nature*, v. 284, p. 604-607, 1980.

PEDRA, J. H. F. et al. Genome-wide transcription profile of field- and laboratory-selected dichlorodiphenyltrichloroethane (DDT)-resistant *Drosophila*. *Proc. Natl. Acad. Sci. USA*, v. 101, p. 7034-7039, 2004.

PETROV, D. A. Evolution of genome size: new approaches to an old problem. *Trends Genet.*, v. 17, p. 23-28, 2001.

POLAVARAPU, N. et al. Evolutionary rates and patterns for human transcription factor binding sites derived from repetitive DNA. *BMC Genomics*, v. 9, 2008.

RANSON, H. et al. Evolution of supergene families associated with insecticide resistance. *Science*, v. 298, p. 179-181, 2002.

REISS, D. et al. Domesticated *P* Elements in *Drosophila montium* Species Subgroup Have a New Function Related to a DNA Binding Property. *J. Mol. Evol.*, v. 61, p. 470-480, 2005.

SCHLENKE, T. A.; BEGUN, D. J. Strong selective sweep associated with transposon insertion in *D. simulans*. *Proc. Natl. Acad. Sci. USA*, v. 101, p. 1626-1631, 2004.

SELA, N. et al. Comparative analysis of transposed element insertion within human and mouse genomes reveals *Alu*'s unique role in shaping the human transcriptome. *Genome Biol.*, v. 8, p. R137, 2007.

SPRADLING, A. C. et al. Gene disruptions using *P* transposable elements: An integral component of the *Drosophila* genome project. *Proc. Natl. Acad. Sci. USA*, v. 92, p. 10824-10830, 1995.

THORNBURG, B. G.; GOTEA, V.; MAKALOWSKI, W. Transposable elements as a significant source of transcription regulating signals. *Gene*, v. 365, p. 104-110, 2006.

TIJET, N.; HELVIG, C.; FEYEREISEN, R. The cytochrome P450 gene superfamily in *Drosophila melanogaster*: Annotation, intron-exon organization and phylogeny. *Gene*, v. 262, p. 189-198, 2001.

TOWER, J. et al. Preferential Transposition of *P* Elements to Nearby Chromosomal Sites. *Genetics Society of America*, v. 133, p. 347-359, 1993.

VAN DE LAGEMAAT, L. M. et al. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.*, v. 19, p. 530-536, 2003.

WICKER, T. et al. A unified classification system for eukaryotic transposable elements. *Nature Rev. Genet.*, v. 8, p. 973-982, 2007.

YANG, H-P.; BARBASH, D. A. Abundant and species-specific *DINE-1* transposable elements in 12 *Drosophila* genomes. *Genome Biol.*, v. 9, p. R39.1-14, 2008.

YANT, S. R. et al. High-Resolution Genome-Wide Mapping of Transposon Integration in Mammals. *Mol. Cell. Biol.*, v. 25, p. 2085-2094, 2005.