



Evolução do Retroelemento **gypsy** em espécies de
Drosophila* e *Zaprionus indianus:

Uma Abordagem Filogenética

Fabiana Herédia

Orientadora: Dra. Vera L. da S. Valente Gaiety
Co-Orientador: Dr. Élgion L. da S. Loreto

Porto Alegre, 2002

Evolução do Retroelemento
gypsy em espécies de *Drosophila*
e *Zaprionus indianus*:

Uma Abordagem Filogenética

Fabiana Herédia

Porto Alegre, 2002

Evolução do Retroelemento *gypsy* em espécies de *Drosophila* e *Zaprionus indianus*:

Uma Abordagem Filogenética

Tese apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal do Rio Grande do Sul, para obtenção do grau de Doutor

Fabiana Herédia

Orientador: Dra. Vera L. da S. Valente Gaiesky

Co-Orientador: Dr. Élgion L. da S. Loreto

Porto Alegre, 2002.

Este trabalho foi realizado no laboratório de *Drosophila* do Departamento de Genética da Universidade Federal do Rio Grande do Sul com recursos do CNPq, PROPESQ-UFRGS, FINEP e FAPERGS.

Agradecimentos

A minha querida orientadora Verinha por ensinar-me o equilíbrio entre o profissionalismo, a ética, a criatividade e a sensibilidade.

Ao meu querido orientador Élgion por ser meu exemplo de pesquisador, pelo seu voto de confiança e pelo estímulo em todos os momentos deste trabalho.

À Lenira e à Julinha por acolherem-me em sua casa de forma tão carinhosa.

Aos meus amigos e companheiros do laboratório de Drosophila por estes sete anos de delicioso convívio e momentos inesquecíveis:

À Rosane por seus conselhos sábios e o ombro amigo.

À Marisa por suas conversas e gargalhadas que contagiam e alegam o nosso ambiente.

Ao Vitinho por ser um exemplo de dedicação, pela sua alegria e cantorias...

Ao Fabiano/Fabício pela grande amizade que construímos.

À Norma pelo seu exemplo de força de vontade e responsabilidade que, em muitos momentos, serviram-me de inspiração.

À Adri pelo seu jeito meigo de ser e pela grande ajuda nos últimos detalhes desta tese.

À Chirlei e à Aninha por serem ótimas, queridas e fiéis companheiras de trabalho, principalmente, nestes últimos meses de sufoco.

À Dani, minha grande amiga de Floripa, por nossas longas conversas e confidências, e por ter me cedido muitas amostras de suas coletas para a realização deste trabalho.

Ao Luciano por, também, compartilhar comigo de várias das suas coletas, e por suas palavras e críticas que me fizeram crescer muito, profissional e pessoalmente.

Aos demais amigos Cláudia, Monica, Mel, André, Marília, Cris, Tatiane, Tiago e Carina que são pessoas maravilhosas e completam o melhor time de trabalho do mundo.

Ao Elmo pelo apoio técnico, por fazer tudo funcionar ao nosso redor, e pela grande amizade.

Aos meus pais e a minha avó pelo apoio incondicional e o enorme amor que sempre dedicaram a mim.

Ao Tiago por toda a força, o carinho e o companheirismo que temo, nem sempre, retribuir à altura.

À família Berger Kuschick pelo apoio e confiança.

“... the presented thesis, and evidence for it, could not be accepted by the majority of geneticists or by others biologists”

Comentário de Bárbara McClintock na introdução de uma revisão sobre seus primeiros esforços de divulgação da descoberta dos elementos transponíveis em 1950. (Mc Clintock 1987).

Aproximadamente

50 anos

depois...

Mobile elements inserted in the distant past have taken on important functions

(Britten 1997)

Transposable elements as source of variation in animals and plants

(Kidwell e Lisch 1997)

Selfish genetic elements and speciation

(Hurst e Schilthuizen 1998)

Stress and transposable elements: co-evolution or useful parasites?

(Capy et al. 2000)

Transposons and genome evolution in plantas

(Fedoroff 2000)

Perspective: Transposable Elements, Parasitic DNA and Genome Evolution

(Kidwell e Lisch 2001)

**Transposable elements are found in a large number of human protein-coding
genes**

(Nekrutenko e Li 2001)

Genome evolution: Sex and the transposable element

(Wright e Finnegan 2001)

SUMÁRIO

Capítulo 1

Introdução.....	1
ELEMENTOS TRANSPONÍVEIS E O GENOMA HOSPEDEIRO.....	1
O RETROELEMENTO TRANSPONÍVEL <i>gypsy</i>	36
O GÊNERO <i>Drosophila</i> E <i>Zaprionus indianus</i>	51
Objetivos.....	56

Capítulo 2

Evolução complexa de <i>gypsy</i> em <i>Drosophila</i> e <i>Zaprionus indianus</i>	58
--	----

Capítulo 3

Analyses of the <i>gypsy</i> transposable element in neotropical <i>Drosophila maculifrons</i> and <i>Drosophila griseolineata</i> using <i>in situ</i> hybridization in polytene chromosomes and Southern blot.....	141
--	-----

Capítulo 4

Discussão Geral e perspectivas.....	159
Resumo.....	166
Summary.....	168
Referências Bibliográficas.....	170

CAPÍTULO 1

INTRODUÇÃO

ELEMENTOS TRANSPONÍVEIS E O GENOMA HOSPEDEIRO

Até a primeira metade do século XX, a ciência tinha o genoma como uma entidade estática, mudando apenas numa escala evolutiva. Após esta época, dados de instabilidade genômica em vários seres vivos começaram a permitir o questionamento de tal dogma. Hoje, sabemos da existência de fatores, antes inimagináveis, que são capazes de produzir variabilidade genética de uma geração para outra. Os elementos transponíveis (TEs) hoje são aceitos como um importante fator evolutivo e como seqüências capazes de reorganizar o genoma dos seres vivos. Por isto, estas entidades são evocadas em discussões que abordam os mais diferentes temas de interesse sobre a organização biológica. Pesquisadores do mundo inteiro tentam entender a dimensão do papel destes elementos móveis para a manutenção da vida, investigam sua distribuição e representatividade nos mais diferentes genomas, as relações evolutivas entre eles e deles com o genoma hospedeiro, além de sua função como vetor de troca de informação genética entre os seres vivos. A relação evolutiva dos elementos transponíveis com os vírus coloca estas seqüências numa posição bastante intrigante. Se vírus são genes que “fugiram de sua morada em busca de promiscuidade”, os elementos transponíveis optaram por estratégias talvez mais inteligentes... Este

trabalho busca contribuir para o conhecimento sobre a distribuição, representação e evolução do retroelemento *gypsy*, um elemento transponível extremamente relacionado filogeneticamente com os retrovírus, em espécies do gênero *Drosophila*, assim como na espécie invasora *Zaprionus indianus*.

DEFINIÇÃO DE ELEMENTOS TRANSPONÍVEIS

Elementos transponíveis (TEs) são seqüências de DNA incluídas na fração de DNA medianamente repetido do genoma. No entanto, algumas peculiaridades os distinguem das demais seqüências presentes nesta categoria. A principal característica dos TEs é a capacidade (intrínseca ou não) de trocar de posição no genoma. Outras características importantes são: 1) causar uma duplicação do sítio de inserção; 2) apresentar polimorfismo de sítios de inserção, e, finalmente, 3) apresentar uma variabilidade no número de cópias dentro e entre espécies (Capy et al. 1998).

COMO OS ELEMENTOS TRANSPONÍVEIS SE ENCONTRAM NO GENOMA

Desde que foram descobertos, por Bárbara McClintock na década de 40, estas seqüências móveis tem sido encontradas nos mais diferentes organismos investigados, sendo, inclusive, parte considerável de seus genomas (revisão em Kidwell e Lisch 2000).

O número de cópias de TEs estende-se desde poucas até milhares delas. Estas seqüências representam mais do que 50% do genoma do milho. No homem, existem mais do que um milhão de cópias da família *Alu* e estas seqüências tem se movimentado tão recentemente que esta porção do

genoma difere a nível populacional, sendo este polimorfismo útil para inferências filogenéticas (Capy et al. 1998).

Apesar dos elementos transponíveis estarem presentes em todos os genomas investigados, sua representatividade varia, enormemente, entre os taxons. Em Angiospermas, diferenças entre espécies chegam a atingir cerca de 4 ordens de magnitude (Kumar et al. 1997). Entre os animais, um exemplo desta variação é revelada quando observamos que TEs chegam a representar cerca de 50% do genoma humano e, apenas, 5% do genoma do verme *Caenorhabditis elegans* (revisão em Kidwell e Lisch 2000). A explicação para estas discrepâncias é alvo de discussão por vários autores, e o modo de reprodução da espécie investigada parece influenciar, fortemente, este aspecto (Wright e Finnegan 2001).

Num mesmo genoma, podem coexistir dois tipos de cópias de um dado TE: os elementos ou cópias completas, também chamados de autônomos, contendo todas as seqüências necessárias para sua transposição, e aqueles deletados, também referidos como incompletos, não autônomos ou defectivos. Nestes últimos, observamos deleções, e, em função disto, estes elementos necessitam de fatores providos por cópias completas para sua transposição. Como os TEs estão sujeitos a altas taxas mutacionais, a freqüência de elementos não autônomos é bastante alta (Kidwell e Lisch 2001).

Famílias e subfamílias diferentes de elementos transponíveis podem, ainda, coexistir no mesmo genoma (Flavell 1992; Clark et al. 1994; Lohe et al. 1995), e o número de famílias e de cópias de cada família varia conforme a espécie hospedeira. Parece não existir qualquer correlação entre o número de

cópias e o tamanho do genoma (Capy et al. 1998). No entanto, Petrov (2001) discute a atividade de elementos transponíveis como uma das forças que operam no aumento do genoma de uma espécie.

Dentro do genoma, a distribuição das cópias de elementos transponíveis parece não ser aleatória, e muitas famílias apresentam preferências insercionais (Kidwell e Lisch 2001). Mesmo dentro da heterocromatina, região apontada como povoada de TEs, sua distribuição não é homogênea (Dimitri et al. 1997). A heterocromatina parece ser um alvo de preferência dos TEs, e alguns autores postulam que esta especificidade pode estar relacionada com a alta densidade de cortes no DNA desta região, cortes que são reparados com a entrada de elementos transponíveis (Dimitri e Junakovic 1999). Esta distribuição não aleatória de agrupamentos de TEs, em vários organismos, sugere que os genomas possam ser divididos dentro de uma variedade de "nichos ecológicos", os quais são explorados de distintas formas por diferentes hospedes. Os transposons que evitam a heterocromatina e preferem seqüências gênicas podem valer-se da vantagem da cromatina local estar aberta e de fatores de transcrição mais disponíveis. No entanto, arcam com grandes custos de seleção contra novas inserções (Kidwell e Lisch 2001).

Tem sido proposto que elementos não infecciosos tendem a apresentar uma certa preferência insercional como um comportamento adaptativo, resultado de um processo evolutivo para um mecanismo de reconhecimento de regiões não importantes no genoma. Ao contrário, retrovírus estão livres desta pressão seletiva devido à sua habilidade infecciosa. Tal especificidade é

caracterizada para elementos da família *Ty1*, *Ty3* e *Ty5* e, também, para *gypsy* de *D. melanogaster* (Dej et al. 1998).

O impacto de TEs pode adquirir proporções distintas dependendo se sua localização for eucromática ou heterocromática. Cópias ativas são encontradas principalmente na eucromatina, enquanto que na heterocromatina observamos muitas cópias defectivas e rearranjadas. No entanto, permanece ainda a dúvida se existe uma preferência insercional de TEs por determinadas regiões ou se este padrão é resultado de diferentes pressões seletivas (Capy et al. 1998).

CLASSIFICAÇÃO DOS ELEMENTOS TRANSPONÍVEIS

A classificação dos TEs baseia-se, principalmente, no modo como o elemento se transpõe. Além disto, suas características estruturais e moleculares são, também, consideradas. Segundo estas premissas, os elementos são agrupados em classes, subclasses, superfamílias, famílias e subfamílias. Cada um destes níveis segue os seguintes critérios:

- classe: mecanismo de transposição;
- subclasse: estrutura molecular (presença ou ausência de domínios e assinaturas);
- superfamília: existência de motivos ou assinaturas similares;
- família: existência de similaridade em motivos e em seqüência (entre 30% e 50% de similaridade em seqüências de aminoácidos);
- subfamília: similaridade em seqüência (entre 50% e 70% de similaridade a nível de aminoácidos)

Os TEs têm sido divididos em duas grandes classes de acordo com o mecanismo usado para a sua transposição. Elementos da Classe I são TEs que se movimentam através de um intermediário de RNA que é transcrito, reversamente, por uma transcriptase reversa antes da nova inserção. Estas seqüências são chamadas de retrotransposons e incluem os retrotransposons com LTRs (longas repetições terminais) que são estruturalmente similares aos retrovírus, e elementos sem LTRs ou retroposons não virais. Elementos da Classe II movem-se via um intermediário de DNA através de um mecanismo de excisão do sítio original e reinserção do elemento em outro ponto do genoma.

Elementos da Classe II têm sido encontrados tanto em procariotos como em eucariotos, enquanto retrotransposons parecem estar restritos aos eucariotos (Capy et al. 1998).

Além destas duas classes bem documentadas, existem elementos transponíveis cujo mecanismo de transposição permanece a ser esclarecido. Muitos autores os colocam em uma terceira classe (Classe III).

ELEMENTOS CLASSE I

A divisão deste grupo em retrotransposons, apresentando ou não LTRs, coincide com outro critério que é a presença ou não da assinatura DDE em suas integrases. Enquanto domínios são seqüências conservadas de poucos aminoácidos consecutivos, assinaturas são consideradas resíduos conservados, mas não consecutivos em uma proteína.

RETROTRANSPOSONS COM LTRs

A organização geral destes elementos é muito similar à dos retrovírus. O tamanho dos mesmos varia de 5-9 kb. Possuem longas repetições terminais

diretas que flanqueiam uma região central contendo de um a três módulos abertos de leitura (ORFs). A primeira ORF constitui-se no gene *gag* que especifica uma poliproteína processada em três proteínas maduras: matriz, capsídeo e nucleocapsídeo, que correspondem aos componentes virais do capsídeo. A segunda ORF corresponde ao gene *pol* que codifica todas as proteínas requeridas para transposição, incluindo protease (Pr), transcriptase reversa (RT), RNaseH e integrase (Int). A terceira ORF, presente em algumas famílias deste grupo, corresponde ao gene *env* que codifica a proteína do envelope em retrovírus (revisão em Capy et al. 1998).

As LTRs destes elementos contém todos os elementos regulatórios requeridos para o início da transcrição e poliadenilação do RNA.

Os retrotransposons com LTRs tem sido subdivididos em dois grandes grupos (subclasses) com base no arranjo das seqüências codificantes dentro da segunda ORF. As duas subclasses são: *Ty1-copia* e *gypsy-Ty3*. No primeiro grupo, a ordem na ORF2 é 5' Pr-Int-RT-RnaseH 3' e, no segundo, temos 5' Pr-RT-RnaseH-Int 3'.

O grupo *gypsy-Ty3* pode, ainda, ser dividido em dois subgrupos (duas superfamílias): aquele que contém famílias de TEs possuindo a terceira ORF (superfamília *gypsy*) e aquele em que o gene *env* está ausente (superfamília *Ty3*). Alguns autores distinguem ainda uma terceira superfamília, *Tf1*, encontrada em fungos, que não utiliza o iniciador de tRNA presente nas outras subfamílias para a transcrição reversa.

Os elementos que codificam a proteína do envelope ativa podem ser considerados retrovírus endógenos como é o caso de *gypsy*. No entanto, análises filogenéticas baseadas em seqüências da transcriptase reversa e

integrase agrupam este elemento com os demais retrotransposons com LTRs. É por esta razão, que elementos da família *gypsy* são considerados intermediários entre retrovirus e retrotransposons com LTRs (Lerat e Capy 1999).

Estima-se que os retrotransposons constituem em pelo menos 2% do genoma de *Drosophila* e mais do que 40% do genoma de certas plantas (Capy et al. 1998).

RETROTRANSPOSONS SEM LTRs ou RETROPOSONS

Esta subclasse tem sido dividida em duas superfamílias: LINEs (longos elementos nucleares dispersos) e SINEs (curtos elementos nucleares dispersos). É característica deste grupo a falta de repetições terminais, diretas ou invertidas. LINEs e SINEs têm sido identificados em plantas, fungos e animais. O número de cópias pode ser muito grande, ultrapassando as 100.000, no caso de SINEs e 10.000 para LINEs, em eucariotos superiores. Elementos LINE 1 chegam a representar mais do que 15% do genoma humano (Smit e Riggs 1996).

A Superfamília LINEs possui uma seqüência rica em adeninas no seu término 3' e seu tamanho varia de 5-8 kb. Nestes elementos, observamos, apenas, alguns genes, tais como os que codificam para transcriptase reversa e RNaseH. A família *L1* em mamíferos, as famílias *I*, *F*, *TART*, *Het-A* e *jockey* em *Drosophila*, *Del2* e *Cin4* em plantas e *Tad* em fungos são bons exemplos deste grupo. Observamos uma especificidade insercional dos elementos deste grupo. Desta forma, os elementos *R1* e *R2* de insetos inserem-se em genes ribossomais e *Het-A* e *TART* são encontrados nas pontas dos cromossomos de *Drosophila melanogaster*, mantendo sua integridade telomérica (revisão em Capy et al. 1998).

A Superfamília dos SINEs compreende elementos com tamanho menor do que 500 pb. Não são encontradas seqüências correspondentes ao gene para transcriptase reversa, sendo a maioria destes elementos derivados de tRNA. Assim, sua estrutura é constituída por uma região homóloga a um tRNA, uma região não relacionada à tRNA e um final 3' rico em adeninas e timinas. Exceção a esta regra de homologia à tRNAs é a família *Alu* em primatas e a família *B1* em roedores, que parecem ter se originado do RNA 7SL (Capy et al. 1998).

Os SINEs não apresentam ORFs, portanto, não produzem suas próprias proteínas. Devido à homologia encontrada entre os terminos 3' de LINEs e SINEs presentes no mesmo genoma, tem sido sugerido que os primeiros supram as necessidades dos segundos quanto às proteínas requeridas para sua mobilidade (Okada e Hamada 1997).

ELEMENTOS CLASSE II

Todos os elementos desta classe contém repetições invertidas terminais (ITRs) de 10 a 500 pb de tamanho, flanqueando uma região central que codifica para uma enzima transposase que cataliza a transposição. Estes elementos transponíveis constituem em uma classe mais heterogênea do que a da Classe I. São divididos em várias famílias de acordo com a seqüência da transposase, o tamanho e a composição das ITRs, bem como a natureza das duplicações que produzem no DNA alvo. Exemplos desta classe são: *Ac/Ds* no milho, *P* e *mariner* em *Drosophila*, *Tc1* em *Caenorhabditis elegans* e *Fot1* em fungos.

Verificamos em termos classificatórios, a formação de duas subclasses dentro da Classe II em função da estrutura da transposase (presença ou ausência da assinatura DDE).

Dentro das subclasses, famílias destes elementos se agrupam em superfamílias por identidade de motivos e seqüências nucleotídicas. Algumas destas superfamílias são: *hAT*, *mariner-Tc1*, *pogo-Fot1*, *P*, *Mu* e *CACTA* ou *En/Spm*.

A Figura 1 apresenta um esquema dos diferentes níveis classificatórios de elementos transponíveis, bem como os critérios subjacentes, de acordo com Capy et al. (1998).

ELEMENTOS CLASSE III

Vários elementos são difíceis de classificar, pois apresentam características incomuns e mecanismos de transposição desconhecidos. Estes elementos constituem-se em duas superfamílias: MITEs e *Foldback*. Estas são diferenciadas pelo tamanho das suas repetições terminais, curtas nos primeiros e longas no segundo grupo.

Os elementos do tipo *Foldback* carregam amplas ITRs e são, assim, chamados, pois suas repetições terminais formam estruturas do tipo grampo. Foram, primeiramente, identificados em *Drosophila*, mas já têm sido detectados em outros organismos como ouriços-do-mar e *Caenorhabditis elegans* (Capy et al. 1998).

A estrutura interna dos elementos MITEs (miniatura de elementos transponíveis com repetições invertidas) segue a classe II de elementos transponíveis, mas seu alto número de cópias e sua preferência insercional

Figura 1 no Anexo.

lembra elementos da Classe I. São pequenos e possuem ITRs imperfeitas. Têm sido detectados em uma ampla variedade de plantas, geralmente flanqueando as extremidades de genes (Capy et al. 1998).

MOVIMENTAÇÃO DE ELEMENTOS TRANSPONÍVEIS

Os elementos transponíveis podem transpôr-se de dois modos: de uma forma conservativa ou de uma replicativa. Na primeira, observamos a manutenção do número de cópias do elemento, sendo o mecanismo usual dos elementos de Classe II que se transpõem via um intermediário de DNA. O mecanismo subjacente à forma conservativa envolve a excisão do sítio original e a reinserção da seqüência num novo sítio. O modo de transposição replicativo leva a um aumento do número de cópias. De um modo geral, neste último mecanismo, o elemento original faz uma cópia de si mesmo e esta é inserida num novo sítio. Este último é o mecanismo típico dos retrotransposons. Um aumento no número de cópias, também, é observado para os elementos Classe II por intermédio de eventos de recombinação e reparação.

CONSEQÜÊNCIAS E IMPLICAÇÕES EVOLUTIVAS DE SUA PRESENÇA E MOBILIDADE

Elementos Transponíveis despertam profundas questões sobre sua relação com o genoma hospedeiro e sua relevância para a evolução dos seres vivos. As discussões e opiniões vão desde a rotulagem destas seqüências como parasitas extremos, DNA lixo ou simbiontes (revisão em Kidwell e Lisch 2001). Na revisão citada, os autores apontam vários aspectos a favor de uma

relação elemento-hospedeiro dinâmica e íntima em que qualquer distinção entre os dois é sem sentido. Os autores colocam, ainda, que somente um elemento invasor recente tem fortes tendências parasíticas. Neste caso, se fortes defesas do hospedeiro surgem, uma “guerra de braço” pode se estabelecer. No entanto, se isto não ocorre, o elemento é transmitido verticalmente, e uma ampla associação TE-hospedeiro ou uma coevolução em direção a um mutualismo evolui. Esta coevolução pode estabelecer-se de três formas: 1) minimização dos efeitos negativos pelo TE através de autorregulação; 2) evolução dos mecanismos de defesa do hospedeiro; 3) recrutamento de funções do elemento ou de partes de sua seqüência. A coevolução é dita existir quando observamos por exemplo, regiões regulatórias de genes do hospedeiro tendo incorporado seqüências de elementos transponíveis, assim como regiões codificantes dos transposons terem evoluído de genes funcionais do hospedeiro.

O impacto de TEs no genoma hospedeiro está relacionado com o aumento da variabilidade genética que eles produzem. Este aumento na variação genética é devido à sua movimentação (transposição) ou à recombinação homóloga entre cópias de elementos de uma mesma família. Como resultado destes fatores, temos mudanças que vão desde mutações pontuais até modificações no tamanho e no arranjo de genomas inteiros. Assim, estas mudanças dão-se tanto na estrutura (rearranjos cromossômicos) como na função (organização e mudança na expressão de genes) do genoma.

A grande variação no tamanho do genoma entre os seres vivos, principalmente plantas e invertebrados, conhecido como o paradoxo do valor

C, tem sido explicada, em parte, por amplificação de elementos transponíveis (Gregory e Hebert 1999; Fedoroff 2000; Petrov 2001).

Uma fração significativa das doenças humanas (0,3%) está relacionada com recombinação homóloga desigual de elementos transponíveis (Deininger e Batzer 1999). Pelo menos 1% de seqüências humanas foi transportado de um ponto a outro no genoma por intermédio do movimento de TEs (Pickeral et al. 2000).

O possível papel dos TEs para o surgimento de mecanismos de supressão de sua atividade é explorado por Hurst e Werren (2001). Exemplos destes mecanismos incluem fenômenos mutacionais induzidos por repetição, metilação, interferência de RNA, entre outros. Muitos destes eventos foram, recentemente descobertos e, ainda, permanecem a ser totalmente elucidados. Interferência de RNA (RNAi) é um bom exemplo. Este mecanismo consiste em um silenciamento da expressão gênica por degradação do mRNA em resposta à RNA dupla fita, dependente de homologia. O mesmo já foi observado em vários organismos, e tem sido implicado no controle de infecção viral. Assim, tais fenômenos teriam surgido para defender o genoma contra DNA parasítico como vírus e elementos transponíveis (Plasterk e Ketting 2000).

A relação entre evolução do genoma e presença de TEs, envolve, ainda, mecanismos de manutenção de telômeros, rearranjos de cromossomos, mecanismos de recombinação somática, variabilidade do sistema imune, entre outros (Kidwell e Lisch 2000).

Outro aspecto interessante é a relação de TEs com a origem de microssatélites. Recentemente, um trabalho de Wilder e Hollocher (2001)

demonstrou que uma classe de microssatélites é originada em regiões internas de seqüências altamente repetidas no genoma de dípteros, e contém aspectos característicos de elementos transponíveis. Estes elementos chamados *mini-me* estão presentes em várias espécies de dípteros, drosofilídeos e não-drosofilídeos, e representam 1,2% do genoma eucromático de *Drosophila melanogaster*.

Kidwell e Lisch (2001), revisando o papel dos TEs na evolução de genomas hospedeiros, mencionam que muitos autores tem especulado sobre o papel dos elementos transponíveis durante o surgimento e a evolução da vida na terra. Neste contexto, a transposição teria sido um mecanismo importante para a evolução de proteínas (Dwyer 1998). Além disso, tem sido discutido que transposons estariam envolvidos com o estabelecimento de um ciclo sexual nos organismos eucariotos (Arkipova e Meselson 2000; Wright e Finnegan 2001). Os elementos móveis estão implicados, ainda, no surgimento de eventos macroevolutivos como a origem dos eucariotos e dos vertebrados pelo seu envolvimento na evolução da formação da cromatina e no processo de metilação, o que possibilitou a expansão do número de genes (Bird 1995; Fedoroff 2000).

Alguns autores sugerem que íntrons nucleares evoluíram de transposons que faziam seu próprio processamento. Estes íntrons seriam, então, os elementos móveis mais ancestrais e adaptados, estando sua capacidade de processamento associada a uma forma de minimização dos efeitos deletérios na expressão de proteínas interrompidas (Kidwell e Lisch 2001).

O papel dos TEs no rearranjo de seqüências genômicas tem sido bem documentado em vários estudos. Tem sido reportada a relação entre pontos

de quebra para inversões cromossômicas em *Drosophila* e presença de elementos transponíveis nestes locais (Lim 1988; Tsubota et al. 1989; Regner et al. 1996; Cáceres et al. 1999; Cáceres et al. 2001), o que sugere uma relação de causa e efeito. Inversões podem levar à ruptura de genes, se estes encontrarem-se nos pontos de quebra, à alteração no padrão de segregação de genes localizados dentro da inversão e à mudança na expressão destes genes devido ao efeito de posição.

TEs podem ser inseridos dentro de genes ou em suas regiões regulatórias, o que leva a mudanças na expressão destes genes ou à sua inativação. O potencial de TEs como modificadores da expressão gênica tem sido investigado e, em plantas, já foram relatados mais de 100 genes que possuem partes de elementos transponíveis ou vestígios deles em suas seqüências regulatórias (White et al. 1994; Capy et al. 1998).

A maioria das mutações produzidas por TEs é neutra ou deletéria, mas estudos já reportaram contribuições destas mudanças para seus hospedeiros (Shrimpton et al. 1990). Para *D. melanogaster* foi evidenciado que mutações induzidas por TEs afetam características adaptativas de modo a tornar estas moscas mais longevas e mais resistentes a estresses ambientais (Lin et al. 1998).

DINÂMICA DOS ELEMENTOS TRANSPONÍVEIS

Compreender a dinâmica de um TE dentro de um genoma hospedeiro é compreender como estas seqüências surgem, permanecem, são reguladas e evoluem em meio às pressões seletivas a que são submetidas. Muitas questões permanecem abertas sobre estes tópicos, mas algum panorama geral tem sido montado.

A origem dos TEs dentro do genoma de uma espécie, acredita-se, dá-se por rearranjo de seqüências genômicas, por transmissão horizontal ou ainda por transmissão vertical. Uma vez estabelecido dentro do genoma, o primeiro passo é caracterizado por uma fase invasora em que o número de cópias aumenta. A taxa, na qual este fenômeno ocorre, depende de vários fatores que regulam a transposição de um elemento transponível. São eles: o próprio elemento, o genoma hospedeiro, fatores ambientais e interação entre todos estes parâmetros. O segundo passo seria a inativação ou imobilização do TE que depende, em grande parte, do seu impacto sobre o valor adaptativo do hospedeiro. Assim, TEs podem ser mantidos em estado de latência ou perdidos estocasticamente (Pinsker et al. 2001).

ORIGEM E EXPANSÃO DOS ELEMENTOS TRANSPONÍVEIS NO GENOMA HOSPEDEIRO

A aquisição de um TE já existente por uma espécie, dá-se por transmissão vertical a partir de um ancestral, ou por transferência horizontal, através de introgressão ou entre espécies distantemente relacionadas, via um vetor.

A primeira aparição de um elemento transponível dentro de um genoma é produzida por vários mecanismos complexos a partir de seqüências genômicas. Um bom exemplo disto são os TEs da superfamília SINEs. A família *Alu* de humanos e *B1* de roedores, representantes deste grupo, originaram-se de uma seqüência de RNA 7SL por retrotransposição (Capy et al. 1998). Parece que a evolução destas seqüências tem ocorrido, predominantemente, por aquisição serial de domínios onde processos reversos de deleções ou

perda de domínios também podem ocorrer. Integrases, transposases, assim como transcriptases reversas apresentam domínios montados em eucariotos, mas que evoluíram de bactérias, levando a retrotransposons e, então, a retrovírus pela aquisição do gene *env* (Capy et al. 1998; Kidwell e Lisch 2001).

Em um modelo proposto por Wright e Finnegan (2001), elementos transponíveis são incapazes de espalhar-se para outras linhagens em uma população assexual (a não ser que assumíssemos a ocorrência de transmissão horizontal), sendo perdidos por deriva genética ou seleção purificadora. Em uma população sexual, exocruzamento e recombinação permitiriam que estes elementos se expandissem ao longo da população. Esta consideração é também reforçada no trabalho de Arkipova e Meselson (2000), onde os autores investigaram superfamílias dos principais grupos de TEs, na maioria dos filos animais, e observaram que táxons que abandonaram a reprodução sexual, há muito tempo atrás, são desprovidos de retrotransposons.

Em uma revisão de Hurst e Werren (2001) sobre o papel dos elementos transponíveis na evolução eucariótica, os autores discutem a importância destas seqüências para a origem e evolução dos seus hospedeiros. Em primeiro lugar, os autores consideram os TEs como uma das classes de elementos genéticos egoístas ou parasitas genômicos, o que pressupõe que este tipo de seqüência vise sua expansão, somente trazendo danos para o organismo hospedeiro. Os autores estabelecem duas regras básicas de convivência entre TEs e o genoma hospedeiro. Uma destas duas regras relata que a diversidade e a expansão destes elementos estão correlacionadas com a taxa de exocruzamento de uma espécie, o que prevê que estas seqüências devem ser menos comuns em taxons endocruzados. A segunda regra diz que

estas seqüências tendem a expandir-se em direção à fixação em uma população. Neste contexto, estas seqüências são reguladas e sua manifestação restringe-se à produção de híbridos entre populações distintas. Também, por causa desta propriedade, TEs são especulados como fatores de isolamento reprodutivo e possível especiação (Ginzburg et al. 1984). Eventos de especiação também têm sido postulados como o resultado das mutações induzidas por TEs levando uma população a escapar de sua estase evolutiva (Dimitri e Junakovic 1999; Kidwell e Lisch 2001).

Vários elementos transponíveis mostram-se capazes de expansão por eventos de transmissão horizontal. Em função desta capacidade, TEs têm sido, amplamente, estudados como vetores de transformação desde o início da década de 80 (Carareto et al. 1997; Sherman et al. 1998; Plasterk et al. 1999). Nos últimos dez anos, tem se conseguido utilizar transposons de *Drosophila* para transformar outros insetos. Numa revisão, Handler (2001) traçou um panorama desta área, salientando a importância de um sistema de estudo como o elemento *P* para a espécie padrão *Drosophila melanogaster*. O autor revisa outros TEs como potenciais vetores para transformação, principalmente transposons propriamente ditos, bem como as dificuldades encontradas em conseguir transformação em uma maior amplitude de organismos utilizando mecanismos de controle confiáveis e taxas razoáveis de sucesso.

Após a origem de um TE no genoma, ocorre o aumento no seu número de cópias, levando a uma diversidade de sítios de inserção. Neste momento, interações entre TE/genoma hospedeiro são muito importantes para a sobrevivência de ambos. Muitos elementos móveis podem produzir mutações letais levando à sua própria extinção junto com a do seu hospedeiro. Desta

forma, um equilíbrio é necessário para a coexistência destas seqüências. Tal equilíbrio é resultado de uma complexa regulação que começa com uma auto-regulação dos próprios TEs.

REGULAÇÃO DE ELEMENTOS TRANSPONÍVEIS

Elementos transponíveis possuem suas próprias seqüências regulatórias que modulam sua transcrição. Além disso, outros mecanismos são conhecidos. Para elementos de Classe II, mecanismos de competição envolvendo cópias completas e deletadas são observados. Ainda neste grupo, temos o exemplo bem documentado do elemento *P*. Para este TE, a regulação tem um efeito materno e sua transposição é limitada na linhagem germinal por um processamento alternativo do mRNA gerador da transposase. Este processo é regulado pelo hospedeiro e herdado maternalmente. Em *mariner*, observamos uma regulação em que o excesso de transposase inibe sua atividade devido a um fenômeno de multimerização. Interações RNA-RNA “anti-sense” podem ainda regular a transposição de TEs, como visto para *micropia* e *P* em *Drosophila*, assim como para elementos móveis do milho (Hartl et al. 1997; Capy et al. 1998).

Existem bons exemplos de fatores do hospedeiro regulando as seqüências de TEs. Em *Drosophila*, um exemplo clássico é o do gene *flamenco* regulando a atividade do retrotransposon *gypsy*. Outro exemplo é a de proteínas do hospedeiro (IRBP e PSI) que interagem com o elemento *P* para sua regulação somática (Capy et al. 1998).

Para a maioria dos TEs, a transposição é regulada para ocorrer somente nas linhagens germinais. Esta restrição parece ser uma estratégia altamente

eficiente para garantir sua transmissão para a próxima geração, minimizando, ao mesmo tempo, os danos ao hospedeiro (Kidwell e Lisch 2001).

A partir de sua invasão no genoma, a dinâmica e evolução de um TE depende de fatores permissivos do hospedeiro. Assim, diferenças interpopulacionais e inter-específicas para elementos transponíveis são, em grande parte, devido a estes fatores.

Com exceção do elemento *P* de *Drosophila*, elementos classe II parecem ser menos dependentes de fatores do hospedeiro do que os demais. Por este motivo, postula-se que estas seqüências sejam capazes de invadir uma gama maior de hospedeiros. Este é o caso do elemento *mariner* (Robertson e Lampe 1995).

Retrotransposons com LTRs em *Drosophila* podem ter sua transcrição modulada por hormônios esteróides de insetos. Tal tipo de regulação foi documentada para *412*, *micropia* e *Ulysses* (Capy et al. 1998).

Mecanismos de inativação de seqüências repetidas como é a metilação e interferência de RNA são reportados para vários organismos (Hurst e Werren 2001). Possivelmente, este mecanismo parece ter evoluído para prevenir rearranjos cromossomais devido a recombinações ectópicas entre seqüências homólogas repetidas. Outros mecanismos epigenéticos como remodelagem da cromatina também parecem ser utilizados na regulação de elementos transponíveis (Kidwell e Lisch 2001).

TEs são regulados de uma forma bastante refinada e, quando ocorre um escape deste controle, fenômenos visíveis podem surgir. Tais fenômenos têm sido bem documentados para *Drosophila melanogaster*, no caso dos elementos *I*, *hobo* e *P*. Nestes casos, observamos o desenvolvimento de uma

síndrome que é chamada de síndrome da disgenesia do híbrido (Kidwell e Kidwell 1976). Esta é caracterizada por um quadro de características anormais correlacionadas geneticamente, induzidas em híbridos produzidos por cruzamentos específicos, onde machos e fêmeas diferem quanto à presença/ausência destes elementos (usualmente, isto ocorre quando o cruzamento é obtido a partir de fêmeas desprovidas do elemento com machos contendo o elemento).

O fenômeno da disgenesia do híbrido parece estar relacionado com uma invasão recente do elemento transponível no genoma. Este é o caso dos elementos acima citados que invadiram o genoma de *D. melanogaster* nos últimos 50 anos. Esta síndrome, provavelmente, desaparece quando o equilíbrio entre habilidade transposicional e regulação é alcançado (Capy et al. 1998).

A taxa de transposição de um dado elemento transponível contrapõe-se a fatores que o regulam (próprios ou do hospedeiro) e o resultado é a manutenção destas seqüência em equilíbrio. Muitas vezes, este equilíbrio pode ser quebrado por fatores ambientais.

FATORES AMBIENTAIS INFLUENCIANDO A ATIVIDADE DOS ELEMENTOS TRANSPONÍVEIS

Capy et al (2000) revisam a relação entre ativação de elementos transponíveis e estresse. TEs podem transpôr-se em resposta a sinais ambientais externos. Este fenômeno já foi identificado para vários transposons, em várias espécies. Diferentes populações e linhagens parecem diferir nesta resposta, e esta variação parece estar relacionada a um "background" genômico. Os

autores postulam que a resposta dos TEs pode ser devido à presença de sítios de fixação para ativadores de transcrição em suas regiões regulatórias. Isto já foi observado para alguns TEs, como *copia* e *mariner*. Nestes elementos, encontramos regiões em suas seqüências muito similares àquelas de promotores de "hsp's", conhecidas proteínas que respondem a choque de temperatura. Quanto à origem destas similaridades, os autores discutem que pode ser um caso de domesticação de TEs ou que, ao contrário, os elementos móveis adquiriram estas seqüências do genoma hospedeiro.

Em plantas, TEs podem ser ativados por fatores bióticos e abióticos tais como estresse químico e físico, assim como ataques de parasitas. Esta resposta parece estar correlacionada com a expressão de genes de estresse em plantas (Wessler 1996; Grandbastien 1998).

O estresse de temperatura tem se mostrado relacionado, pelo menos, indiretamente, com a atividade de elementos transponíveis. Vários ensaios têm sido delineados para a checagem deste fator, mas conclusões definitivas têm se mostrado difíceis. Este fator parece afetar vários TEs favorecendo, inclusive, fenômenos de disgenesia do híbrido (Strand e McDonald 1985; Junakovic et al. 1986).

A teoria evolutiva que explica a resposta de TEs a fatores estressantes considera que uma taxa de transposição alta nestas condições tem evoluído por garantir a permanência de uma cópia ativa do elemento na prole do indivíduo estressado. Por outro lado, para o hospedeiro, isto também representaria uma vantagem evolutiva, à medida em que sua prole tem uma variabilidade aumentada no momento em que esta propriedade é necessária para adaptações a novas condições ambientais (Kidwell e Lisch 2001). Além

disto, ter seqüências no genoma que se expressam condicionalmente, representa uma alternativa de regulação importante que pode ser utilizada pelo hospedeiro (Li et al. 1999).

DESTINO DOS ELEMENTOS TRANSPONÍVEIS NO GENOMA HOSPEDEIRO

Vários mecanismos podem levar à perda de um elemento transponível de um genoma. Isto pode acontecer por mudanças em sua seqüência nucleotídica, através do acúmulo de mutações. TEs podem, também, ser perdidos por deriva genética ou seleção, em função dos seus efeitos deletérios para o hospedeiro.

Além destes destinos, um TE pode ser “recrutado” pelo genoma, um fenômeno conhecido como “domesticação” de elementos transponíveis. Este processo é facilmente explicável se pensarmos que qualquer função do hospedeiro que necessite de quebra e junção de seqüências pode valer-se da maquinaria enzimática especializada para realizá-la. Esta forma de recrutamento pode se dar em vários níveis: 1) regulatória; 2) estrutural; 3) enzimática e 4) novas seqüências codificantes. Vários achados têm sido documentados na literatura, principalmente, para retrotransposons com LTRs. Esta domesticação pode refletir novos padrões de expressão impostos pelo TE ou pelo “uso” de seqüências internas codificantes do elemento (Britten 1997).

Um bom exemplo de domesticação de elementos transponíveis é o que ocorre com *Het-A* e *TART* de *D. melanogaster*, dois retrotransposons que funcionam como telomerasas, mantendo o telômero dos cromossomos desta espécie. Em função desta descoberta, especula-se a origem de telomerasas a partir da domesticação de antigos TEs (Pardue e DeBaryshe 1999).

Os elementos transponíveis assumem funções estruturais, à medida em que formam grande parte de regiões específicas do genoma como centrômeros e telômeros. Além disto, uma proteína conhecida como CENP-B, altamente conservada nos mais diferentes organismos desde leveduras até humanos, que se liga a motivos repetidos nos centrômeros e tem uma função na segregação cromossômica durante a divisão celular, é relacionada a uma ampla e ancestral família de TE (Smit e Riggs 1996; Kidwell e Lisch 2001).

Em *Saccharomyces cerevisiae*, o reparo do tipo quebra cromossômica de fita dupla parece ser dirigido por transcriptase reversa de retrotransposons (Moore e Haber 1996; Teng et al. 1996; Britten 1997).

A diversidade observada em receptores das células do sistema imune em vertebrados é adquirida pela ação de duas enzimas através de um processo de rearranjos de seqüências gênicas, processo este que é muito semelhante ao da transposição. O que parece possível é que transposons teriam abastecido o genoma dos vertebrados com funções enzimáticas necessárias para atingir este grau de hipermutabilidade somática (Smit 1999; Kidwell e Lisch 2001). Ainda, com relação ao sistema imune, a presença de várias famílias de elementos transponíveis dentro do loco do MHC aponta para uma provável importante função da mobilidade destas seqüências na geração de diversidade desta região (Gaudieri et al. 1997).

Outro exemplo é o da aquisição, por camundongos, de resistência retroviral por domesticação de um gene que teve sua origem no genoma de uma família de retrovírus (Benit et al. 1997).

Existem, ainda, fortes evidências de que seqüências retrovirais endógenas expressem-se e atuem como imunossupressores em tecidos da

placenta, evitando uma resposta imune ao feto em animais vivíparos. Desta forma, tem sido proposto que alguns retroelementos em mamíferos tenham sido recrutados para tornar possível a viviparidade (Villarreal 1997; Denner 1998). Assim, uma função que inicialmente evoluiu nestas seqüências invasoras para driblar a defesa de seu hospedeiro é domesticada.

Em plantas, seqüências degeneradas de retrotransposons são encontradas em cerca de 21 genes e desempenham importantes papéis na sua regulação (White et al. 1994).

Um exemplo recente de um TE tornando-se parte integral da regulação de um gene é aquele observado no gene humano da apolipoproteína, onde o "enhancer" deste gene encontra-se dentro de um elemento transponível da superfamília dos LINEs (Yang et al. 1998). Em humanos, um trabalho recente (Nekrutenko e Li 2001) analisou mais de 13.000 genes para a inserção de elementos transponíveis e encontrou em 4% destas regiões codificantes TEs ou fragmentos deles. Os autores observaram uma alta taxa de localização de seqüências homólogas a TEs em éxons. Estes achados são explicados, provavelmente, pelos sítios de processamento existentes nos TEs. Desta forma, inserções de elementos transponíveis podem estar intimamente relacionadas com uma alta freqüência de processamento alternativo em genes humanos. Além disto, TEs fixados em regiões gênicas diferentes, em populações ou espécies, podem levar a uma aceleração da evolução gênica e provocar rápida divergência entre táxons. Neste contexto, os autores especulam o papel dos TEs nas diferenças entre homens e chimpanzés.

O número médio de cópias de TEs num genoma reflete a ação de forças antagônicas. De um lado, transposição e seleção positiva aumentam o

número de cópias. De outro, excisão, seleção negativa e perda por deriva genética ou acúmulo de mutações, diminuem o número de cópias. O modelo que dá conta destes parâmetros explica como elementos transponíveis permanecem num genoma, embora aparentemente sem nenhum efeito vantajoso para seu hospedeiro. A demonstração de que TEs podem expandir-se no genoma hospedeiro, a despeito da perda adaptativa, tem sido explicada como um exemplo de seleção natural agindo em direções opostas em níveis diferentes da hierarquia biológica (DNA e organismo), onde a seleção a nível do hospedeiro para reduzir a transposição é mais fraca se comparada com a seleção a nível do TE para aumento da transposição (Biémont 1992; Kidwell e Lisch 2001).

FILOGENIA E EVOLUÇÃO DE ELEMENTOS TRANSPONÍVEIS

Filogenias obtidas de elementos transponíveis são, freqüentemente, comparadas com filogenias de suas espécies hospedeiras, na tentativa de entender a evolução destas seqüências móveis. As congruências entre as duas filogenias são encaradas como indicativos de transmissão vertical dos elementos transponíveis. No entanto, quase todas estas comparações levam a discrepâncias que podem ser explicadas por várias hipóteses.

Uma destas hipóteses é a de transmissão horizontal de TEs. Este fenômeno assume uma transmissão assexual de material genético entre espécies, distantemente relacionadas. As diferenças encontradas também podem evocar transmissão vertical como explicação, desde que, pelo menos, uma das seguintes premissas esteja presente: polimorfismo ancestral, taxas de

evolução de TEs diferentes entre hospedeiros e/ou amostragem aleatória durante o processo de especiação (Capy et al. 1994).

Um dos exemplos mais interessantes e consistentes de divergência entre filogenia de TEs e de suas espécies hospedeiras é o caso do elemento *P*. Da análise dos dados, surgem vários argumentos para um evento de transmissão horizontal do elemento *P* de *D. willistoni* para *D. melanogaster*. Os elementos *P* de *D. melanogaster* e *D. willistoni* diferem em apenas um nucleotídeo, fato completamente inconsistente se levarmos em consideração os estimados 50 milhões de anos que separam estas duas espécies. Outro interessante aspecto é o fato de que as demais espécies do grupo *melanogaster* não possuem estas seqüências. Além disto, neste caso específico, *D. melanogaster* adquiriu este elemento por volta de 1950, já que linhagens mantidas em laboratório antes desta época são desprovidas de *P*, e aquelas coletadas posteriormente já encontram-se com esta seqüência. Assim, este parece ser um caso bem documentado de transmissão horizontal, onde a extensa representatividade de coletas deste organismo modelo para estudos evolutivos permitiu o acompanhamento da invasão do elemento *P* nesta espécie. Outros eventos de transmissão horizontal são utilizados para explicar outros pontos inconsistentes entre as filogenias obtidas para este elemento, embora não descartem as hipóteses alternativas.

Casos de transmissão horizontal também são reportados para o elemento *mariner* e explicam como elementos tão similares existem no genoma de espécies distantemente relacionadas (Robertson e MacLeod 1993). Contudo, também existem evidências de transmissão vertical e perda

estocástica nos diferentes táxons analisados para este elemento (Lohe et al.1995).

Para a superfamília *Ty1-copia*, a transmissão vertical parece prevalecer na sua história evolutiva. As análises das filogenias em questão distinguem dois grandes grupos que conhecemos com os elementos de metazoários de um lado, e aqueles de plantas Angiospermas de outro. Esta diferença sugere que os elementos deste grupo evoluíram por transmissão vertical, a partir de um ancestral comum, antes da diversificação destes dois ramos. No entanto, quando elementos dentro desta superfamília são analisados de uma forma mais cuidadosa nestes grupos, algumas inconsistências surgem e eventos de transmissão horizontal não podem ser facilmente descartados (Capy et al. 1998).

Stuart-Rogers e Flavell (2001) mostram que várias famílias de retrotransposons da superfamília *Ty1-copia*, dentro do ramo das plantas Gimnospermas, apresentam um padrão de distribuição complexo, onde alguns destes retrotransposons são específicos de alguns gêneros e outros são amplamente distribuídos e conservados entre as Gimnospermas. A taxa de substituição para alguns retroelementos chega a ser menor do que as taxas de genes nucleares considerados conservados nestes genomas. Os autores defendem uma hipótese evolutiva que aponta dois processos dominantes na evolução destes elementos neste grupo: transmissão vertical com forte seleção para conservação da seqüência e extinção de retrotransposons em linhagens particulares. É ressaltado, neste trabalho, que os dados não excluem a possibilidade de transmissão horizontal, mas este mecanismo, além de ser desnecessário, é menos parcimonioso. Os autores defendem que este grau de

conservação é característico de todos retrotransposons eucariotos, de animais e vegetais. Eles alegam que as taxas evolutivas observadas para TEs em Gimnospermas não são evidenciadas para outros táxons pela dificuldade de se obter seqüências não parálogas ou homólogas. Nas Angiospermas, isto pode ser explicado pela história evolutiva do genoma deste grupo, caracterizado por uma grande instabilidade, o que propicia a extinção de linhagens de TEs.

Vários trabalhos sobre a evolução dos elementos *R1* e *R2* têm mostrado que eles são verticalmente transmitidos no gênero *Drosophila*, evoluindo a taxas constantes (Eickbush et al. 1995). Surpreendentemente, este que é um dos exemplos mais sólidos de transmissão vertical foi re-analisado num trabalho de Gentile et al. (2001) demonstrando que a coexistência de linhagens divergentes na região de rDNA em algumas espécies de *Drosophila*, bem como discrepâncias entre TEs a nível de grupo de espécies, têm se mostrado difíceis de explicar à luz da evolução em concerto existente nesta região.

Blesa et al. (2001) analisaram seqüências homólogas ao elemento *bilbo* (um retrotransposon sem-LTRs de *Drosophila subobscura*) amplificadas a partir do genoma de sete espécies do grupo *obscura*, seguindo três subgrupos. As relações evolutivas que emergiram desta análise indicaram um padrão de transmissão vertical destas seqüências, neste grupo de espécies. Quando estes dados são analisados à luz dos resultados de *Southern blot*, um cenário mais amplo pode ser visualizado apontando para a presença de elementos *bilbo* em ramos basais da família Drosophilidae. Os dados indicam, ainda, que um ancestral desta família de elementos móveis esteve presente no início da radiação *Scaptodrosophila*.

Em bactérias, eventos de transmissão horizontal entre linhagens e, até mesmo, entre espécies, explicam as inconsistências obtidas nas análises de diferentes subfamílias e famílias de elementos IS (Lawrence e Hartl 1992).

É claro que devemos ter cuidado ao concluir pela existência de transmissão horizontal ou não. Capy et al. (1998) coloca pelo menos quatro situações que devem ser consideradas ao levantarmos a hipótese de transmissão horizontal. São elas: 1) inconsistência entre as filogenias (do TE e do hospedeiro); 2) motivos conservados entre elementos de espécies muito distantes; 3) distribuição não homogênea de um elemento entre espécies; 4) invasão recente de um genoma por um elemento. Na primeira situação, é necessária a existência de uma diferença estatisticamente significativa entre as filogenias obtidas, o que pressupõe uma filogenia de espécies bem robusta. Quanto à existência de regiões conservadas, é fundamental uma análise mais profunda para testar se esta conservação não é restrita a pequenas regiões codificantes conservadas dentro da superfamília ou até mesmo entre elementos de mesma classe. Padrões onde algumas espécies, no meio da filogenia, não apresentam sinais de determinados elementos devem ser confirmados utilizando-se vários "primers" diferentes para amplificação ou outras metodologias, e podem ser explicados, ainda, por perda estocástica. Atenção especial deve ser dada para ramos inteiros onde uma ou poucas espécies, posicionadas internamente, possuem o elemento considerado, o que implica que vários eventos independentes de perda estocástica devem ser utilizados para explicar o que um único evento de transmissão horizontal resolveria, algo, parcimoniosamente, indesejável.

Dados disponíveis para a sugestão de invasão recente são escassos pois requerem uma análise extensa de muitas populações naturais de vários pontos da distribuição geográfica de uma dada espécie, antes, durante, e após a invasão, o que é bem documentado apenas no caso do elemento *P* em *D. melanogaster*.

A análise cuidadosa de todas estas situações permite-nos inferir sobre casos de transmissão horizontal. No entanto, outras hipóteses não podem ser descartadas e sabemos, ainda, que as alternativas não são mutuamente exclusivas.

Em uma outra abordagem, filogenias de seqüências de diferentes elementos transponíveis são analisadas em uma tentativa de traçar a história evolutiva e as relações de parentesco existentes entre as muitas famílias de TEs. A análise de elementos transponíveis Classe I e Classe II demonstra domínios comuns. Tais achados levam à hipótese de uma origem comum entre estas seqüências móveis. A razão desta homologia, detectada entre TEs pertencentes a diferentes classes, pode estar relacionada ao fato desta similaridade residir em regiões que representam centros catalíticos de proteínas ou seqüências importantes para regulação e movimentação do TE.

Entre os retrotransposons, o domínio mais conservado e utilizado em comparações é o da transcriptase reversa. Utilizando esta região, é possível incluir na análise retrovírus, íntrons do grupo II e retrons. Os resultados desta análise demonstraram que retrotransposons com LTRs são claramente mais relacionados com retrovírus, enquanto os retrotransposons sem LTRs agrupam-se com íntrons do grupo II e retrons (Capy et al. 1998).

Comparações entre elementos da Classe I e da Classe II são feitas alinhando domínios da integrase dos retrotransposons e da transposase dos transposons propriamente ditos. Árvores filogenéticas, obtidas deste alinhamento, apresentam retrotransposons com LTRs e retrovírus formando um agrupamento monofilético. Um aspecto interessante deste estudo é que esta relação é similar àquela obtida a partir de análises da transcriptase reversa (Capy et al. 1996).

Xiong e Eickbush (1990) propõem um cenário onde todos os elementos transponíveis contendo transcriptase reversa derivaram de uma seqüência ancestral possuindo um domínio de *gag* e de *pol*. Os retrotransposons com LTRs teriam emergido após a aquisição das LTRs. Estes elementos teriam dado origem aos retrovírus pela aquisição de um gene *env*. O mesmo ancestral pode ter dado origem aos retrotransposons sem LTRs. No entanto, uma análise mais cuidadosa das integrases destes últimos retrotransposons tem demonstrado que elas seguem um sistema diferente daquele visto em integrases dos retrotransposons com LTRs. A partir disto, tem sido sugerido que os retrotransposons sem LTRs formam o grupo mais antigo de retrotransposons que, por sua vez, parecem ter sido originados de introns do grupo II. O que parece mais provável, compilando todos os dados, é que a integrase de retrotransposons com LTRs e de retrovírus e a transposase de elementos classe II com a assinatura DDE (superfamília Mariner-Tc1) divergiram de uma seqüência comum (Capy et al. 1996; Capy et al. 1997).

Uma importante questão sobre evolução de elementos transponíveis diz respeito à direção deste processo, por aquisição ou deleção de domínios. Comparações entre retrotransposons com LTRs e retrovírus indicam que ambas

direções são possíveis. Um retrotransposon com LTRs pode tornar-se um retrovírus pela aquisição de um gene *env* funcional e o inverso também é possível, à medida que ocorre a inativação ou deleção deste gene (Lerat et al. 1999).

Lerat e Capy (1999) compararam proteínas a partir do gene *env* de retrovírus e retrotransposons com LTRs e identificaram dois motivos similares. No entanto, estas seqüências não ocorrem na mesma ordem nos dois grupos, o que leva a duas hipóteses: 1) elas são derivadas de uma mesma seqüência ancestral que passou por aquisições independentes destes domínios; 2) a outra possibilidade envolve uma reorganização desta região após sua aquisição. De qualquer forma, os autores não encontraram nenhum retrovírus endógeno, ou seja, todos retrovírus possuem um padrão de motivos conservados e estes são diferentes nos retrotransposons com LTRs, inclusive em seqüências *gypsy* e "gypsy-like". A análise filogenética da seqüência de aminoácidos destes motivos também distingue retrovírus de retrotransposons com LTRs. Por estarem em regiões distintas na molécula protéica, se possuem alguma função importante, estas funções são diferentes nos dois grupos. Em algumas das análises feitas pelos autores, parece que o ramo *gypsy/yoyo* está mais relacionado com os retrovírus. Se o modelo sugerindo que retrotransposons com LTRs deu origem aos retrovírus for verdadeiro, parece que estas seqüências são as candidatas.

Marín e Lloréns (2000) realizaram uma extensa análise filogenética de elementos da superfamília Ty3/Gypsy a partir de seqüências depositadas no Genbank, incluindo o genoma de *Arabidopsis thaliana* e *Drosophila melanogaster*. Os dados obtidos demonstram que o número de linhagens

incluídas nesta superfamília é bastante variável entre as espécies. Os autores assumem que transmissão horizontal é muito infreqüente ou ausente e, baseando-se nesta premissa, estimam que tenham existido múltiplas (pelo menos 4) linhagens antes da diversificação entre mono e dicotiledôneas. Para o genoma de *D. melanogaster*, os dados indicam a coexistência de cerca de 20 tipos de retrotransposons *Ty3/gypsy*. Este número cai para apenas um tipo de seqüência no genoma de *S. cerevisiae*, 3 linhagens em peixes e nenhuma encontrada em mamíferos. Os autores chegam a propor uma correlação negativa entre o sucesso deste grupo de retrotransposons e o tamanho do genoma, considerando o tamanho reduzido do genoma de *A. thaliana* e *D. melanogaster*.

O RETROTRANSPOSON GYPSY

O retrovírus endógeno *gypsy*, também conhecido como *mdg4*, foi o primeiro identificado em *Drosophila melanogaster*, possui 7,5 kb de comprimento, tendo LTRs com 482 pb (Georgiev et al. 1981; Bayev et al. 1984). Este elemento pertence à superfamília *gypsy*/Ty3 de retrotransposons (Capy et al. 1998). A relação íntima entre este grupo e os retrovírus tem sido bastante discutida (Kim et al. 1994; Pélisson et al. 1997; Lerat e Capy 1999).

Assim como em outros retrovírus, *gypsy* possui um genoma que codifica três genes diferentes chamados de *gag*, *pol* e *env*, que são requeridos para sua replicação e infectividade. Os primeiros dois genes são traduzidos a partir de um único RNA. O produto de *gag* é o precursor de proteínas estruturais do capsídeo viral. O gene *pol* é expresso fusionado à *gag* sendo posteriormente clivado para dar origem à transcriptase reversa e enzimas como protease, RNaseH e integrase. O terceiro gene, localizado na porção 3' do genoma (Figura 2), é expresso como um transcrito processado e codifica os componentes do envelope viral que media a entrada do vírus na célula por interação com receptores celulares seguido de fusão da partícula viral com a membrana da célula (Pélisson et al. 1994).

A ORF 3 encontrada em *gypsy* de *D. melanogaster* é transcrita num mRNA de 2,1 kb de comprimento e prevê uma proteína de 54kD (Pélisson et al. 1994). Este polipeptídeo dá origem, por clivagem em sítios conservados, a uma proteína de superfície de 32 kD e uma proteína transmembrana de 20 kD. Modificações pós-traducionais, bem como clivagens proteolíticas adicionais, são esperadas ocorrer, gerando variação nos tamanhos protéicos (Song et al.

Figura 2 no anexo.

1994). O RNA *env* é obtido através de um processamento do transcrito completo de 7kb inicialmente produzido e envolve uma região 5' que compreende parte da LTR 5' e uma região "ladder" antes da ORF1 (nucleotídeos 238 a 568), com uma região 3' de *gypsy*, cujo sítio para o processamento é o nucleotídeo 5551. Esta última porção se estende até o nucleotídeo 7277 (Avedisov e Ilyin 1994; Pélisson et al. 1994).

A seqüência primária da *env* de *gypsy* de *D. melanogaster* é bastante variável e não mostra similaridades óbvias com as proteínas de retrovírus. No entanto, esta variação é típica desta proteína. Além disto, os domínios esperados são encontrados: um domínio transmembrana perto da região carboxi-terminal, sítios de glicosilação e sítios de clivagem em posições conservadas (Song et al. 1994). Este polipeptídeo é processado e gera uma proteína transmembrana (TM) e um amplo peptídeo (SU) contendo funções como ligação ao receptor de membrana na célula-alvo e sítios antigênicos. A taxa de evolução do gene *env* de retrovírus pode ser muito alta para escapar do reconhecimento, via anticorpos, do hospedeiro. Entretanto, como um todo, este gene pode ser representado como uma sucessão de domínios variáveis flanqueados por regiões conservadas (Lerat e Capy 1999).

Segundo Alberola et al. (1997), seqüências com a terceira ORF teriam a habilidade de expandir-se mais rapidamente entre populações e entre espécies do que os retrotransposons sem esta ORF.

A superfamília *gypsy-like* constitui um grupo heterogêneo amplamente distribuído entre os eucariotos (Marín e Lhoréns 2000). Miller et al. (1999) identificaram e relacionaram filogeneticamente, seqüências de

retrotransposons da superfamília “gypsy-like” nos mais diferentes grupos de organismos, incluindo várias classes de vertebrados. Os autores não encontraram seqüências relacionadas apenas em aves e mamíferos. Os resultados obtidos evidenciaram uma relação entre seqüências homólogas à *gypsy* de vertebrados e fungos, ao invés de vertebrados e insetos (invertebrados), como seria o esperado. Eventos de transmissão horizontal ancestral entre filós ou a existência de diferentes linhagens destas seqüências podem ser possíveis explicações para estes achados.

Gypsy foi identificado e seqüenciado a partir do genoma de *D. subobscura* (Alberola e De Frutos 1996). A análise destas seqüências evidenciou que estes elementos podem ser transcricionalmente ativos e móveis. No entanto, enquanto a estrutura geral é mantida, quando comparamos estes elementos com aqueles de *D. melanogaster*, a ORF 3 codifica uma proteína *env* não funcional, sendo, portanto, estas seqüências classificadas como formas degeneradas de retrovírus (Alberola et al. 1997).

Mizrokhi e Mazo (1991) seqüenciaram e analisaram *gypsy* de *D. virilis*. Suas análises evidenciaram um alto grau de homologia com o elemento canônico de *D. melanogaster*, indicando condições para uma possível atividade desta seqüência. Confirmando esta conclusão, os autores demonstraram que *gypsy* é transcrito nesta espécie. Este elemento é também distribuído em espécies relacionadas à *D. virilis*, embora demonstre uma menor homologia nestes genomas. Diferentes linhagens de *D. virilis* também foram investigadas e um número de cópias aproximadamente constante entre as amostras foi obtido.

Quando a sonda de *gypsy* de *D. melanogaster* é utilizada para investigar a presença destas seqüências em genomas de outras espécies, novamente a ampla distribuição deste elemento é evidenciada (Stacey et al. 1986; Alberola et al. 1997; Loreto et al. 1998a). No entanto, o sinal da hibridização é bastante variável, indicando considerável divergência das espécies analisadas. Quando diferentes linhagens são investigadas com o propósito de comparar os padrões obtidos e inferir estabilidade ou movimentação destas seqüências, Alberola e De Frutos (1996) e Mizrokhi e Mazo (1991) identificaram, em linhagens de *D. subobscura* e *D. virilis* respectivamente, um padrão de fragmentos diferencial, o que parece indicar a atividades destes elementos. Para diferentes linhagens de *D. melanogaster*, Bayev et al. (1984) e Pfeifer e Bender (1988) demonstraram um bom grau de conservação no que se refere ao mapa de restrição de *gypsy*.

Uma fração substancial de elementos *gypsy* de *Drosophila melanogaster* exibe um polimorfismo de sítios de restrição. Foi caracterizada uma família ativa, possuidora de um sítio para a enzima *HindIII* na região central do elemento, além de sítios de reconhecimento para a enzima *XbaI* nas suas LTRs. As evidências que suportam estes achados vem da análise de elementos ativos de várias linhagens mutáveis, quando comparados com elementos não envolvidos com eventos de transposição. Tais achados sugerem que estes sítios de restrição sejam diagnósticos e associados com determinantes funcionais compartilhados por uma subfamília de elementos *gypsy* ativos (Bayev et al. 1984; Lyubomirskaya et al. 1993; Chalvet et al. 1998). Quando o padrão esperado para esta subfamília de *gypsy* ativo é investigado entre machos e fêmeas, bem como para regiões eucromáticas e

heterocromáticas, observou-se que os elementos ativos parecem se distribuir preferencialmente no cromossomo Y de linhagens permissivas, com relação ao gene *flamenco* (Chalvet et al. 1998).

No caso de *gypsy*, parecem coexistir no mesmo genoma retrovírus ativos (codificando uma *env* funcional), retrovírus degenerados, mas também retrotransposons ativos, assim como formas deste último também degeneradas (Alberola et al. 1997).

Gypsy apresenta características de retrovírus endógenos que, em determinadas linhagens permissivas, forma partículas virais (Kim et al. 1994). Sua transposição e/ou propriedades infectivas parecem ser controladas pelo gene do hospedeiro chamado de *flamenco* (*flam*). O locus *flamenco* encontra-se no cromossomo X e a transposição só ocorre na linhagem germinal da prole de fêmeas homocigotas para o alelo permissivo. A regulação de *gypsy* por *flamenco* é tecido-específica, ocorrendo nos ovários (Prud'homme et al. 1995). Pélisson et al. (1997) discutem a relação entre este vírus endógeno e o genoma de *Drosophila*, bem como a evolução para a aquisição do gene *flamenco* com a habilidade de controlar a multiplicação deste elemento. A evolução deste sistema seria o resultado de uma longa coexistência entre parasita e hospedeiro. Estes autores sugerem que *flamenco* é um gene de *Drosophila* que inicialmente encontrava-se em um estado permissivo, antes da invasão de *gypsy* e que foi subsequentemente selecionado para adquirir a função de prevenir a atividade de *gypsy*. Experimentos de Prud'homme et al. (1995) demonstraram que, embora a instabilidade do genoma por transposição de *gypsy* seja controlada principalmente pelo produto do gene *flamenco*, esta atividade também sofre

influências de outros fatores, como idade materna e temperatura, além do número de cópias ativas do elemento.

A proteína codificada pelo gene *flamenco* parece reprimir não só a transposição de *gypsy*, mas também sua expressão. Esta proteína parece agir principalmente no processamento correto do transcrito *env*, embora os demais genes de *gypsy* também tenham sua expressão diminuída em linhagens permissivas com relação a *flamenco* (Pélisson et al. 1994; Bucheton 1995; Alberola et al. 1997).

A existência de linhagens estáveis permissivas de *D. melanogaster* na natureza evidencia o estado defectivo de cópias de *gypsy*, incluindo aquelas presentes em regiões eucromáticas destas linhagens. Vários alelos para *flamenco* já foram identificados e este polimorfismo é visto tanto para a forma restritiva como para a permissiva. Os alelos diferem quantitativamente em sua habilidade de reprimir a expressão de *gypsy* (Pélisson et al. 1997). Este polimorfismo é considerado pelos autores como o resultado de sucessivas ondas de variação em *gypsy* (o modelo é copiado de outro vírus de RNA de insetos, por exemplo, sigma) resultantes da alta probabilidade de erro em seu processo replicativo, característico de qualquer retroelemento/retrovírus. Assim, o modelo postula que quando um alelo restritivo eventualmente surge, pode não haver tempo para sua fixação na população antes que *gypsy* gere uma forma resistente. Este processo levaria a um aumento de variabilidade tanto para *gypsy*, quanto para *flamenco*.

Para a transposição de *gypsy* ocorrer é necessário um estado permissivo materno em células somáticas do ovário. Estes requisitos são indiferentes para as células germinativas. Desta forma, a inserção de novas cópias de *gypsy*,

nos cromossomos da progênie, depende de material genético do elemento vindo das células somáticas ao redor do óvulo (células foliculares do ovário), sendo observado um acúmulo de RNAs de proteína *env* e de partículas virais de *gypsy* em regiões perto da membrana apical, próxima ao ovócito, um processo típico realizado por vírus (Pélisson et al. 1994; Bucheton, 1995; Lécher et al. 1997). Surpreendentemente, a transferência de elementos *gypsy* para o embrião parece não depender da proteína *env* pois linhagens com *gypsy* mutante para este gene infectam a linhagem germinal normalmente. Esta proteína provavelmente tem como papel central a transmissão horizontal de *gypsy* para novos hospedeiros (Chalvet et al. 1999). Os trabalhos realizados até o momento não conseguiram determinar como *gypsy* entra nos ovócitos pois partículas extracelulares não foram ainda identificadas.

Reversão de mutações por inserção de retroelementos são conseguidas por integração de outros elementos ou excisão, através de recombinação entre LTRs. No entanto, eventos de excisão precisos são raros para ambos os casos. Este fenômeno foi estudado por Kuzin et al. (1994) para *gypsy* e foi observado, surpreendentemente, que este mecanismo ocorria em frequências muito superiores às esperadas e caracterizavam-se pela sua precisão. Os autores postulam que um mecanismo diferente e ainda desconhecido possa estar envolvido neste processo.

A superfamília *gypsy-like* apresenta LTRs peculiares. Enquanto todos os retrovírus e demais retrotransposons com LTRs possuem estas seqüências terminais conservadamente flanqueadas por TG...CA, os elementos "gypsy-like" não possuem estas seqüências em suas LTRs. Além disso, seqüências terminais de *gypsy* de diferentes espécies de *Drosophila* divergem nesta

região, o que implica em uma menor rigorosidade de suas integrações (Dej et al. 1998).

A inserção de *gypsy* parece não ser aleatória, já tendo sido registrados vários casos de "hotspot" para este elemento, principalmente, nos locos *ovo* e *cut* (Pélisson et al. 1994; Song et al. 1994; Prud'homme et al. 1997).

O trabalho de Dej et al. (1998) aponta para uma especificidade insercional de *gypsy*. Neste trabalho, os autores analisaram várias inserções de *gypsy* no locus *ovo* ligado ao X, um conhecido "hot spot" para *gypsy*, e verificaram que parece existir uma seqüência consenso preferencial de reconhecimento pelo elemento, rica em pirimidinas(Y) e purinas(R), formando, desta forma, YRYRYR. Este padrão alternado é típico de seqüências de DNA que assumem a forma Z. Talvez a maquinaria de integração de *gypsy* tenha a capacidade de reconhecer esta estrutura (Song et al. 1994).

Gypsy é estável na maioria das linhagens de *Drosophila melanogaster* investigadas mas se transpõe em altas freqüências em linhagens hipermutáveis como MS e MG (Kim et al. 1990; Lyubomirskaya et al. 1993; Ilyin et al. 1991; Bucheton 1995; Pélisson et al. 1997), em que o número de cópias é aumentado.

Experimentos envolvendo a contaminação de linhagens comprovadamente desprovidas de *gypsy* com este elemento ativo foram executadas através de várias abordagens. A utilização de métodos de microinjeção de células providas de *gypsy* em embriões ou o cultivo de larvas em meio de cultura contendo homogeneizado de indivíduos possuidores de *gypsy* ativos, demonstraram a capacidade deste elemento de mover-se de

uma célula para outra, infectando linhagens livres de *gypsy* (Kim et al. 1994; Song et al. 1994).

A ampla distribuição de *gypsy* entre espécies do gênero *Drosophila* (Stacey et al. 1986; Loreto et al. 1998a; Biémont e Cizeron 1999) é atribuída, por muitos autores, à existência de elementos *gypsy* ancestrais nestes insetos. No entanto, se levarmos em consideração as características infectivas de *gypsy* de *D. melanogaster*, eventos infectivos podem ter ocorrido e propiciado esta grande expansão (Alberola e De Frutos 1996; Vásquez-Manrique et al. 2000). Assim, fenômenos de transmissão horizontal podem explicar o cenário de distribuição deste elemento.

Em várias espécies do subgrupo *melanogaster* de *Drosophila*, seqüências homólogas a *gypsy* parecem localizar-se, principalmente, em regiões centroméricas e pericentroméricas, com um pequeno número de cópias nos braços cromossômicos (em torno de 5) em *D. melanogaster* (Pélisson et al. 1997; Junakovic et al. 1998; Biémont e Cizeron 1999; Vieira et al. 1999). Bayev et al. (1984) estimaram o número total de cópias de *gypsy* para *D. melanogaster*, incluindo aquelas sub-replicadas nos cromossomos politênicos, e chegaram a um número próximo de 20 cópias. Lambertson et al. (1989) observaram uma grande heterogeneidade de cópias dentro do mesmo genoma. Neste trabalho, os autores examinaram uma única linhagem mutante de *D. melanogaster* e evidenciaram três tipos de seqüências: 1) *gypsy* completo, com os sítios de restrição diagnósticos; 2) cópias levemente divergentes e 3) cópias altamente divergentes. Nestes dois últimos grupos, foi evidenciado um considerável polimorfismo de restrição. Os autores

argumentam que a maioria das cópias não se enquadram na primeira categoria e representam, provavelmente, elementos defectivos.

A primeira ampla investigação sobre a distribuição de *gypsy* de *D. melanogaster* em outras espécies do gênero *Drosophila* foi realizada por Stacey et al. (1986), através do método de *Southern blot*. Estes autores observaram a ampla distribuição com descontinuidades ocasionais e salientaram que o grau de similaridade medido pelo sinal de hibridização não se correlacionava com o grau de divergência entre sonda e a espécie investigada, no que se refere à similaridade em seqüência e parentesco filogenético. Estes dados foram confirmados diante da seqüência completa de *gypsy* de *D. subobscura* e *D. virilis*. Similaridade entre estas seqüências são discrepantes em relação aos seus tempos de divergência e autores sugerem que eventos de transmissão horizontal podem ter ocorrido na história evolutiva de *gypsy* entre estas espécies (Alberola e De Frutos 1993a).

Pélisson et al. (1997) revisam a distribuição inter e intraespecífica de *gypsy* em várias espécies do gênero *Drosophila* e apontam a ampla distribuição deste elemento. Entretanto, os autores também salientam possíveis eventos de transmissão horizontal, explicando discrepâncias em análises comparativas de elementos *gypsy* de diferentes espécies. Dentro do subgrupo *melanogaster*, observa-se um considerável polimorfismo de tamanho de fragmento de restrição em análises de *Southern blot*. Variação interpopulacional em *Drosophila melanogaster* foi investigada e um considerável compartilhamento de fragmentos de restrição entre linhagens (mais de 50% dos fragmentos) sugere que uma alta proporção de cópias desta família consiste de elementos estáveis desde o ancestral das

populações analisadas. Além disso, a maioria destes elementos parece ser defectiva em função do mapa de restrição alterado. Entretanto, algumas poucas cópias, localizadas nos braços euromáticos, apresentam uma localização polimórfica entre linhagens, o que sugere uma movimentação recente. Alberola e De Frutos (1993b) investigaram *gypsy* em 8 espécies do subgrupo *obscura* por *Southern blot* e detectaram, além da ubiquidade esperada, um alto nível de conservação de seqüências de *gypsy*. Estes resultados apontam para um modelo que assume a presença de *gypsy* antes da divergência deste subgrupo e sua transmissão vertical. Estes mesmos autores realizaram análises similares com diferentes populações naturais de *D. subobscura* e encontraram padrões de hibridização variáveis, indicando a possível presença de elementos ativos. Em concordância, os autores também detectaram seqüências completas. Além disso, a estrutura molecular interna destes elementos parece estar conservada, embora em algumas linhagens tenham sido detectados padrões de fragmentos explicáveis por polimorfismo de sítios de restrição. Para esta espécie, os autores detectaram um número de cópias que varia de 4 a 7 cópias euromáticas entre linhagens e um alto polimorfismo de localização das marcações, o que parece sugerir extensa transposição de *gypsy*.

Vásquez-Manrique et al. (2000) realizaram um estudo filogenético baseado na análise de uma região do gene *env* de *gypsy* proveniente de dez espécies do grupo *obscura* do subgênero *Sophophora* de *Drosophila*. As análises mostram que elementos *gypsy* destas espécies formam um grupo monofilético e bem divergente do elemento *gypsy* de *D. melanogaster*. As relações obtidas para o grupo em questão suportam a hipótese de

transmissão vertical. No entanto, seqüências de *D. hydei* e *D. virilis*, incluídas na análise, agrupam-se com espécies do subgrupo *affinis*, que pertencem a outro subgênero de *Drosophila*, o que dá suporte para a existência de eventos de transmissão horizontal quando a análise torna-se mais ampla do ponto de vista taxonômico. Ainda neste estudo é verificada a coexistência de cópias bastante divergentes no genoma de *D. hydei*. Outro dado interessante de suas análises é que a porção 5' do gene do envelope mantém características funcionais nas diferentes espécies investigadas.

Terzian et al. (2000) analisaram filogeneticamente uma porção do gene da integrase de *gypsy* de espécies do subgrupo *melanogaster* e detectaram a formação de duas linhagens de *gypsy*, além de um complexo padrão de transferência vertical com múltiplos eventos de transmissão horizontal. No caso de *gypsy*, este fenômeno poderia ocorrer sem um vetor, como um parasita ou vírus, já que este retrotransposon pode codificar uma partícula infecciosa.

Song et al. (1994) mencionam que inserções de *gypsy* em alta freqüência parecem correlacionar-se com fenômenos de instabilidade genética similares àqueles provenientes de disgenesia do híbrido. Em linhagens recém infectadas com *gypsy* observou-se o aparecimento de mutantes visíveis por várias gerações, além de altas freqüências de rearranjos cromossômicos, inserções somáticas e esterilidade.

Gypsy tem o potencial de mudar a expressão de genes próximos ao seu local de inserção. Esta atividade tem sido extensamente estudada e os achados provenientes destes estudos levaram a fazer deste elemento um modelo para a atividade de "insulator" em *Drosophila*. Trabalhos demonstraram que a inserção de *gypsy* entre promotores e "enhancers"

previne a ativação gênica. Uma região de 340 pb do *gypsy* foi caracterizada como a responsável por este fenômeno (Cai e Levine 1997). Reversões de mutações causadas por *gypsy* ocorrem quando esta região é deletada ou é alvo de inserções de outros elementos (Modolell et al. 1983; Freund e Meselson 1984; Pfeifer e Bender 1988; Flavell et al. 1990; Smith e Corces 1992). Os "insulators" são responsáveis por construir barreiras genômicas, limitando e protegendo regiões de efeitos de posição promovidos por ativadores ou silenciadores, inclusive heterocromatinização. Estudos posteriores mostraram que a região de *gypsy* responsável por este efeito tinha proteínas específicas ligadas a ela e esta ligação era capaz de modular sua atividade. Foi descoberta a proteína supressora de cerdas da asa – [*Su(Hw)*] que se liga a esta região, além de uma segunda proteína produzida pelo gene *mod(mdg4)* que pode interagir com a primeira, regulando seu efeito. O mecanismo proposto de ação desta região ligada à [*Su(Hw)*] propõe que o complexo protéico *Su/mod* liga-se tanto ao DNA de *gypsy*, como à lâmina nuclear ou outros componentes da arquitetura do núcleo, criando uma série de domínios em "loops" e compartimentalizando a região genômica associada a ela. Várias evidências têm apontado para a localização destas proteínas envolvidas na periferia nuclear. Regiões com inserções de *gypsy* assumem uma conformação de vários "loops" e estes são dispostos na periferia. Quando cópias deste elemento são inseridas próximas a duas seqüências do DNA hospedeiro, originalmente distantes, elas passam a localizar-se juntas. O modelo pressupõe que, para um "enhancer" atuar sobre uma região promotora, ambas seqüências devem permanecer no mesmo "loop". Assim, se *gypsy* é inserido entre estas regiões, promovendo um centro de ancoragem

em seu "insulator", ativação do gene seria interrompida pois as regiões regulatórias ficariam em "loops" diferentes. O mesmo princípio pode ser pensado para se deter a heterocromatinização, pois o "loop" interferiria com a extensão do processo de condensação da cromatina (revisão em Bell et al. 2001). Talvez seqüências de *gypsy* possam ser importantes agentes estruturais na compartimentalização nuclear dos seus genomas hospedeiros.

O GÊNERO *Drosophila* E *Zaprionus indianus*

Segundo a revisão de Wheeler (1981), o gênero *Drosophila* pertence à família *Drosophilidae*, subfamília *Drosophilinae*. Dentro desta subfamília, existem 40 gêneros descritos. O maior deles é o gênero *Drosophila*, composto de 15 subgêneros, onde se distribuem mais de 1400 espécies já descritas. Outro gênero abordado neste trabalho é o gênero *Zaprionus*, ao qual *Zaprionus indianus* pertence.

Dos quinze subgêneros existentes no gênero *Drosophila*, merecem destaque, pelo número de espécies, o subgênero *Drosophila*, com 759 espécies seguido pelo subgênero *Sophophora* com 233 espécies (dados de Wheeler 1981). O presente trabalho investiga espécies que seguem, ainda, os subgêneros *Dorsilopha* e *Scaptodrosophila*. Throckmorton (1975) aponta várias evidências que sugerem uma origem tropical para a família *Drosophilidae*. O que vimos hoje, é que a grande parte destas espécies são encontradas em regiões neotropicais (Wheeler 1981).

Registros fósseis datam a origem da família para, pelo menos, 50 milhões de anos atrás (Throckmorton 1975). A primeira radiação de espécies pertencentes ao gênero *Drosophila* é representada pelo subgênero *Scaptodrosophila* e parece ter ocorrido nos trópicos do Velho Mundo. Outra importante radiação é a *Sophophora* que compreende, principalmente, a expansão das linhagens *saltans-willistoni*, *melanogaster* e *obscura*. A primeira é postulada originar-se nos trópicos do Novo Mundo, sendo as demais no Velho Mundo. Acredita-se que as primeiras são derivadas (*willistoni* o último grupo a diversificar-se) em relação às demais, sendo *obscura* e *melanogaster* grupos

irmãos. Assim, sugere-se que a radiação *Sophophora* surgiu também em regiões tropicais do Velho Mundo (Throckmorton 1975).

A radiação *Drosophila* é compreendida de várias subradiações que ocorreram posteriormente àquelas já mencionadas acima. Dentro deste cenário, merecem destaque as radiações *virilis-repleta*, pioneira neste subgênero, além das *immigrans-tripunctata* e *Hirtodrosophila*. Da radiação *immigrans-tripunctata*, temos o grupo *immigrans*, originário do Velho Mundo, e o grupo *tripunctata*, do Novo Mundo. O gênero *Zaprionus* relaciona-se à radiação *immigrans*. Por último, temos a radiação *Hirtodrosophila*, da qual o subgênero *Dorsilopha* faz parte. Muitos grupos de espécies ainda são associados com estas grandes e complexas radiações (Throckmorton 1975). A Figura 3 mostra um esquema destas radiações, salientando a posição das espécies analisadas neste trabalho. As relações apresentadas nesta Figura baseiam-se nos dados apresentados por Throckmorton (1975), Wheeler (1981), e filogenias obtidas a partir de dados moleculares (Pelandakis e Solignac 1993; Kwiatowski et al. 1994; Russo et al. 1995; Kwiatowski e Ayala 1999), que confirmam as relações estabelecidas pelos primeiros trabalhos.

Os dados de Throckmorton (1975) tendem a mostrar subgêneros e gêneros surgindo a partir do subgênero *Drosophila* (como pode ser observado na Figura 3), o que torna a classificação usual parafilética. Grimaldi (1990) amplia o número de caracteres morfológicos, reavalia os dados de Throckmorton (1975) e apresenta, em suas análises, o gênero *Drosophila* como derivado dentro de sua família e um grupo irmão de vários outros grupos incluindo *Zaprionus*, *Hirtodrosophila* e *Scaptodrosophila*. Baseado nestes resultados, ele propõe uma nova classificação, onde o *status* destes dois

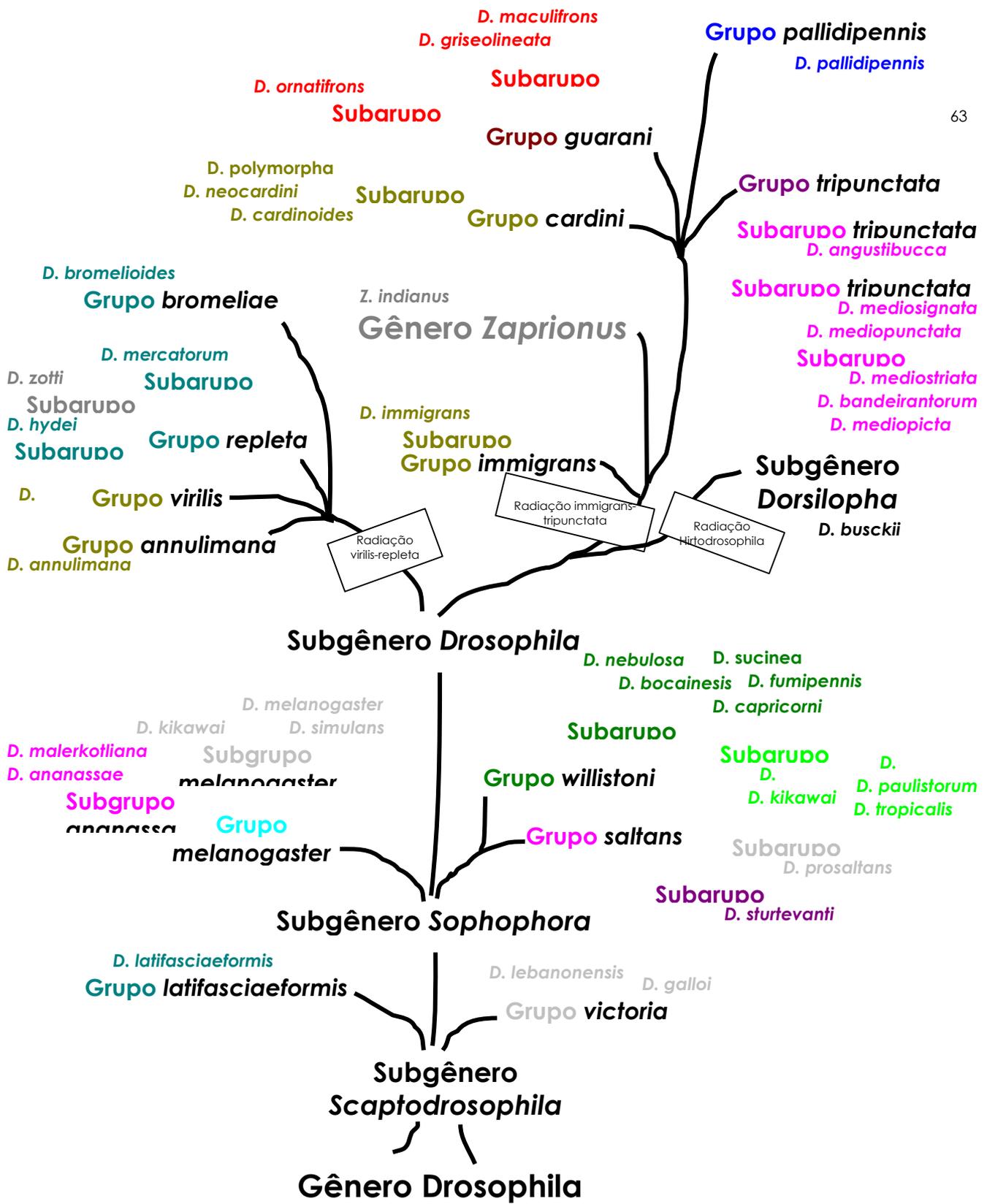


Figura 3. Esquema das relações de parentesco entre as espécies investigadas no presente trabalho (Baseada na filogenia de Throckmorton 1975 e na filogenia proposta de Kuimovskii e Avdey 1999).

Últimos grupos é revisado e classificado como gênero, ao invés de subgênero como é colocado na classificação tradicional.

Powell e DeSalle (1995) revisam as filogenias construídas a partir de vários tipos de dados moleculares e apontam suas congruências e discrepâncias. A posição mais confusa é a do gênero *Zaprionus*. Dados morfológicos analisados sob uma perspectiva fenética por Throckmorton (1975) e corroborados por vários trabalhos com filogenias construídas a partir de dados moleculares, colocam este grupo dentro do subgênero *Drosophila*. Contrariando estes resultados, outros dados obtidos de análises de caracteres de comportamento, mtDNA e das conclusões de Grimaldi, posicionam *Zaprionus* fora do gênero *Drosophila*.

Em uma recente revisão sobre a classificação de *Drosophila* e gêneros relacionados, Kwiatowski e Ayala (1999) compilam vários dados moleculares e chegam a uma filogenia consenso, em que, novamente, o gênero *Zaprionus* encontra-se relacionado ao subgênero *Drosophila*, derivado em relação ao subgênero *Sophophora*. Apesar disso, *Zaprionus* é considerado gênero por todas classificações propostas até então.

Zaprionus indianus é a mosca mais comum do seu gênero no continente africano, local de provável origem do grupo. As espécies de *Zaprionus* ocorrem principalmente nas regiões Afrotropical, Oriental e Australiana. A espécie utilizada neste estudo está expandindo-se através dos trópicos, provavelmente devido à intensificação do comércio de frutas. O primeiro registro desta mosca no Brasil foi em 1999 (Vilela 1999). Hoje ela é conhecida no Brasil como a mosca-africana-do-figo e parece ter invadido todo o

continente americano, tornando-se uma considerável praga desta cultura. Sua capacidade de desenvolver-se em uma grande gama de substratos, de diferentes origens geográficas tem, provavelmente, facilitado sua rápida expansão neste continente.

Dentre as espécies investigadas neste trabalho, as seguintes são consideradas cosmopolitas: *D. melanogaster*, *D. simulans*, *D. ananassae*, *D. hydei*, *D. immigrans* e *D. busckii*. Espécies de ampla distribuição, consideradas como semi-cosmopolitas, são citadas a seguir: *D. kikkawai*, *D. mercatorum*, *D. virilis*, *D. latifasciaeformis* e *D. malerkotliana* (Wheeler 1981).

Com exceção de *D. lebanonensis*, cuja distribuição é restrita às regiões neárticas e paleárticas, as demais espécies presentes neste estudo são consideradas espécies neotropicais (Wheeler 1981).

OBJETIVOS

Este trabalho tem o objetivo geral de contribuir para o entendimento da evolução do retroelemento transponível *gypsy* em espécies do gênero *Drosophila* e na espécie invasora *Zaprionus indianus*.

Tendo em mente a meta acima, delineamos diferentes abordagens que foram desenvolvidas com os seguintes objetivos específicos:

Capítulo 2

1) Analisar por *Southern blot* a distribuição e a conservação de sítios de restrição de seqüências homólogas a *gypsy* em várias espécies do gênero *Drosophila*, especialmente aquelas neotropicais, e *Z. indianus*.

2) Estimar o número de cópias, através da técnica de *Southern blot*, de *gypsy* nas espécies investigadas.

3) Avaliar a conservação de regiões do retroelemento *gypsy* nas diferentes espécies pelo anelamento de oligonucleotídeos degenerados e posterior amplificação por PCR.

4) Construir uma filogenia, a partir de seqüências amplificadas e clonadas de *gypsy* obtidas do genoma das espécies, a fim de estabelecer as relações de parentesco existentes.

5) Comparar a filogenia obtida das seqüências homólogas a *gypsy* com as relações filogenéticas entre as espécies hospedeiras, a fim de detectar coincidências e incongruências.

Capítulo 3

6) Investigar, por hibridização "in situ" sobre cromossomos politênicos e por *Southern blot*, a presença do retroelemento *gypsy* em *Drosophila maculifrons* e *Drosophila griseolineata*, duas espécies onde este elemento parece ter evoluído verticalmente.

CAPÍTULO 2

EVOLUÇÃO COMPLEXA DO RETROELEMENTO

TRANSPONÍVEL *gypsy* EM *Drosophila*

E *Zaprionus indianus*

HERÉDIA, Fabiana. (Dep. de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil)

LORETO, Elgion L. da S. (Dep. de Biologia, Universidade Federal de Santa Maria, Santa Maria, RS, Brasil)

VALENTE, Vera L. da S. (Dep. de Genética, Caixa Postal 15053; fax (051) 3316-7311; Universidade Federal do Rio Grande do Sul, CEP 91501-970; Porto Alegre, RS, Brasil)

*e-mail para correspondência: fheredia@terra.com.br

**Trabalho em preparação a ser enviado para a revista Molecular Biology e Evolution

INTRODUÇÃO

Os elementos retrotransponíveis são encontrados em todos os eucariotos, sendo vislumbradas várias e heterogêneas famílias dentro desta classe. A superfamília de retroelementos *Ty3/gypsy* é amplamente distribuída nos seres vivos (Miller et al. 1999; Marin e Lloréns 2000) e sua relação com os retrovírus tem sido inferida em vários trabalhos (Kim et al. 1994; Péliçon et al. 1997; Lerat e Capy 1999). Análises filogenéticas de domínios conservados, produzidas por diferentes regiões destes elementos, confirmam sua íntima relação com os retrovírus (Xiong e Eickbush 1990; Capy et al. 1998).

O retroelemento *gypsy*, também conhecido como *mdg4*, foi o primeiro identificado em *Drosophila melanogaster*, possui 7,5 kb de comprimento, tendo LTRs (longas repetições terminais) com 482 pb (Georgiev et al. 1981; Bayev et al. 1984). Kim et al. (1994) reportaram as primeiras evidências que culminaram na caracterização deste elemento móvel como o primeiro retrovírus em invertebrados. Como outros retrovírus, *gypsy* possui um genoma composto de três ORFs (módulos abertos de leitura) chamadas de *gag*, *pol* e *env*, que codificam proteínas responsáveis pela sua replicação e infectividade. O terceiro gene, localizado na porção 3' do elemento, apresenta uma sucessão de regiões variáveis, flanqueadas por seqüências conservadas (Lerat e Capy 1999). A proteína *env* é formada de um transcrito processado a partir do RNA completo de 7 kb, formado da transcrição do elemento inteiro. *Env* é glicosilada e processada como em retrovírus e é encontrada associada a partículas virais de linhagens hipermutáveis de *D.*

melanogaster (Avedisov e Ilyin 1994; Pélisson et al. 1994; Song et al. 1994; Prud'homme et al. 1995).

Gypsy foi posteriormente isolado e seqüenciado a partir do genoma de *D. virilis* e *D. subobscura* (Mizrokhi e Mazo 1991; Alberola e De Frutos 1996). Análises comparativas destes elementos, em relação a *gypsy* de *D. melanogaster*, demonstram que embora sua terceira ORF encontre-se interrompida, o restante de suas seqüências possibilitam sua atividade e mobilidade. Além disso, as relações filogenéticas inferidas das seqüências destes três elementos móveis, sugerem inconsistências quando comparadas às relações de parentesco de suas espécies hospedeiras (Alberola e De Frutos 1996).

Amplios levantamentos por seqüências homólogas à *gypsy* de *D. melanogaster* têm sido feitos, através das técnicas de *Southern blot* e *Dot blot* (Stacey et al. 1986; Alberola et al. 1997; Loreto et al. 1998a) e evidências da ampla distribuição deste TE, no gênero *Drosophila*, têm sido acumuladas. No entanto, inconsistências entre sinais de hibridização e a filogenia das espécies apontam para as mesmas perspectivas vislumbradas pela análise comparativa, feita por Alberola e De Frutos (1996), acima referida.

Gypsy é estável na maioria das linhagens de *Drosophila melanogaster* investigadas, localizando-se principalmente em regiões centroméricas e pericentroméricas. Um pequeno número de cópias é reportado nos braços eucromáticos desta espécie (em torno de 5) (Pélisson et al. 1997), enquanto o número total de cópias tem sido estimado por Bayev et al. (1984) em 20, incluindo aquelas que se encontram em regiões sub-replicadas nos

cromossomos politênicos e, por isso, não são detectadas em ensaios de hibridização “*in situ*”.

Em linhagens hipermutáveis, como MS e MG, o retrotransposon *gypsy* apresenta alta atividade e um número de cópias elevado. O *status* especial destas linhagens, em relação a este elemento, é caracterizado pela existência de alelos permissivos (presentes em homozigose) para o gene do hospedeiro *flamenco*, cujo produto é conhecido por reprimir a atividade de *gypsy* (Prud'homme et al. 1995; Pélisson et al. 1997).

Tem sido demonstrada, por análises de polimorfismo de restrição, a coexistência, em um mesmo genoma, de diferentes subfamílias de *gypsy* (Lambertson 1989). Através desta abordagem, uma subfamília ativa foi detectada e diagnosticada como possuidora de sítios de restrição peculiares para as enzimas *HindIII* e *XbaI* (Chalvet et al. 1998). Acredita-se que a maioria das cópias de *gypsy* existentes no mesmo genoma é defectiva e bastante divergente entre si, o que contribui para a heterogeneidade observada (Lambertson 1989).

A ubiquidade de seqüências homólogas à *gypsy* no Gênero *Drosophila* tem sido relacionada à sua presença antes da diversificação dos principais ramos que compõem este grupo de insetos, com posterior expansão por transmissão vertical (Alberola e De Frutos 1996). Por outro lado, algumas evidências, obtidas por *Southern blot* e análises filogenéticas de seqüências de *gypsy*, em determinados grupos de espécies, apontam para um quadro evolutivo um pouco mais complexo, em que possíveis eventos de transmissão horizontal na história destes grupos específicos são sugeridos (Alberola e De Frutos 1996; Terzian et al. 2000; Vázquez-Manrique et al. 2000).

Com o objetivo de vislumbrar um cenário mais abrangente sobre mecanismos subjacentes ao processo evolutivo que regem a história do retroelemento *gypsy* no Gênero *Drosophila*, nós implementamos uma análise de amplo espectro investigando 40 espécies, pertencentes a diferentes grupos e subgêneros deste Gênero, além de uma espécie do Gênero *Zaprionus*. Neste trabalho, seqüências homólogas à *gypsy* foram investigadas por *Southern blot* e PCR. Além disso, um amplo estudo filogenético, baseado na análise de uma região de aproximadamente 500 pb da terceira ORF foi realizado. Nossos resultados, tomados em conjunto, sugerem um padrão de evolução bem mais complexo do que os trabalhos, até hoje publicados para *gypsy*, têm proposto.

MATERIAL E MÉTODOS

Manutenção dos estoques de *Drosophila*

Isolinhagens de todas as espécies trabalhadas foram estabelecidas e mantidas em meio de cultura de farinha (Marques et al. 1966) dentro de câmaras de temperatura e umidade constante (a $17^{\circ}\text{C} \pm 1$ grau e 60% de umidade relativa). A Tabela 1 apresenta as espécies, bem como sua precedência e data de coleta.

Algumas espécies foram analisadas por *Southern blot*, intra-especificamente. As diferentes populações utilizadas neste trabalho são ressaltadas na Tabela 1.

Tabela 1 no anexo.

Extração de DNA

Aproximadamente 100 moscas foram maceradas em nitrogênio líquido dentro de um tubo de microcentrífuga de 1,5ml. Acrescentou-se à massa homogeneizada 750µl de tampão de lise com a seguinte composição: Tris-HCl 0,1M, EDTA 0,1M, SDS 1% e NaCl 0,06M. A amostra foi incubada por 30 min a 65°C. A extração de proteínas e lipídios foi realizada com solventes orgânicos da seguinte forma: acrescentou-se um volume de fenol equilibrado pH 8,0 e agitou-se por 10 min, lentamente. Centrifugou-se por 10 min a 8000 g. Foi passado o sobrenadante do tubo original para um novo e repetiu-se o procedimento com um volume de fenol-clorofórmio (24 clorofórmio:1 álcool isoamílico). O sobrenadante também foi retirado após a centrifugação e submetido a uma nova etapa nas mesmas condições que a primeira, agora com clorofórmio (24 clorofórmio:1 álcool isoamílico) apenas. A precipitação do DNA foi realizada retirando-se, novamente, o sobrenadante e acrescentando-se 20% do volume inicial com NaCl 1M e 2 volumes de álcool etílico absoluto. Após a precipitação, centrifugou-se por 5 min para precipitar o DNA. O líquido sobrenadante foi descartado e o precipitado foi lavado por 3 vezes com álcool etílico 70%. O DNA foi deixado secar e acrescentou-se 30µl de TE para a ressuspensão. O DNA foi, então, submetido a um tratamento com RNase e testado com enzimas de restrição para avaliação de sua qualidade antes de ser utilizado no trabalho.

Sonda

A sonda utilizada foi o plasmídeo pGGHS, contendo o retroelemento *gypsy* completo, isolado do genoma de *D. melanogaster*, gentilmente cedida pelo Dr. Dale Dorsett (Dorsett et al. 1989). A Figura 1 apresenta um esquema da estrutura molecular do retroelemento *gypsy* de *D. melanogaster*, mostrando suas ORFs, LTRS e os sítios de restrição para as enzimas utilizadas neste trabalho. A Figura ainda ressalta o polimorfismo de restrição entre os elementos completos conhecidos na literatura (provenientes de *D. virilis* e *D. subobscura*) para estas enzimas. Para a construção do mapa de restrição, foi utilizado o programa pDRAW 32 1.0 revision 1.1. 39 desenvolvido por Kjeld Olesen.

As enzimas utilizadas neste trabalho foram: *Bgl*III, *Hind*III e *Bam*HI (GIBCO/BRL). Esta última, por não apresentar sítios de restrição internos ao elemento *gypsy* de *D. melanogaster*, produz um número de fragmentos que estima o número de cópias do elemento no genoma desta espécie. *Bgl*III, por sua vez, reconhece seqüências nas LTRs deste elemento. Assim, fragmentos de tamanho esperado (7 kb) indicam a presença de possíveis elementos completos.

Southern blot

Dez µg de DNA de cada amostra foram clivados com as enzimas de restrição acima descritas, utilizando-se as condições recomendadas pelo fabricante. Os fragmentos produzidos foram fracionados em gel de agarose a 0.8%, transferidos para uma membrana Hybond N+ (Amersham Pharmacia Biotech) e hibridizados de acordo com o protocolo do "kit" Gene Images® da Amersham Pharmacia Biotech. As membranas foram hibridizadas com uma

Figura 1 no anexo.

sonda marcada (através do método de "random prime") a 60 °C em uma solução contendo 0.1% SDS, 5% dextran sulfato e líquido bloqueador diluído 20 vezes em 5xSSC. A membrana foi lavada duas vezes a 60° C, primeiramente com 1x SSC e 0.1% SDS e após com 0.5x SSC e 0,1% SDS. Em ambos os casos, agitando-se por 15 minutos. Para a detecção, foi utilizado o método do "kit" CPD-Star (Amersham Pharmacia Biotech). O tamanho molecular (em kb) dos fragmentos detectados foi determinado usando os fragmentos do marcador de peso molecular 1 kb Plus DNA ladder (GIBCO/BRL) como controle.

Amplificação por PCR

Três pares de "primers" degenerados foram desenhados a partir de regiões conservadas, baseadas no alinhamento de seqüências do retroelemento *gypsy* publicadas no Genebank (NCBI) de *D. melanogaster* (acesso M12927), *D. virilis* (acesso M38438) e *D. subobscura* (acesso X72390). Os "primers" obtidos são os seguintes (utilizando como referência a seqüência de *D. melanogaster*): GYP1S (sense) GAGTTTGCAGGTGGARGCRCC e GYP1AS (antisense) GCRAACARGCTTCTCTCWATGCTWGC que correspondem às regiões da ORF1 1313-1333 e 1869-1893, respectivamente, produzindo um fragmento de 580 nucleotídeos; GYP3S1 (sense) YCTMGATTCTTAGGYACWGC e GYP3AS1 (antisense) GTCYTCGTCGAGKCGCARRAT que correspondem às regiões da ORF3 5839-5859 e 6336-6358, respectivamente, produzindo um fragmento de 519 nucleotídeos; GYP3S2 (sense) AAAGGCGAYTTGGTTGACTCC e GYP3AS2 (antisense) CARGTGGCTRGGTTGRGTGTG que correspondem às regiões da ORF3 6026-

6048 e 6491-6511, respectivamente, produzindo um fragmento de 485 nucleotídeos. Para todos os pares de "primers", os componentes da reação foram: 50ng de DNA foram submetidos a uma reação final de 50µl contendo 1U Taq polimerase (GIBCO/BRL), 5µl de tampão 10X fornecido pelo fabricante, 50µM de cada nucleotídeo, 20pmol de cada primer e 1,5mM de MgCl₂. As condições de amplificação para todas as reações foram: 96°C por 2 min seguidos de 35 ciclos de 96°C por 15 seg, 55°C por 30 seg e 72°C por 1,5 min, seguidos por 72°C por 5 min para finalizar a reação.

Clonagem e Sequenciamento

Os produtos de tamanho esperado, referentes ao terceiro par de "primers", foram clonados no plasmídeo pCR®-TOPO® utilizando-se o "kit" TOPO TA Cloning® Kit for Sequencing da Invitrogen. Plasmídeos das colônias obtidas foram extraídos pelo método de "miniprep" por lise alcalina e purificados com GFX™ PCR DNA and Gel Band Purification "Kit" da Amersham Pharmacia Biotech. Para o sequenciamento, dois a três clones, foram selecionados.

O sequenciamento automatizado foi realizado na Universidade da Geórgia (Athens, GA) pelo Molecular Genetics Instrumentation Facility, utilizando os "primers" T7 e M13 que anelam em regiões do plasmídeo que flanqueiam o fragmento clonado.

As seqüências de *D. virilis*, *D. subobscura* e *D. melanogaster*, utilizadas nas análises filogenéticas, foram obtidas a partir das seqüências publicadas no Genebank (NCBI), utilizadas para o desenho dos "primers" deste trabalho.

Para *D. willistoni*, nós estendemos a análise de seqüências homólogas à *gypsy* para duas populações, a fim de confirmar a identidade desta seqüência. Utilizamos as populações de Morro Santana/RS-Brasil e Pará/Brasil por serem as mais distantes geograficamente, entre as disponíveis em nosso laboratório.

Análises Filogenéticas

As seqüências de nucleotídeos e aminoácidos obtidas foram alinhadas pelo programa Clustal V 1.3b (Higgins et al. 1992). Ajustes manuais do alinhamento, principalmente aquele que gerou a seqüência de aminoácidos, foram necessários. A obtenção da provável seqüência de aminoácidos, a partir das seqüências de nucleotídeos obtidas, foi realizada pelo programa GeneDoc Version 2.6.001 (Nicholas e Nicholas 1997). Todas as análises filogenéticas foram conduzidas, usando-se o programa MEGA versão 2.1 (Kumar et al. 2001). Os métodos utilizados para a inferência das árvores foram Máxima Parcimônia e Neighbor Joining. Este último foi escolhido por não pressupor taxas evolutivas constantes. Teste de bootstrap com 100 amostragens foi realizado para todas as árvores obtidas. Para o método de Máxima Parsimônia, utilizamos o método de pesquisa Close-Neighbor-Interchange (CNI) com nível 3 e opção "random addition" (10 amostragens). A taxa de substituição nucleotídica foi calculada pelo método de Kimura (1980), modelo de dois parâmetros. A distância p foi utilizada para a análise de seqüência de aminoácidos. Foi calculado dN e dS das seqüências traduzíveis, usando o método de Nei e Gojobori (1986). Foi realizado o teste Z de Seleção que compara a abundância relativa de substituições sinônimas e não

sinônimas para a avaliação da existência de seleção na seqüência codificante. Os demais parâmetros de análise foram aqueles estabelecidos como padrão pelo programa utilizado.

RESULTADOS

Southern blot

Com o objetivo de avaliar a distribuição de seqüências homólogas à *gypsy* em diferentes espécies do gênero *Drosophila*, assim como a conservação dos sítios de restrição dessas seqüências em relação à encontrada em *D. melanogaster*, empregamos a técnica de *Southern blot*. As espécies e populações investigadas para as diferentes enzimas de restrição são mostradas na Tabela 2.

Detectamos seqüências homólogas à sonda utilizada, em todas espécies investigadas, indicando que seqüências homólogas ao retrotransposon *gypsy* são übiquas no genoma do gênero *Drosophila*. Entretanto, obtivemos uma grande diversidade de padrões de hibridização entre as diferentes espécies, caracterizando-se, principalmente, por um elevado número de bandas e baixo sinal de hibridização. A maioria das bandas produzidas mostraram-se fracas quando comparadas com o controle (*D. melanogaster*). Cabe salientar, também, que foram necessários longos tempos de exposição para a obtenção de um bom sinal de hibridização nos blots. Além disso, quantidades menores de DNA (metade daquelas utilizadas para as espécies investigadas) foram utilizadas para o controle, a fim de

Tabela 2 . Espécies analisadas por Southern blot.

GÊNERO	SUBGÊNERO	GRUPO	ESPÉCIE	HindIII	BglIII	BamHI		
<i>Drosophila</i>	<i>Drosophila</i>	<i>guarani</i>	<i>D. maculifrons</i>	X	*	□		
			<i>D. griseolineata</i>	□	* □	X		
			<i>D. ornatifrons</i>	X	□	X		
		<i>cardini</i>	<i>D. polymorpha</i>	X	X	X		
			<i>D. cardinoides</i>	X	X	X		
			<i>D. neocardini</i>	X	X	X		
			<i>D. immigrans</i>	X	□	X		
			<i>D. pallidipennis</i>	X	□	□		
		<i>tripunctata</i>	<i>D. bandeirantorum</i>	X	X	X		
			<i>D. angustibucca</i>	X	X	□		
			<i>D. mediopunctata</i>	X	□	X		
			<i>D. mediosignata</i>	na	□	X		
			<i>D. mediopicta</i>	X	na	X		
			<i>D. mediotriata</i>	*	X	X		
			<i>D. virilis</i>	*	*	*		
			<i>virilis</i>	<i>D. hydei</i>	□	□	X	
				<i>D. mercatorum</i>	□	□	X	
			<i>Sophophora</i>	<i>willistoni</i>	<i>D. willistoni</i>	X	X	X
		<i>D. nebulosa</i>			X	□	na	
		<i>D. fumipennis</i>			na	X	na	
		<i>D. paulistorum</i>			X	X	X	
		<i>D. bocainensis</i>			X	X	X	
		<i>D. capricorni</i>			X	na	na	
		<i>D. tropicalis</i>			□	X	X	
		<i>D. sucinea</i>			X	X	na	
		<i>D. insularis</i>			X	X	X	
		<i>saltans</i>			<i>D. prosaltans</i>	na	□	X
					<i>D. melanogaster</i>	*	*	*
		<i>melanogaster</i>			<i>D. kikkawai</i>	X	X	X
					<i>D. ananassae</i>	X	□	X
					<i>D. malerkotliana</i>	X	□	X
					<i>D. simulans</i>	* □	X	X
			<i>D. lebanonensis</i>	X	X	X		
<i>D. galloi</i>	X		X	X				
<i>Scaptodrosophila</i>	<i>D. latifasciaeformis</i>	□	X	X				
	<i>D. busckii</i>	□ ⊕	□	X				
	<i>Zaprionus</i>	<i>Zaprionus</i>	<i>armatus</i>	<i>Z. indianus</i>	na	*	X	

na – Espécies não avaliadas para a respectiva enzima de restrição.

X – Espécies que não produziram fragmentos esperados ou bandas predominantes em sinal.

* – Espécies que produziram fragmentos de tamanho esperado, de acordo com o mapa de restrição de *gypsy* de *D. melanogaster*.

⊕ – Espécies que produziram fragmentos de tamanho esperado, de acordo com o mapa de restrição de *gypsy* de *D. subobscura*.

□ – Espécies que produziram fragmentos com sinal de hibridização predominante forte.

* Espécies utilizadas como controle.

obtermos um sinal razoavelmente equivalente. Uma seleção dos blots obtidos neste trabalho é mostrada nas Figuras 2, 3 e 4.

Os fragmentos de 1,7 kb e 7,0 kb são observados quando o elemento *gypsy* de *D. melanogaster* é submetido à digestão com as enzimas *HindIII* e *BglIII*, respectivamente (ver Figura 1 no Material e Métodos). Fragmentos com estes tamanhos não foram evidenciados para a maioria das espécies investigadas, conforme pode ser visto nas Figuras 2 e 3. A ausência de fragmentos de tamanho correspondente aos encontrados em *D. melanogaster* indica a divergência das seqüências homólogas a *gypsy* nas espécies investigadas, com relação aos sítios de restrição encontrados em *gypsy* de *D. melanogaster*. Resultados similares foram encontrados para as enzimas *KpnI* e *XbaI* (dados não mostrados).

A decisão pela existência de homologia entre o fragmento obtido pela clivagem do DNA genômico da espécie usada como controle (*D. melanogaster*) com aqueles de tamanho molecular aproximado, às vezes observados nas espécies investigadas, é bastante difícil. Principalmente, quando obtemos um grande número de bandas. Decidiu-se por estabelecer o critério de considerar como homólogas, somente aquelas bandas que apresentaram um sinal forte em relação aos demais fragmentos. Este critério foi baseado na suposição de que as bandas cujo sinal é alto são decorrentes do acúmulo de fragmentos produzidos pelo reconhecimento de sítios internos de várias cópias do retroelemento *gypsy* presentes no genoma. Os demais fragmentos produziram sinais fracos por representarem fragmentos de apenas uma cópia ou que tenham perdido um ou ambos os sítios para a

Figura 2 no anexo.

Figura 3 no anexo.

Figura

4

no

anexo.

enzima, ou, ainda, que tenham sítios únicos alternativos de corte para a enzima. Na Tabela 2, salientamos as espécies que produziram possíveis fragmentos homólogos àqueles obtidos na espécie padrão. Foi analisada, também, a ocorrência de fragmentos alternativos, como aqueles esperados para *D. virilis* e *D. subobscura*.

Para *gypsy* de *D. melanogaster*, esperamos um fragmento de 1,66 ou 1,77 kb quando o DNA é clivado com *HindIII*. A produção do primeiro fragmento está relacionada à existência de um sítio de restrição que é marcador de uma subfamília ativa de *gypsy* (Bayev et al. 1984; Chalvet et al. 1998). O fragmento de tamanho esperado (aproximadamente 1,7 kb) parece ser produzido para *D. mediotriata* (Figura 2), *D. simulans* e *D. busckii* (Figura 5). Fragmentos com os tamanhos alternativos e sinal forte foram encontrados em *D. busckii* com tamanho aproximado de 3,4 kb. Este fragmento pode corresponder àquele esperado para *gypsy* de *D. subobscura*. Neste caso, o fragmento computado como o de tamanho 1,7 kb poderia ser o de 1,5kb esperado para *D. subobscura* compondo, assim, o padrão de restrição completo desta espécie. Assim, *D. busckii* e *D. subobscura* podem compartilhar estes sítios de clivagem. As demais bandas de forte sinal (salientadas nas figuras) são, provavelmente, produtos de novos padrões de restrição internos, envolvendo mais de um sítio de restrição no elemento *gypsy*, nestas espécies. As espécies que apresentaram este tipo de banda foram *D. griseolineata*, *D. hydei*, *D. mercatorum*, *D. busckii*, *D. simulans*, *D. tropicalis* e *D. latifasciaeformis*. A diversidade de sítios de restrição entre as espécies é ressaltada na comparação destas bandas predominantes. Nossas análises não

Figura 5 no anexo.

evidenciaram coincidências destas bandas nas diferentes espécies, o que nos leva a postular diferentes padrões de restrição entre diferentes espécies. Com exceção de *D. melanogaster* e *D. simulans*, que apresentaram um padrão de bandas muito semelhantes, todas as outras espécies, inclusive espécies crípticas do grupo *willistoni*, apresentaram padrões diferentes de bandas.

Nos “blots” em que DNAs genômicos das espécies estudadas foram clivados com *Bgl*III (Figura 3), o fragmento esperado de aproximadamente 7 kb para *gypsy* de *D. melanogaster* foi observado, também, em *D. griseolineata*, *D. maculifrons* e *Z. indianus*. Espécies onde são evidenciadas bandas predominantes em sinal são: *D. immigrans*, *D. mediosignata*, *D. hydei*, *D. mercatorum*, *D. prosaltans*, *D. busckii*, *D. pallidipennis*, *D. mediopunctata*, *D. nebulosa*, *D. ananassae*, *D. malerkotliana*, *D. ornatifrons* e *D. griseolineata*. Novamente, estes resultados sugerem que novos padrões de restrição, compostos por dois sítios de reconhecimento, surgiram nas seqüências destas espécies. Nossas análises também não evidenciaram bandas fortes coincidentes entre as espécies investigadas.

A Tabela 3 resume os dados obtidos para a estimativa do número de cópias de seqüências homólogas a *gypsy* nas trinta e duas espécies estudadas. Bandas maiores de 7,5 kb foram consideradas como possíveis cópias completas dos elementos, pressupondo a não existência de sítios internos para a enzima empregada. Assim, os demais fragmentos podem ser considerados cópias deletadas de *gypsy*. A presente estimativa tem algumas limitações, em função de polimorfismos de restrição que possam existir nas seqüências investigadas. Neste sentido, se surgirem sítios para *Bam*HI internos

Tabela 3. Estimativa do número de cópias do retroelemento *gypsy*.

GÊNERO	SUBGÊNERO	GRUPO	ESPÉCIE	Estimativa do número de cópias	Fragmentos maiores que 7,0 kb (prováveis cópias completas)			
<i>Drosophila</i>	<i>Drosophila</i>	<i>guarani</i>	<i>D. maculifrons</i>	09	04			
			<i>D. griseolineata</i>	12	04			
			<i>D. ornatifrons</i>	06	03			
		<i>cardini</i>	<i>D. polymorpha</i>	16	02			
			<i>D. cardinoides</i>	12	04			
			<i>D. neocardini</i>	16	03			
			<i>immigrans</i>	<i>D. immigrans</i>	10	01		
				<i>pallidipennis</i>	15	03		
			<i>tripunctata</i>	<i>D. bandeirantorum</i>	12	01		
				<i>D. angustibucca</i>	13	04		
				<i>D. mediopunctata</i>	10	02		
				<i>D. mediosignata</i>	12	02		
				<i>D. mediopicta</i>	15	03		
		<i>D. mediostriata</i>		12	03			
		<i>repleta</i>		<i>D. hydei</i>	10	03		
				<i>D. mercatorum</i>	07	02		
				<i>D. willistoni</i>	07	01		
		<i>Sophophora</i>		<i>willistoni</i>	<i>D. paulistorum</i>	06	0	
			<i>D. bocainensis</i>		08	02		
			<i>D. tropicalis</i>		13	04		
			<i>D. insularis</i>		10	02		
			<i>saltans</i>		<i>D. prosaltans</i>	08	02	
					<i>melanogaster</i>	<i>D. kikkawai</i>	11	03
			<i>D. melanogaster</i>			23	06	
			<i>D. ananassae</i>			16	02	
			<i>Scaptodrosophila</i>		<i>Scaptodrosophila</i>	<i>D. malerkotliana</i>	11	02
						<i>D. simulans</i>	04	0
						<i>D. lebanonensis</i>	13	03
						<i>D. galloi</i>	13	03
		<i>D. latifasciaeformis</i>		10		01		
		<i>Dorsilopha</i>	<i>Dorsilopha</i>	<i>D. busckii</i>	04	03		
				<i>Zaprionus</i>	14	02		
		<i>Zaprionus</i>	<i>Zaprionus</i>	<i>armatus</i>	<i>Z. indianus</i>	14	02	

ao elemento, o número de cópias estará sendo superestimado. Por outro lado, bandas de alto peso molecular podem conter mais de uma cópia do elemento, principalmente, se eles encontrarem-se num arranjo "in tandem". Esta estimativa torna-se menos confiável para espécies que apresentam bandas predominantes como é o caso de *D. maculifrons*, *D. angustibucca* e *D. pallidipennis*. As demais espécies demonstram um padrão de bandas com sinais homogêneos (Figura 4).

Na tentativa de avaliar a divergência a nível intraespecífico e uma possível atividade deste elemento, foi realizada uma análise com a enzima *HindIII* para as diferentes populações de algumas espécies que dispúnhamos em estoque. Na Tabela 1, são ressaltadas as populações das espécies utilizadas nesta análise. Com exceção de *D. sturtevanti*, onde observamos um certo grau de divergência entre as 6 populações investigadas, para as demais seis espécies não foi observada variação no padrão de bandas. Isto nos leva a inferir a conservação dos sítios internos existentes e uma estabilidade destas seqüências nestes genomas. Contudo, é importante ressaltar que, para a maioria das espécies investigadas, o baixo número e a proximidade das populações (geralmente, coletadas em um raio de 800 km no Sul do Brasil) pode ser um fator importante para a obtenção destes resultados. No entanto, abordagens deste tipo já foram implementadas por nós para *D. simulans* (Loreto et al. 1998b) e *D. willistoni* (Sassi 2000) onde várias populações foram analisadas, de diferentes e distantes localidades, e seus resultados apontam, também, para uma homogeneidade no padrão de bandas obtido entre as amostras, o que sugere estabilidade de *gypsy* a nível populacional.

Embora entre as populações de *D. sturtevanti* (Figura 5) tenham sido encontradas algumas bandas discordantes, os perfis das linhagens COL e BRA são os mesmos, assim como o padrão das populações RP, APA e MAT. A população I27 é bastante similar ao primeiro grupo (COL e BRA), mas contém algumas bandas de tamanhos distintos. Os grupos de padrões formados não demonstram relação geográfica óbvia, a não ser pelas duas populações do México que apresentam o mesmo padrão.

Análise por PCR

De um total de 40 espécies testadas por PCR, empregando 3 diferentes pares de "primers", obtivemos amplificação em 28 dessas espécies, para pelo menos um dos pares de "primers" testados (Tabela 4).

No entanto, apesar de terem apresentado sinais de hibridização com a sonda de *gypsy* de *D. melanogaster*, onze espécies não apresentaram amplificação com nenhum par de "primer". Foram elas: *D. cardinoides*, *D. immigrans*, *D. angustibucca*, *D. mercatorum*, *D. fumipennis*, *D. tropicalis*, *D. sucinea* e *D. insularis*, *D. ananassae* e *D. galloi*. Estes resultados indicam que, nas regiões de anelamento dos "primers", ocorreram substituições nucleotídicas resultando na divergência das seqüências *gypsy* destas espécies.

Um ponto interessante a ressaltar é a distribuição filogenética das espécies que não apresentaram amplificação positiva por PCR. Elas distribuem-se em vários grupos e subgêneros. No entanto, cinco delas

Tabela 4. Resultado das ampliações por PCR para os três pares de primers investigados.

GÊNERO	SUBGÊNERO	GRUPO	ESPÉCIE	GYP1S/ GYP1AS	GYP31S/ GYP31AS	GYP32S/ GYP32AS		
<i>Drosophila</i>	<i>Drosophila</i>	<i>guarani</i>	<i>D. maculifrons</i>	X		X		
			<i>D. griseolineata</i>			X		
			<i>D. ornatifrons</i>			X		
		<i>cardini</i>	<i>D. polymorpha</i>			X		
			D. cardinoides					
			<i>D. neocardini</i>			X		
		<i>immigrans</i>	<i>D. immigrans</i>					
			<i>pallidipennis</i>		X	X	X	
			<i>tripunctata</i>		X	X	X	
		<i>Sophophora</i>	<i>annulimmana</i>	<i>willistoni</i>	<i>D. angustibuca</i>			
					<i>D. mediopunctata</i>	X		X
					<i>D. mediosignata</i>			X
					<i>D. mediopicta</i>	X		X
					<i>D. mediotriata</i>	X		
					<i>D. virilis *</i>	X	X	X
	<i>D. hydei</i>				X	X	X	
	<i>D. zotti</i>				X	X	X	
	<i>D. mercatorum</i>							
	<i>D. bromelioides</i>						X	
	<i>Scaptodrosophila</i>	<i>melanogaster</i>	<i>willistoni</i>	<i>D. annulimmana</i>	X	X	X	
				<i>D. willistoni</i>			X	
				<i>D. nebulosa</i>		X	X	
				<i>D. fumipennis</i>				
				<i>D. paulistorum</i>			X	
				<i>D. bocainensis</i>				
				<i>D. capricorni</i>		X		
				<i>D. equinoxialis</i>	X			
				<i>D. tropicalis</i>				
				<i>D. sucinea</i>				
	<i>Zaprionus</i>	<i>Zaprionus</i>	<i>armatus</i>	<i>D. insularis</i>				
				<i>D. prosaltans</i>		X		
				<i>D. sturtevanti</i>	X	X		
				<i>D. kikkawai</i>	X	X		
<i>D. melanogaster *</i>				X	X	X		
<i>D. ananassae</i>								
<i>D. malerkotliana</i>				X	X			
<i>D. simulans</i>				X	X	X		
<i>D. lebanonensis</i>					X	X		
<i>D. galloi</i>								
<i>D. latifasciaeformis</i>								
<i>D. busckii</i>	X	X	X					
<i>Z. indianus</i>	X	X	X					

* Espécies utilizadas como controle.

pertencem ao grupo *willistoni* (cerca de 40% da amostra). Desta forma, parece que este grupo reúne as seqüências homólogas à *gypsy* mais divergentes, em relação às das espécies utilizadas para o desenho dos “primers”.

D. pallidipennis, *D. bandeirantorum*, *D. hydei*, *D. zotti*, *D. annulimana*, *D. simulans*, *D. busckii* e *Z. indianus* tiveram os fragmentos esperados para as ampliações por PCR com os três conjuntos de “primers” empregados. Além disto, as bandas obtidas como produtos das reações de PCR eram equivalentes às obtidas nos controles (*D. melanogaster* e *D. virilis*) (Figura 6). Cabe salientar, que a maioria destas espécies são distantes filogeneticamente das espécies de onde as seqüências *gypsy* foram usadas para o desenho dos “primers”. Exceções a esta constatação foram *D. simulans*, por ser uma espécie muito próxima à *D. melanogaster*, e *D. hydei* e *D. zotti* por estarem próximas ao grupo da *D. virilis*, fazendo parte da mesma radiação (radiação *virilis-repleta*) (Trockmorton 1975).

Análises Filogenéticas

As seqüências empregadas para as inferências filogenéticas, incluindo os “gaps” necessários para os alinhamentos, totalizaram 499 nucleotídeos de 64 amostras submetidas à análise. Destes, 116 mostraram-se conservados, 29 foram caracterizados como autapomorfias e 354 foram considerados

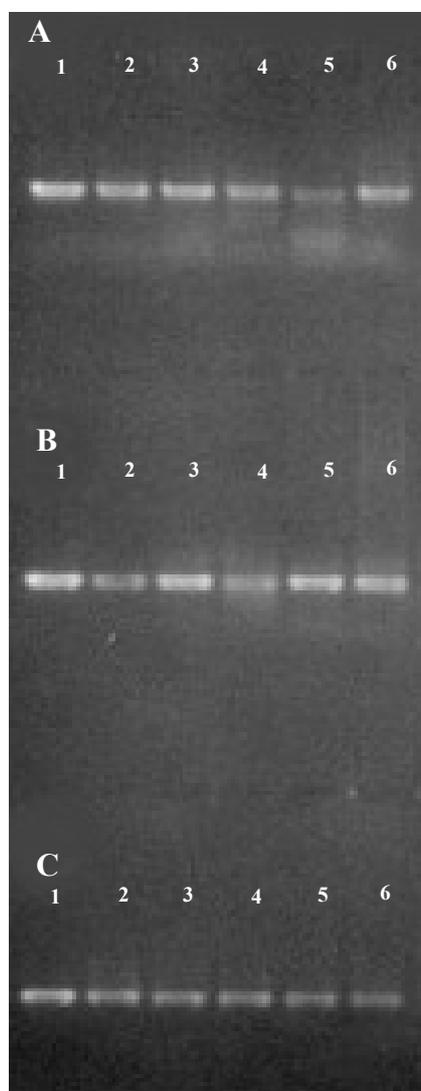


Figura 6. Resultados de algumas reações de amplificação por PCR de seqüências do elemento *gypsy* em espécies do gênero *Drosophila* e DNA controle. A) amplificações utilizando o par de primers GYP1S/GYP1AS: 1) DNA controle - sonda do elemento *gypsy* inserida no plasmídeo pGGHS; 2) *D. melanogaster*; 3) *D. busckii*; 4) *D. simulans*; 5) *D. bandeirantorum*; 6) *D. hydei*; B) amplificações utilizando o primer GYP31S/GYP31AS: 1) DNA controle - sonda do elemento *gypsy* inserida no plasmídeo pGGHS; 2) *D. melanogaster*; 3) *D. busckii*; 4) *D. simulans*; 5) *D. bandeirantorum*; 6) *D. hydei* amplificações utilizando o primer GYP32S/GYP32AS: 1) DNA controle - sonda do elemento *gypsy* inserida no plasmídeo pGGHS; 2) *D. melanogaster*; 3) *D. busckii*; 4) *D. simulans*; 5) *D. bandeirantorum*; 6) *D. hydei*.

informativos para a análise de parcimônia.

A taxa média de substituição entre todos os táxons foi de 32,7%, variando de 0,5% a 63,2%. As seqüências de *D. willistoni* mostraram-se muito diferentes das demais amostras, sendo que a taxa de substituição encontrada entre esta espécie e as demais variou de 51,2% (em relação à *D. griseolineata*) a 63,2% (em relação à *D. mediopunctata*, clones 49 e 86). Assim, para verificar se a alta taxa de substituição encontrada nesta espécie estava prejudicando a análise, testamos sua retirada, a fim de comparar os agrupamentos obtidos com e sem *D. willistoni*. A comparação dos resultados evidenciou que os principais agrupamentos mantiveram-se constantes e optamos, assim, por conservar *D. willistoni* na análise. Esta decisão foi tomada exatamente por nos parecer extremamente interessante a sua divergência e a sua posição, em relação às demais espécies. Contudo, para confirmar o grau de divergência de nucleotídeos obtido nos clones analisados, provenientes desta espécie (população RS), seqüenciamos e analisamos a mesma região de gypsy de indivíduos de outra população, geograficamente distante daquela inicial (população PA). Estas duas populações encontram-se nos extremos do Brasil, distantes mais de 3.500 km. Mesmo assim, mostraram-se extremamente homogêneas quanto à seqüência nucleotídica do elemento gypsy (variando de 0 a 1% de divergência).

A árvore obtida pelo método de Neighbor Joining (Figura 7) é constituída de quatro grandes grupos, que podem, ainda, ser divididos em 8 agrupamentos, de acordo com seus graus de divergência. Foram considerados grupos distintos, aqueles que apresentavam graus de

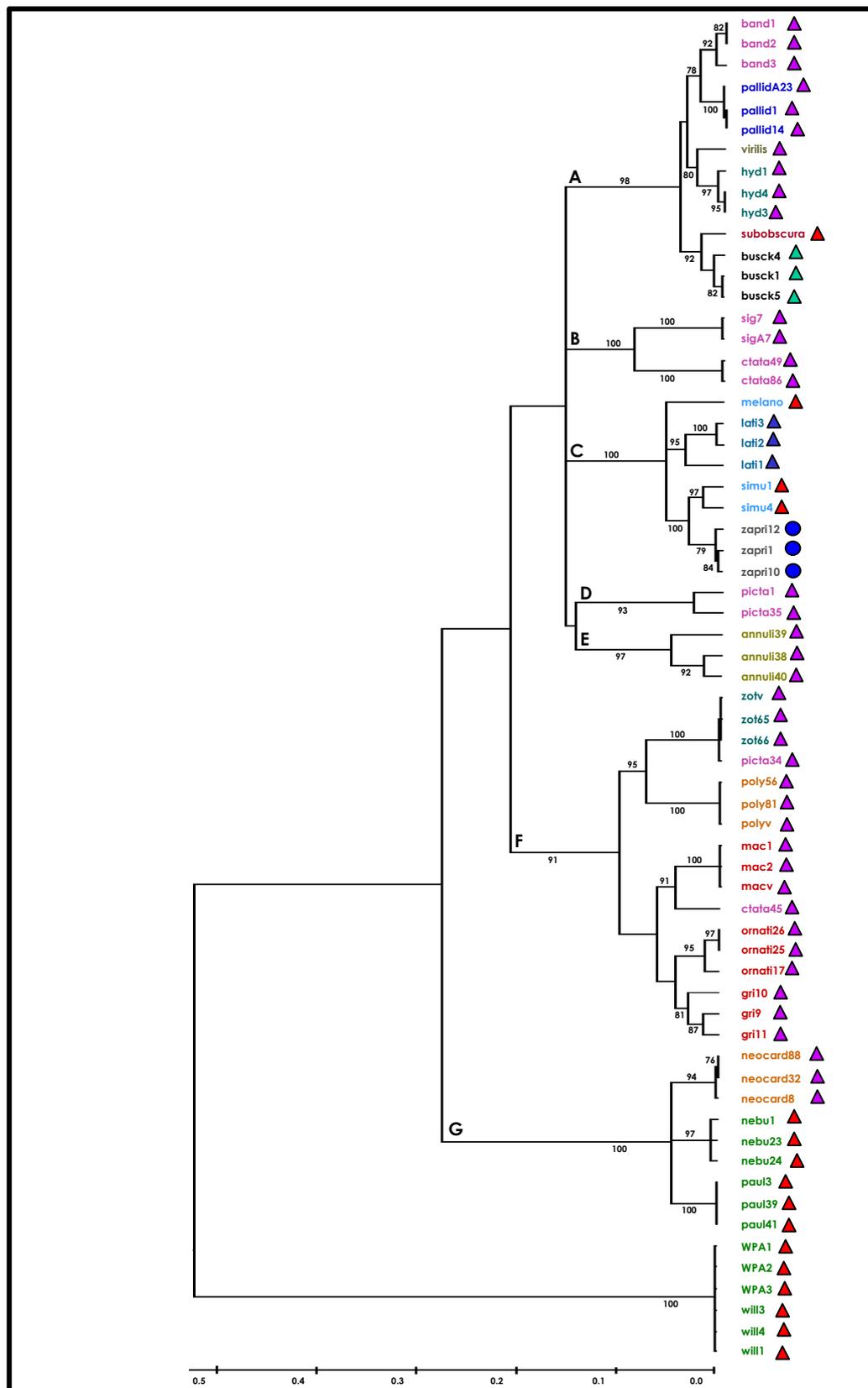


Figura 7. Análise filogenética de 499 nucleotídeos de uma região da ORF 3 do retroelemento *gypsy*. Árvore Neighbor Joining baseada da distância Kimura dois parâmetros. Os números nos ramos representam os valores de bootstrap menores do que 75 para 100 amostragens. Táxons de cores diferentes pertencem a diferentes grupos de espécies. Símbolos diferentes correspondem a gêneros distintos (círculos representam o gênero *Zaprionus* e triângulos o gênero *Drosophila*). Triângulos de cores diferentes representam distintos subgêneros (vermelho - *Sophophora*; lilás - *Drosophila*; azul escuro - *Scaptodrosophila* e verde - *Dorsilopha*). As letras nos ramos indicam as diferentes subfamílias de *gypsy*, propostas neste trabalho.

divergência acima de 20%. Assim, os 8 grupos foram denominados por letras de A a H, de acordo com sua posição na Figura 7 (de cima para baixo). Os agrupamentos compreendem os seguintes táxons: A) três clones de *D. bandeirantium* (band1, band2 e band3) agrupados com três clones de *D. pallidipennis* (pallidA23, pallid1e pallid14), este grupo junta-se com outro formado por *D. virilis* (virilis) e três clones de *D. hydei* (hyd1, hyd 4 e hyd3), por fim, estes primeiros agrupamentos unem-se à *D. subobscura* (subobscura) e aos três clones de *D. busckii* (busck4, busck1 e busck5); B) formado por dois clones de *D. mediosignata* (sig7 e sigA7) e dois clones de *D. mediopunctata* (ctata49 e ctata86); C) compreende os dois clones de *D. simulans* (simu1 e simu4) agrupados com *Zaprionus indianus* (zapri12, zapri1 e zapri10), seguido do agrupamento dos três clones de *D. latifasciaeformis* (lati3, lati2 e lati1) e, por fim, de *D. melanogaster* (melano); D) formado por dois clones de *D. mediopicta* (picta1 e picta35); E) agrupamento formado pelos três clones de *D. annulimana* (annuli39, annuli38 e annuli40); F) constituído dos grupos formados pelos três clones de *D. zotti* (zotv, zot65 e zot66) e um clone de *D. mediopicta* (picta34), agrupando-se com os três clones de *D. polymorpha* (poly56, poly81 e polyv), estes táxons formam um ramo que se agrupa com outro formado pelos três clones de *D. ornatifrons* (ornati26, ornat25 e ornat17), com os três clones de *D. griseolineata* (gri10, gri9 e gri11) estando, ainda, este grupo junto aos três clones de *D. maculifrons* (mac1, mac2 e macv) e um clone de *D. mediopicta* (picta 34); G) agrupamento formado pelos três clones de *D. paulistorum* (paul3, paul 39 e paul 41), junto com os três clones de *D. nebulosa* (nebu1, nebu23 e nebu24) e três clones de *D. neocardini* (neoc88,

neoc32 e neoc8); H) grupo formado exclusivamente pelos clones provenientes de *D. willistoni*, três deles sendo da população do Rio Grande do Sul (will3, will4 e will1) e três da população do Pará (WPA1, WPA3 e WPA2). A divergência média encontrada dentro destes grupos é de 6,7%, 10,8%, 9,8%, 5,5%, 7,5%, 13%, 7,2% e 0,4%, respectivamente. Na Tabela 5 encontram-se as taxas de divergências médias entre os grupos estabelecidos.

No genoma das espécies *D. mediopunctata* e *D. mediopicta* existem seqüências homólogas a *gypsy* com graus de divergência que as colocam em agrupamentos distintos na árvore filogenética. Para *D. mediopunctata*, a divergência entre o clone tacta45, interno ao grupo estabelecido F, em relação aos outros, pertencentes ao grupo B, é de 34,8%. Para *D. mediopicta*, a distância nucleotídica estimada entre o clone picta34 do agrupamento F e os demais (picta1 e picta35 do grupo D) é de, em média, 30,8%.

Na Figura 8, podemos observar a árvore consenso estrito obtida a partir de 143 árvores igualmente mais parcimoniosas, todas contendo 594 passos. Podemos evidenciar em sua topologia, os mesmos 8 grupos citados, anteriormente, obtidos por Neighbor Joining. Valores de "bootstrap" menores do que 75 foram omitidos de todas as árvores deste trabalho. Um ponto importante a ser ressaltado é o fato de que os ramos que levam aos 8 agrupamentos são bem suportados por altos valores de "bootstrap" nas duas árvores (Figura 7 e 8).

Para a análise de seqüências de aminoácidos, alguns táxons iniciais foram excluídos. Retiramos as seqüências de *D. willistoni*, pois o número de substituições existentes produziu uma seqüência de aminoácidos com sete "stop" códons e com um elevado grau de divergência em relação às demais

Tabela 5. Percentagem de divergência entre os agrupamentos (de A a H) formados na árvore filogenética obtida a partir dos dados de seqüência de nucleotídeos de *gypsy*.

	A	B	C	D	E	F	G	H
A								
B	27,2%							
C	24,1%	30,8%						
D	22,6%	28,4%	23,3%					
E	23,0%	32,9%	28,4%	23,5%				
F	30,0%	36,8%	32,3%	30,0%	34,4%			
G	38,2%	40,8%	41,5%	44,3%	40,6%	34,9%		
H	55,9%	62,1%	57,8%	52,3%	54,8%	53,6%	57,5%	

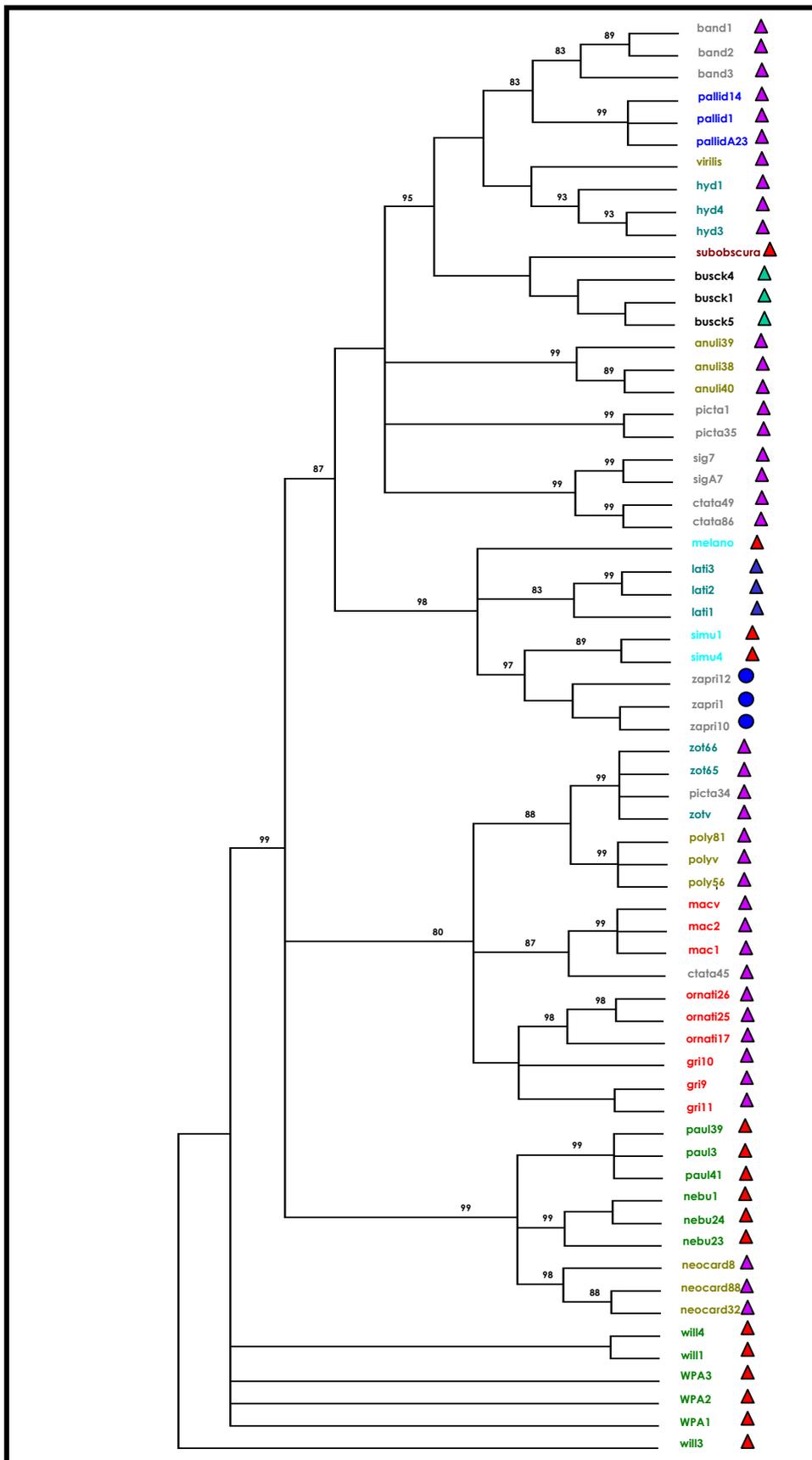


Figura 8. Análise filogenética de 499 nucleotídeos de uma região da ORF 3 do retroelemento gypsy. Árvore consenso, não enraizada, gerada pelo método de máxima parcimônia. Os números nos ramos representam os valores de bootstrap maiores que 75 para 100 amostragens. Táxons de cores diferentes pertencem a distintos grupos de espécies. Símbolos diferentes correspondem a gêneros distintos (círculos representam o gênero *Zaprionus* e triângulos o gênero *Drosophila*). Triângulos de cores diferentes representam distintos subgêneros (vermelho - *Sophophora*; lilás - *Drosophila*; azul escuro - *Scaptodrosophila* e verde - *Dorsilopha*).

espécies. Além disto, na análise das prováveis seqüências de aminoácidos, as espécies foram representadas somente por clones que diferiam em mais de 1% de suas seqüências nucleotídicas. Assim, resultaram 40 táxons que foram alinhados códon a códon, de acordo com seu módulo de leitura. Para este ajuste, as seqüências de *D. melanogaster*, *D. virilis* e *D. subobscura*, cuja tradução em aminoácidos é conhecida, foram utilizadas como padrão. Foi necessário que incluíssemos alguns "gaps" no alinhamento original para compensar inserções existentes. Os seguintes ajustes foram efetuados: dois "gaps" entre as posições 54 e 55 de *D. paulistorum*; dois "gaps" entre as posições 359 e 360 de todas seqüências pertencentes ao agrupamento F e *D. paulistorum*; e, finalmente, um "gap" entre os nucleotídeos de número 384 e 385 para *D. mediopunctata* (clones 49 e 86), para todas as seqüências do agrupamento F e G, com exceção de *D. paulistorum*.

A região do gene *env*, analisada neste estudo, corresponde a uma porção com potencial para codificar 151 aminoácidos. Das seqüências analisadas, várias continham "stop" códons como por exemplo, *D. hydei* (hyd1, hyd3 e hyd4), *D. polymorpha* (poly81), *D. mediopunctata* (ctata49), *D. annulimana* (annuli38 e annuli40), *D. neocardini* (neoc88), *D. griseolineata* (gri10, gri11), *D. simulans* (simu1), *D. annulimana* (anuli38), *D. paulistorum* (paul39), *D. latifasciaeformis*(lati3) e *D. ornatifrons* (ornati17). No entanto, quando inserções/deleções eram corrigidas e o módulo de leitura era restabelecido, podia-se verificar que essas seqüências permaneciam relativamente conservadas após o códon de terminação. Assim, estes táxons

permaneceram na análise, inclusive para cálculo de taxa de substituição sinônima e não sinônima, bem como para o teste de seleção.

O alinhamento produzido pela provável seqüência de aminoácidos revelou 20 posições conservadas, 103 informativas e 25 autapomorfias. As filogenias obtidas (pelos métodos de Neighbor Joining e Máxima Parcimônia) a partir da análise das seqüências de aminoácidos, resultaram em topologias muito semelhantes entre si, que separam os sete primeiros grupos observados na primeira análise (Figura 9), mas que demonstram alguns pontos internos de divergência, quando comparadas com aquelas obtidas das seqüências de nucleotídeos. As principais mudanças internas foram: seqüências de *Z. indianus*, que nesta análise se agrupa com *D. latifascaeiformis* e não mais com *D. simulans*, esta última agora formando um grupo com *D. melanogaster* que apresentava-se distante nas análises anteriores; seqüências de *D. pallidipennis* e *D. virilis* apresentam-se agrupadas, quando nas topologias anteriores ocorriam juntas a *D. bandeirantorum* e a *D. hydei*, respectivamente; seqüências de *D. maculifrons* deixam o agrupamento formado com *D. mediopunctata* (clone49) e agrupam-se com seqüências de *D. griseolineata*. A taxa de substituição média obtida para estes dados foi de 30,6%. A divergência interna dos grupos formados nesta análise foi: 7% (grupo A), 10,1% (grupo B), 8,5% (grupo C), 8,7% (grupo E), 14,5% (grupo F) e 15,5% (grupo G). Este cálculo não foi realizado para o grupo D pois ele é composto por apenas uma seqüência nesta análise. A Tabela 6 contém os índices de divergência entre os grupos.

As seqüências codificantes foram submetidas a uma análise de taxas de substituição sinônimas (dS) e não sinônimas (dN). As taxas médias obtidas

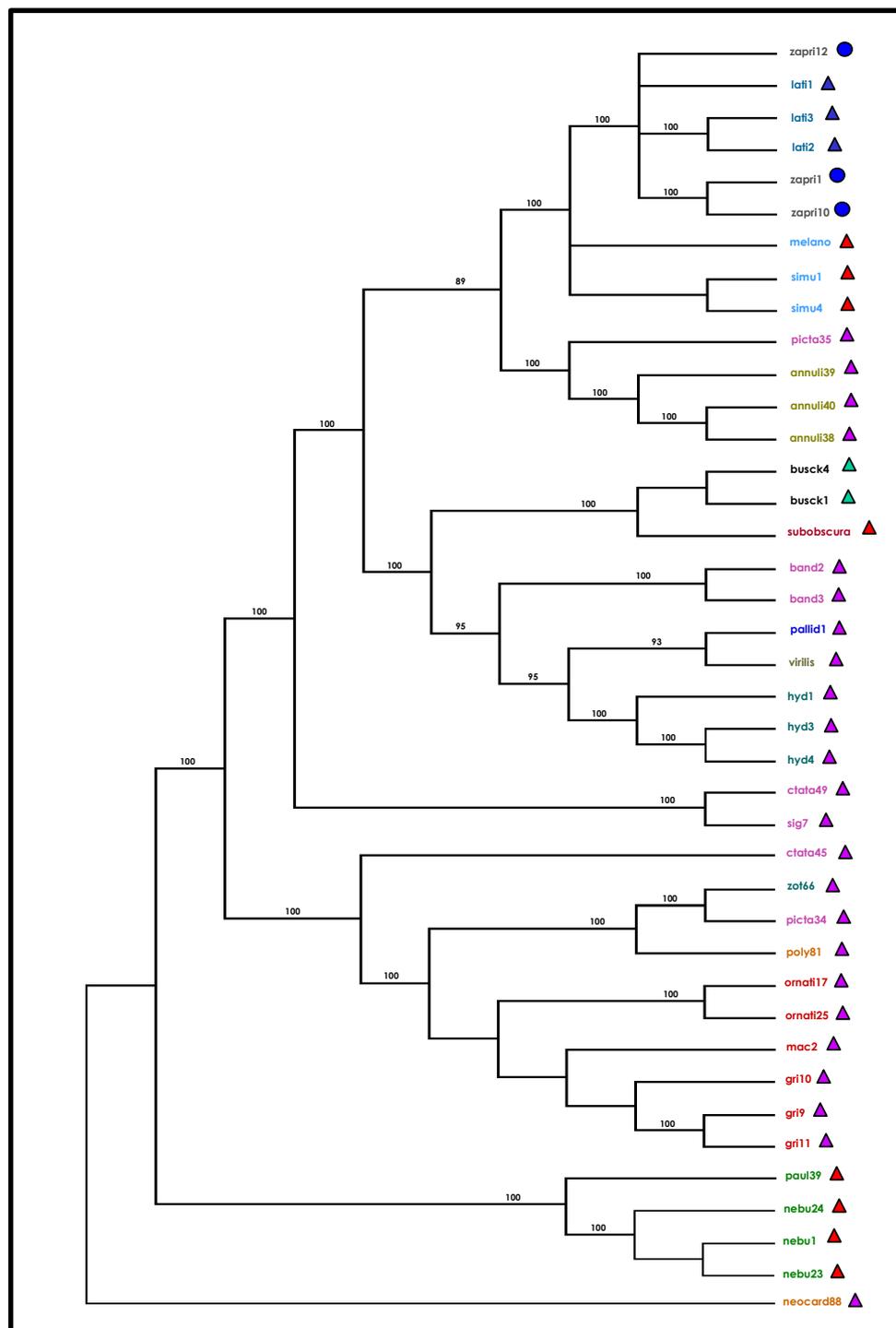


Figura 9. Análise filogenética baseada na provável seqüência de aminoácidos obtida a partir da seqüência de 499 nucleotídeos de uma região da ORF 3 do retroelemento gypsy. Árvore consenso, não enraizada, gerada pelo método de máxima parcimônia. Os números nos ramos representam os valores de bootstrap maiores que 75 para 100 amostragens. Táxons de cores diferentes pertencem a distintos grupos de espécies. Símbolos diferentes correspondem a gêneros distintos (círculos representam o gênero *Zaprius* e triângulos o gênero *Drosophila*). Triângulos de cores diferentes representam distintos subgêneros (vermelho - *Sophophora*; lilás - *Drosophila*; azul escuro - *Scaptodrosophila* e verde - *Dorsilopha*).

Tabela 6. Percentagem de divergência entre os agrupamentos (de A a G) formados na árvore filogenética obtida a partir dos dados da provável seqüência de aminoácidos de *gypsy*.

	A	B	C	D	E	F	G
A							
B	30,0%						
C	20,3%	30,1%					
D	17,0%	30,4%	16,3%				
E	23,6%	36,2%	22,9%	12,6%			
F	35,1%	46,4%	39,0%	39,0%	39,5%		
G	44,9%	52,2%	47,2%	49,9%	50,1%	42,1%	

para estes dois parâmetros foram de 61,5% e 17,7%, respectivamente. A relação dN/dS foi de 0,29. Esta relação nos fornece uma estimativa de que, em média, mudanças sinônimas em sítios sinônimos estão ocorrendo cerca de 3,5 vezes com mais frequência do que aquelas não sinônimas em sítios não sinônimos, o que indica que seleção purificadora possa estar atuando na região analisada. Um teste de seleção foi realizado para cada par de seqüências.

O teste de seleção realizado evidenciou neutralidade para as seguintes comparações: entre *gypsy* de *D. hydei* e *D. virilis*; entre clones provenientes de *D. latifasciaeformis*, *Z. indianus*, *D. annulimana* e *D. griseolineata*; entre as espécies *D. ornatifrons* e *D. griseolineata* e, finalmente, entre todas as seqüências oriundas do genoma das espécies pertencentes ao agrupamento G da filogenia (Figuras 7 e 8). As demais comparações resultaram em rejeição da hipótese nula (neutralidade) e aceitação da hipótese de seleção purificadora.

DISCUSSÃO

Abordagens utilizando a técnica de *Southern blot* e sondas heterólogas têm se mostrado, extremamente, úteis em amplos "screenings" de elementos transponíveis. Talvez este método seja a melhor escolha para se avaliar a distribuição de elementos transponíveis. Análises por PCR, mesmo com o emprego de "primers" degenerados, muitas vezes falham em detectar a presença de um dado elemento, devido a mutações ocorridas na região de

anelamento do "primer", como foi mostrado em nossos resultados. Análises por *Southern blot* prestam-se, também, para estimar a divergência entre sonda e seqüências homólogas em análise, tanto por quantificação do sinal de hibridização em diferentes estringências, como por padrões obtidos com a utilização de diferentes enzimas de restrição. Análises de seqüências realizadas neste trabalho indicam que inferências baseadas em resultados com a metodologia acima referida têm se mostrado bastante confiáveis.

O primeiro trabalho sobre distribuição de *gypsy*, dentro de gênero *Drosophila*, foi realizado por Stacey et al. (1986), utilizando o método de *Southern blot*. Seus resultados já indicavam a ubiqüidade deste elemento. Além disto, os sinais de hibridização que os autores obtinham apontavam um padrão de conservação desta seqüência que era inconsistente com o modelo de origem ancestral única, seguida de transmissão vertical. Os autores chamavam à atenção para o fato de que transmissão horizontal poderia explicar a grande homologia que parecia existir entre seqüências *gypsy* de espécies distantemente relacionadas, como por exemplo, entre *D. busckii* e *D. melanogaster*.

De Frutos et al. (1992) empregaram a mesma metodologia de Stacey et al. (1986) para investigar o grupo *obscura* de *Drosophila*. Os autores demonstraram, utilizando uma enzima de restrição que clivasse nas extremidades (LTRs) do elemento de *D. melanogaster* (*Xho*I), que cópias completas existiam no genoma de várias espécies. Contudo, fragmentos menores, compartilhados por algumas espécies, indicavam elementos deletados ou polimorfismos de sítios de restrição. Os autores ressaltaram a divergência inesperada encontrada entre a sonda de *D. melanogaster* e as

seqüências investigadas do grupo *obscura*, já que estes dois grupos de espécies são filogeneticamente bastante relacionadas (Alberola e De Frutos 1993a). Alberola e De Frutos (1993b) analisaram mais detalhadamente as seqüências de *gypsy* de espécies do subgrupo *obscura*, utilizando várias enzimas de restrição, e sugerem um padrão de transmissão vertical dentro deste subgrupo. Estes autores utilizaram as mesmas enzimas empregadas no presente trabalho (*Bam*HI e *Bg*III) e também observaram um padrão heterogêneo de bandas, de difícil análise e que apontava para um mapa de restrição divergente daquele conhecido para *D. melanogaster*. Apesar da constatação da existência de cópias completas, em nenhuma das oito espécies analisadas por aqueles autores, foi encontrado o fragmento esperado de 7,0 kb, liberado pela ação da enzima *Bg*III.

D. griseolineata, *D. maculifrons* e *Z. indianus* parecem possuir cópias completas de *gypsy*, como evidenciado pelos resultados de *Southern blot* com a enzima *Bg*III. Contudo, não podemos descartar a existência deste tipo de seqüência no genoma das demais espécies investigadas, pois a perda de um dos sítios reconhecidos por esta enzima pode levar a um padrão diferente do esperado para *D. melanogaster*. A julgar pelos nossos resultados, este tipo de variação é bastante freqüente. É interessante ressaltar que as três espécies possuidoras de seqüências homólogas a *gypsy*, cujo padrão de clivagem da enzima *Bg*III é mantido, são extremamente distantes filogeneticamente de *D. melanogaster*. Enquanto isto, os sítios de restrição não são mais observados nas demais espécies do grupo *melanogaster*, nem em outras espécies mais aparentadas (por exemplo aquelas pertencentes ao subgênero *Sophophora*).

Bayev et al. (1984), em um dos primeiros trabalhos de caracterização do elemento *gypsy* de *D. melanogaster*, estimou que 20 seria o número total de cópias desta seqüência no genoma desta espécie. Destas, poucas estariam nos braços cromossômicos (em torno de 5) e o restante seriam cópias presentes no cromocentro, sub-representadas na parte eucromática dos cromossomos politênicos das glândulas salivares. Nossa estimativa do número de cópias para *D. melanogaster* (23), nossa espécie controle, resultou em números bem próximos a estes. Outra conclusão resultante das análises destes autores é que co-existem no mesmo genoma cópias de *gypsy* que divergem quanto aos seus sítios de restrição. As cópias completas, presentes nos braços eucromáticos, apresentariam um padrão de clivagem, enquanto cópias deletadas e divergentes estariam confinadas à região heterocromática. Entre as cópias completas de *gypsy*, os autores caracterizaram duas subfamílias: uma ativa (representada por poucas cópias) e a outra inativa (sendo a maioria das cópias completas), que também diferem em sítios de restrição.

Cópias completas, representantes da subfamília transcricionalmente ativa, possuem sítios de restrição internos para *HindIII*, produzindo um fragmento de 1,7 kb (Bayev et al. 1984; Lambertsson et al. 1989). Assim, nossos resultados apontam para a existência de elementos, potencialmente ativos, somente nas espécies *D. mediotriata*, *D. busckii* e *D. simulans*. Novamente, é notável a distância filogenética entre as duas primeiras espécies, em relação à *D. melanogaster*. Apesar de apresentarem o sítio interno de *HindIII*, as seqüências de *gypsy* de *D. busckii* e *D. simulans*, quando investigadas a nível intra-populacional, evidenciam perfis de estabilidade genômica. Nas demais

espécies, as seqüências investigadas não demonstraram possuir os sítios diagnósticos de reconhecimento desta enzima.

Lambertsson et al. (1989) isolaram 21 clones contendo seqüências homólogas à *gypsy* de uma única linhagem de *D. melanogaster* e detectaram uma heterogeneidade estrutural entre estes clones, caracterizada por um extensivo polimorfismo de restrição. Os autores estabeleceram três subgrupos: 1) elementos completos; 2) elementos levemente divergentes e; 3) elementos altamente divergentes, cada um sendo representado por várias cópias no genoma. Esta heterogeneidade poderia explicar o padrão complexo de bandas obtido em nossas observações e abre algumas questões interessantes sobre o tempo de existência de seqüências *gypsy* no genoma de *Drosophila*, necessário para o padrão de diversificação encontrado. Esta divergência encontrada pode ser conseqüência de uma longa co-existência de TEs e genomas hospedeiros. Por outro lado, mecanismos de controle internos deste genoma hospedeiro, contra DNA parasítico, reconhecendo e marcando seqüências homólogas repetidas, poderiam estar acelerando este processo de divergência. A ampla distribuição de *gypsy* no gênero parece apoiar a primeira hipótese.

Dada a divergência encontrada para seqüências do elemento *gypsy*, sítios de restrição conservados entre diferentes espécies podem indicar cópias ou regiões da seqüência sob pressão seletiva. Ou, ainda, tal conservação pode ser resultado de uma ancestralidade recente. Este parece ser o caso para as espécies próximas. Padrões semelhantes entre *D. melanogaster* e *D. simulans* podem exemplificar esta relação. No entanto, se padrões de clivagem semelhantes forem encontrados em espécies filogeneticamente

distantes, um evento de transmissão horizontal pode ser evocado. O padrão de bandas esperado para *D. subobscura* (subgênero *Sophophora*) encontrado em *D. busckii*, uma espécie pertencente a outro subgênero (*Dorsilopha*), poderia exemplificar esta última hipótese.

O elemento *gypsy* isolado do genoma de *D. virilis* mostrou-se bastante conservado quando comparado com aquele de *D. melanogaster*, sendo esta conservação suficiente para ele ser transcrito. A presença de *gypsy* em outras espécies do mesmo grupo de *D. virilis* indicou, no entanto, divergência de padrões de restrição entre as seqüências *gypsy* destas espécies. *Gypsy* foi investigado em diferentes linhagens de *D. virilis*, de diferentes regiões geográficas, e foram constatadas variações na localização de algumas cópias entre as amostras, tanto em ensaios de *Southern blot* como por hibridização *in situ* (Mizrokhi e Mazo 1991). Estes resultados apontam para uma certa atividade transposicional de *gypsy*, o que condiz com sua capacidade transcricional. Resultados semelhantes foram observados em diferentes populações de *D. subobscura* (Alberola e De Frutos, 1993b). Em nossas análises não constatamos este tipo de fenômeno, talvez porque a capacidade para a transposição esteja faltando nas seqüências investigadas neste trabalho. A conservação dos padrões obtidos para *gypsy* e, provavelmente, de sua localização genômica, entre as diferentes populações de *D. polymorpha*, *D. immigrans*, *D. mercatorum*, *D. simulans* e *D. busckii*, investigadas neste trabalho, indicam inativação destas seqüências móveis. Resultados semelhantes foram observados por Bucheton (1995) para populações distantemente relacionadas de *D. melanogaster*. Loreto et al. (1998b) investigaram a variação existente neste elemento para populações de *D.*

simulans. Neste trabalho, os autores estimam um baixo número de seqüências homólogas a *gypsy* no genoma desta espécie, fato que se repete em nossas estimativas.

Estimativas por hibridização *in situ* apontam um número semelhante de cópias de *gypsy* para *D. melanogaster* e *D. simulans* (Biémont e Cizeron 1999). No entanto, esta abordagem não dá conta das cópias existentes em regiões de heterocromatina. No presente trabalho, nossas estimativas (23 cópias para *D. melanogaster* e 4 para *D. simulans*) podem refletir a variação desta porção genômica entre estas duas espécies. Um menor número de cópias de TEs no genoma de *D. simulans*, quando comparado com os dados provenientes de sua espécie críptica, *D. melanogaster*, é observado para muitas famílias de elementos móveis. Este fenômeno caracteriza a menor quantidade de DNA medianamente repetitivo existente na primeira espécie (Dowsett et al. 1982; Vieira e Young. 1999).

Estimativas confiáveis do número de cópias requerem conservação do mapa de restrição utilizado como padrão para as seqüências investigadas. Assim, nossos resultados de número de cópias para aquelas espécies em que não observamos os sítios esperados, para nenhuma das duas enzimas utilizadas no trabalho, devem ser tomados com cautela. Além disso, a alta divergência pode fazer com que bandas deixem de ser detectadas e contabilizadas como novas cópias.

Quando comparamos os resultados obtidos pela técnica de *Southern blot* e por PCR, verificamos que uma alta homologia parece existir entre as seqüências de *gypsy* de *D. melanogaster* e *Z. indianus*. A conservação de alguns dos sítios de restrição do elemento *gypsy* de *D. melanogaster* em *Z.*

indianus (para a enzima *Bgl*III) e o forte produto de PCR para os três pares de “primers” utilizados neste trabalho corroboram esta afirmativa. Tal constatação nos parece surpreendente, já que estas espécies são classificadas em dois gêneros distintos.

Um cenário de ampla distribuição para elementos transponíveis é bem documentado para outros elementos da Classe I como *copia* (Stacey et al. 1986), *412* (Cizeron et al. 1998), *1731* (Montchamp-Moreau et al. 1993) e *bilbo* (Blesa et al. 2001). Juntamente com *gypsy*, estes TEs são postulados a existir no genoma de espécies ancestrais do gênero *Drosophila*. A divergência encontrada entre elementos de espécies distantes, por padrões diferenciais de clivagem por enzimas de restrição, também justificam esta afirmação. Para *bilbo* e *1731*, os autores detectaram uma relação entre o sinal e o padrão de hibridização obtido nos experimentos com a história evolutiva de suas espécies hospedeiras e sugerem que transmissão vertical seja o mecanismo principal de expansão destas seqüências móveis. Elementos como *412* e *copia* já demonstram padrões mais complexos de divergência, onde o parentesco filogenético das espécies é incongruente com os resultados de análise destes elementos. Nestes casos, eventos de transmissão horizontal são citados como importantes para explicar estas discordâncias. Este último padrão evolutivo parece ser o que mais se adapta aos nossos resultados para o elemento *gypsy* e concordam com análises realizadas por outros autores (Stacey et al. 1986, De Frutos et al. 1992, Terzian et al. 2000).

Quando construímos filogenias, a partir de elementos transponíveis, ou seja, quando tentamos reconstruir as relações existentes entre estas seqüências, presentes nos mais diferentes genomas hospedeiros muitas vezes

nos deparamos com cenários complexos. A evolução de TEs é pensada em ser permeada de diferentes fenômenos tais como: 1) polimorfismo ancestral resultante de cópias que evoluem independentemente no genoma; 2) taxas evolutivas desiguais dependentes de mecanismos de regulação do hospedeiro, de fatores ambientais sobre o organismo, dosando seu valor adaptativo ou da domesticação destas seqüências; 3) transmissão vertical; 4) transmissão horizontal mediada por vetores ou por características intrínsecas do elemento (como é o caso de retroelementos capazes de formar a proteína do envelope e ter uma parte de seu ciclo infectivo de forma extracelular, como *gypsy*); 5) reorganização de elementos inativos e/ou incompletos no genoma, originando elementos ativos com novas propriedades; e, por fim, 6) perdas estocásticas por erro amostral em eventos de especiação. O que torna mais complicado o entendimento da evolução de um dado TE é que, na maioria das vezes, estes fenômenos não são mutuamente exclusivos (revisões em Capy et al. 1998; Pinsker et al. 2001).

Para a análise dos diferentes fenômenos que possam estar atuando na história evolutiva de um dado elemento é necessário que as relações filogenéticas estabelecidas no estudo sejam comparadas com as relações de parentesco das espécies que hospedam, em seu genoma, tais seqüências móveis.

Apesar de *Drosophila* representar um organismo modelo para as mais distintas áreas da biologia, filogenias moleculares que envolvam um amplo espectro de espécies do gênero são escassas. A maioria delas aborda um pequeno número de espécies em torno da espécie referência, que é *Drosophila melanogaster*. Neste sentido, o subgênero *Sophophora* é, de longe,

o mais explorado. O subgênero *Drosophila*, que contém o maior número de espécies neotropicais, é subrepresentado nas filogenias em questão. Assim, relações entre espécies envolvendo todo o gênero são, ainda, citadas a partir dos dados de Throckmorton (1975) com caracteres morfológicos que, mais tarde, foram complementados e re-analisados sob uma perspectiva filogenética por Grimaldi (1990).

As seqüências do genoma mais amplamente estudadas, sob uma perspectiva filogenética, são Adh (gene para enzima álcool desidrogenase) (Russo et al. 1995), rDNA (Pelandakis e Solignac 1993), Sod (gene para superóxido dismutase) (Kwiatowski et al. 1994) e mtDNA (DeSalle 1992). Uma análise comparativa dos principais pontos conflitantes entre filogenias de dados moleculares e entre estas e aquelas obtidas a partir de dados morfológicos é implementada por Kwiatowski e Ayala (1999). Powell e DeSalle (1995) também revisaram os principais resultados obtidos em diferentes abordagens para as relações entre as espécies de *Drosophila*.

Um ponto importante de discussão entre as diferentes filogenias é a posição de *Zaprionus*. Muitas análises colocam *Zaprionus* dentro do subgênero *Drosophila* a partir de dados morfológicos de Throckmorton (1975) e dados moleculares de rDNA de Pelandakis e Solignac (1993), Sod investigada por Kwiatowski et al. (1994) e Adh de Russo et al. (1995), enquanto outras evidenciam um caráter basal para este grupo e o classificam como um gênero a parte do gênero *Drosophila* (dados de Grimaldi 1990 e mtDNA de DeSalle 1992). Uma árvore que compila todos os dados de seqüências nucleares apresentada por Kwiatowski e Ayala (1999) aponta o subgênero *Drosophila* como parafilético pois outros subgêneros como *Hirtodrosophila* e

Dorsilopha, assim como o gênero *Zaprionus*, são incluídos neste grupo, em todas as análises. Cabe aqui salientar, que os primeiros dados de parentesco em *Drosophila*, a partir de análises realizadas por Throckmorton (1975), concordam com muitos dos pontos evidenciados por estas filogenias. Diante das conclusões destes autores, utilizaremos como padrão a filogenia de Throckmorton (1975) suportada pelos trabalhos com dados moleculares citados, como base para nossas discussões. A Figura 10 apresenta uma árvore hipotética que contém as relações descritas acima, mas que restringe seus táxons àqueles investigados neste trabalho.

Tempos de divergência estimados a partir de vários tipos de dados como fósseis, dados biogeográficos, bem como inferências baseadas em relógio molecular são compilados no trabalho de Powell e DeSalle (1995). Russo et al. (1995) também apresentaram este cálculo para algumas espécies e grupos do gênero *Drosophila* que não são apresentados no trabalho de Powell e DeSalle (1995). Algumas destas estimativas constam na árvore da Figura 10.

As relações filogenéticas menos estudadas são aquelas incluídas por Throckmorton (1975) na radiação *immigrans-Hirtodrosophila*. Das 22 espécies investigadas em nossa análise, 11 fazem parte desta radiação e comparações entre os grupos destas espécies e os agrupamentos formados pelas seqüências homólogas a *gypsy*, retiradas de seus genomas, tornam-se mais difíceis.

Se traçarmos uma comparação entre a árvore da Figura 10 e a filogenia obtida em nossas análises (Figuras 7 e 8), em relação à posição das espécies pertencentes aos 4 subgêneros investigados neste trabalho, (ver

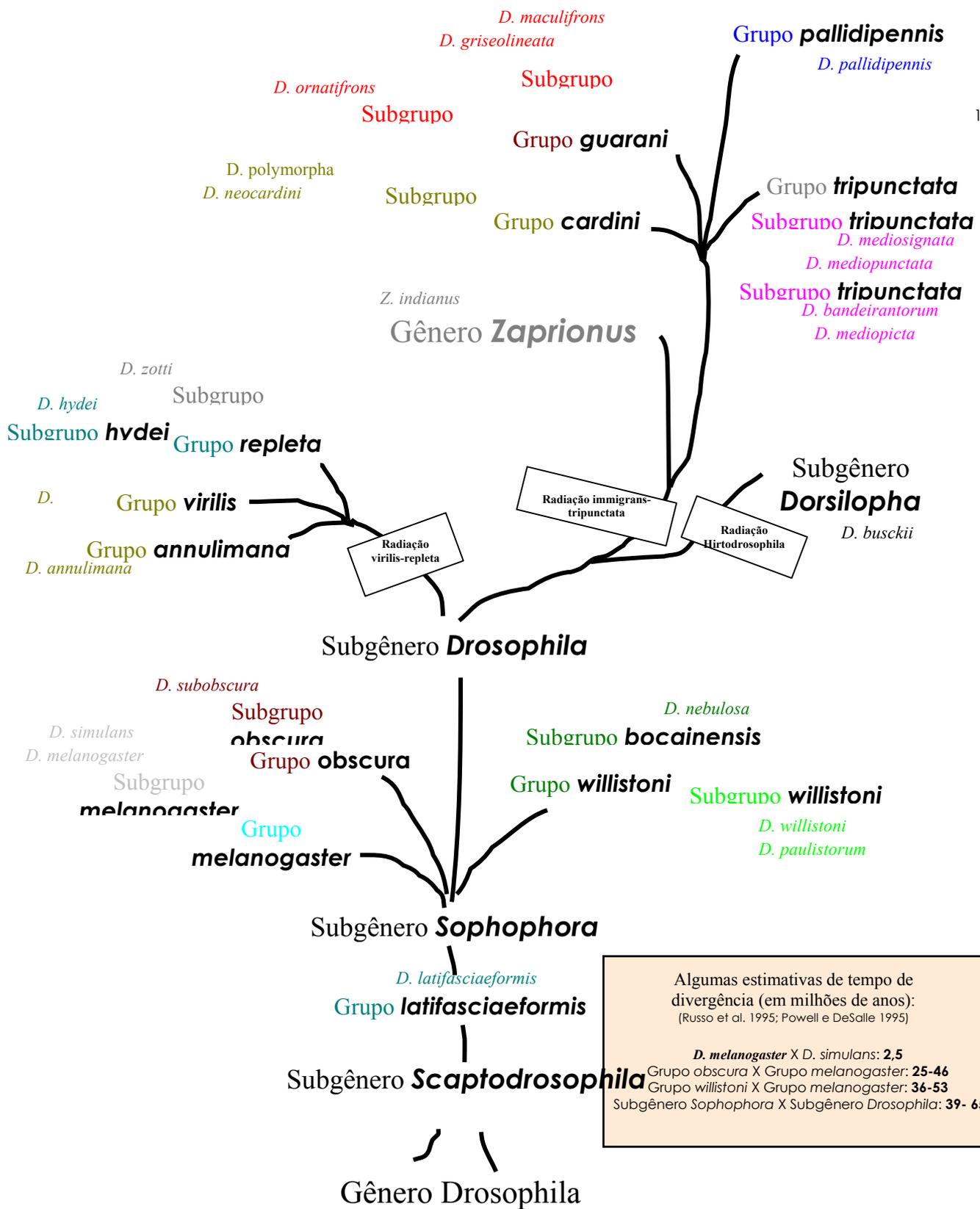


Figura 10. Esquema das relações de parentesco entre as espécies investigadas no presente trabalho (Baseada na filogenia de Throckmorton 1975 e na filogenia consenso de Kwiatowski e Ayala 1999).

símbolos que discriminam os diferentes subgêneros; de acordo com Throckmorton 1975 e Wheeler 1981), veremos que para os arranjos compostos de mais de um tipo de espécie, não ocorreu a separação esperada entre táxons dos diferentes subgêneros. O que podemos observar, na verdade, é a reunião de espécies provenientes de subgêneros distintos. Tal aspecto é bem caracterizado para os agrupamentos A, C e G.

Num segundo momento, traçamos comparações mais detalhadas sobre a posição, nas diferentes árvores, dos grupos de espécies (diferentes grupos são salientados por cores diferentes nas Figuras 7 e 8). Assim, foram feitas as seguintes observações: o agrupamento A reúne *D. subobscura* e *D. busckii*, duas espécies de grupos incluídos em subgêneros diferentes, que diversificaram em radiações muito distantes, a primeira sendo relativamente basal e a segunda sendo derivada do subgênero *Sophophora*. Ainda neste grupo verificamos a presença de espécies relacionadas como *D. hydei* e *D. virilis* que, embora de grupos diferentes, pertencem à mesma radiação *virilis-repleta* dentro do subgênero *Drosophila*. No entanto, este agrupamento não inclui *D. zotti* que é do mesmo grupo de *D. hydei*, o grupo *repleta*. O agrupamento A também inclui *D. bandeirantium* e *D. pallidipennis*, espécies provenientes da radiação *tripunctata*, mas exclui todas as demais espécies de seus grupos de origem. No agrupamento B, embora tenhamos agrupadas seqüências de clones de duas espécies do grupo *tripunctata*, novamente estas encontram-se bastante divergentes de outras espécies deste grupo. O agrupamento C reúne com *D. melanogaster* e *D. simulans*, *D. latifasciaeformis* (subgênero *Scaptodrosophila*), espécie acreditada ser mais primitiva do

gênero *Drosophila*, dentre as analisadas em nosso trabalho. Este agrupamento mostra-se razoavelmente coerente, já que *D. simulans* e *D. melanogaster* são espécies muito relacionadas, pertencentes ao mesmo grupo (grupo *melanogaster*), e cuja origem parece ser o subgênero *Scaptodrosophila*. No entanto, surpreendentemente, seqüências de *Z. indianus*, classificado como um outro gênero derivado da radiação *immigrans*, aparece intimamente relacionado com as espécies representantes do grupo *melanogaster*. Os grupos D e E caracterizam-se por representarem apenas uma espécie altamente divergente, inclusive de outras espécies de seus próprios grupos ou grupos relacionados dentro da mesma radiação. O agrupamento F reúne todas as espécies trabalhadas do grupo *guarani*, mas engloba também espécies do grupo *tripunctata*, *repleta* e *cardini*. Apesar de agrupar apenas espécies do subgênero *Drosophila*, as relações estabelecidas não demonstram coerência com aquelas observadas na filogenia das espécies. *D. zotti* do grupo *repleta* encontra-se agrupada com um clone proveniente de *D. mediopicta* (grupo *tripunctata*) e *D. polymorpha* do grupo *cardini*. Apesar das relações filogenéticas entre as espécies dos grupos da radiação *tripunctata* não serem totalmente compreendidas, a separação da linhagem que origina a radiação *virilis-repleta* é bem estabelecida e separada da radiação *tripunctata*, o que não é verificado nas relações do clado F. O quadro se repete quando observamos o agrupamento G, onde encontram-se associadas *D. neocardini* (grupo *cardini*, radiação *tripunctata* do subgênero *Drosophila*) com espécies do grupo *willistoni*, subgênero *Sophophora*. A análise comparativa empreendida acima ainda chamou-nos à atenção para a extrema divergência observada entre as espécies crípticas *D. willistoni* e *D.*

paulistorum. Aliás, a percentagem de divergência (56%) entre estas duas espécies é muito semelhante aos índices que *D. willistoni* apresenta em relação a praticamente todas as espécies analisadas neste trabalho (51% a 63%), o que leva a linhagem de *gypsy* existente neste genoma a adquirir um *status* muito especial.

O quadro que parece emergir, das comparações realizadas entre as relações de parentesco das espécies e aquelas obtidas das análises das seqüências homólogas a *gypsy*, é a existência de uma série de incompatibilidades entre as duas filogenias. As espécies com as mais diferentes origens misturam-se na árvore produzida, sendo que seqüências de *gypsy* de espécies distantes apresentam, muitas vezes, uma homologia maior do que aquelas oriundas de espécies de mesmo grupo, por exemplo.

Para traçar considerações mais precisas sobre as incongruências acima ressaltadas, procuramos comparar as taxas de divergência obtidas neste trabalho com taxas de divergência retiradas da literatura para genes nucleares. Os dados de referência são dos trabalhos de Pelandakis e Solignac (1993) com rDNA (28S) e de Kwiatowski et al. (1994) e Kwiatowski e Ayala (1999) com o gene Sod (superóxido dismutase). Estas seqüências foram obtidas com a opção "nucleotide" a partir dos trabalhos cujos números de acesso no PubMed (NCBI) são 8283482, 7545938 e 10603260, respectivamente. Calculamos as distâncias conforme os parâmetros utilizados neste trabalho, a fim de traçar as comparações desejadas. As espécies, cujas seqüências estavam disponíveis para esta análise, foram: 1) Sod: *D. melanogaster*, *D. simulans*, *D. willistoni*, *D. subobscura*, *D. hydei*, *D. virilis*, *D. busckii*, *D. paulistorum*, *D. nebulosa*, *Zaprionus*, *Scaptodrosophila* (estes dois últimos grupos foram

representados por espécies distintas da utilizada neste trabalho, *Z. tuberculatus* e *D. lebanonensis*, respectivamente); 2) rDNA (28S): *D. melanogaster*, *D. simulans*, *D. willistoni*, *D. busckii*, *D. nebulosa*, *D. maculifrons*, *D. pallidipennis*, *D. ornatifrons*, *D. latifasciaeformis*, *D. mediopictoides* representando o grupo *tripunctata* e *Z. inermis*, *Z. sepsoides* e *Z. taronus* representando o gênero *Zaprionus*. Os valores comparativos destas últimas três espécies foram representados por uma média. Foram utilizadas médias de taxas de divergência entre clones provenientes de uma mesma espécie que se encontravam agrupados na filogenia produzida a partir de nossos dados. As Tabelas 7 e 8 apresentam as comparações realizadas entre as taxas de divergência das seqüências *gypsy* e de genes nucleares.

Se compararmos os valores obtidos para Sod e rDNA para os mesmos pares de táxons, constataremos que regiões de rDNA mostram-se aproximadamente 3 vezes mais conservadas do que as de Sod. Neste sentido, comparações destes valores entre as seqüências de *gypsy* e de rDNA devem ser consideradas à luz deste conhecimento. Outro ponto a ser ressaltado nesta análise é que taxas de substituição de seqüências retrovirais, como é o caso de *gypsy*, são geralmente maiores em relação aos valores de genes nucleares, devido às altas taxas de má incorporação de nucleotídeos pela transcriptase reversa, quando comparada com a DNA polimerase (Drake 1993). Assim, valores maiores do que aqueles observados para os genes nucleares são esperados, sendo que o grau desta diferença depende da taxa de transposição do elemento no genoma. Quanto maior sua atividade transposicional tanto maior tende a ser a diferença esperada entre as taxas de divergência (Eickbush et al. 1995).

Analisando as Tabelas 7 e 8, observamos que, em um grande número de comparações, os valores para *gypsy* e para genes nucleares são muito semelhantes. Também constatamos, em muitos casos, valores de divergência menores para o elemento em questão. Regiões codificantes são esperadas sofrer forte seleção purificadora em função da sua importância para o organismo. Assim, encontrar uma conservação semelhante ou menor para seqüências de *gypsy* nos é surpreendente. Algumas das comparações mais intrigantes envolvem valores obtidos entre *D. busckii* e *D. subobscura* (0,27 para Sod e 0,05 para *gypsy*), *D. simulans* e *Z. indianus* (0,33 para Sod e 0,08 para rDNA comparada com 0,07 para *gypsy*), *D. subobscura* e *D. hydei* (0,36 para Sod e 0,08 para *gypsy*). Mesmo seqüências de espécies que diversificaram na mesma radiação como *D. virilis* e *D. hydei*, que se encontram agrupadas na filogenia obtida em nossos estudos, apresentam taxas de substituição para *gypsy* muito menores do que aquelas para Sod (0,06 e 0,19, respectivamente). Se tentássemos explicar estes resultados à luz de um padrão evolutivo regido por transmissão vertical, teríamos que pressupor coeficientes de seleção muito altos, muitas vezes maiores do que aqueles que operam sobre uma das regiões mais conservadas do DNA ribossomal (28S). Mesmo assim, não conseguiríamos explicar a relação tão próxima entre seqüências de espécies tão distantes, enquanto seqüências de espécies próximas mostram-se tão divergentes, como é o caso das espécies do grupo *tripunctata*, espalhadas ao longo da filogenia construída por nossas análises.

A maioria dos autores que se deparam com incongruências, do nível das acima mencionadas, reportam que eventos de transmissão horizontal possam reger estas discrepâncias. Gonzales e Lessions (1999) encontraram

Tabela

7

no

anexo.

Tabela

8

no

anexo.

seqüências idênticas de um retroelemento da família *gypsy/Ty3* (*SURL* de ouriço do mar) em espécies distintas. A explicação que os autores postulam para tal conservação entre espécies que sabidamente separaram-se há 3 e 9 milhões de anos atrás, e cujos genes nucleares evoluem a uma taxa de 1,5% por milhão de ano, é a de transmissão horizontal. Os casos documentados de transmissão horizontal para o elemento *mariner* envolvem fortes similaridades (90% - 99% a nível de nucleotídeos) entre táxons pertencentes a ordens e filos distintos (Robertson e MacLeod 1993; Lohe et al. 1995). Ainda em *mariner*, Maruyama e Hartl (1991) detectaram um caso de transmissão horizontal entre *D. mauritiana* e *Zaprionus*, onde a divergência obtida para *mariner* foi de 0,03, índice muito menor do que aquele observado para dados de *Adh* que se encontra em torno de 0,18. Jordan et al. (1999) apontam como evidências fortes para transmissão horizontal do elemento *copia* o fato de ocorrer 99% de identidade nucleotídica, enquanto 80% é observado para genes nucleares (*Adh*), entre *D. willistoni* e *D. simulans*. Os múltiplos casos estabelecidos para o elemento *P* também se baseiam nas comparações acima discutidas (Clark e Kidwell 1997; Silva e Kidwell 2000). Para o elemento *P*, são reportados 8,1% de divergência entre uma espécie do grupo *obscura* e outra do grupo *saltans*, 7,8% entre espécies de gêneros diferentes, além do clássico e pioneiro caso de invasão do genoma de *D. melanogaster* pelo elemento *P* de *D. willistoni* (divergência de 3,1% entre os elementos *P* de espécies que divergiram, há cerca de 50 milhões de anos atrás) (Clark e Kidwell 1997). Mais recentemente, Loreto et al. (2001) documentaram um caso de transmissão horizontal entre espécies de subgêneros distintos, onde os autores reportaram taxas de divergência de apenas 3%.

Muitas vezes, dados de elementos transponíveis são utilizados para construção de filogenias cujos agrupamentos concordam com aqueles observados nas árvores "clássicas" das espécies. No entanto, comparações das divergências obtidas com valores de regiões codificantes não são realizadas. Isso pode subestimar o número de possíveis casos de transmissão horizontal, principalmente aqueles ocorridos entre táxons aparentados (Silva e Kidwell, 2000). Isto ocorre pois muitos agrupamentos esperados são construídos com base em taxas de divergências extremamente baixas, justificadas somente se coeficientes de seleção altíssimos fossem postulados. Forte seleção muitas vezes não condiz com a relação entre mudanças sinônimas e não sinônimas ou com mutações detectadas nas seqüências analisadas, indicando inatividade do elemento.

Além da necessidade de utilização de parâmetros evolutivos das espécies hospedeiras para comparação com aqueles obtidos das seqüências de elementos transponíveis, parece que o número de táxons e, mais especificadamente, de táxons distantemente aparentados, é importante para detectarmos possíveis casos de transmissão horizontal. Isto é bem evidenciado no trabalho de Vásquez-Manrique et al. (2000) onde transmissão vertical é apontada como o modo de evolução de *gypsy* dentro do grupo *obscura*. No entanto, quando seqüências de *gypsy* provenientes de *D. hydei* e *D. virilis* são incluídas na análise, um novo cenário se constrói e transmissão horizontal novamente necessita ser utilizada como explicação para o padrão encontrado.

De acordo com as comparações realizadas, seqüências *gypsy* de espécies do subgênero *Sophophora* parecem ter evoluído acompanhando a

história de suas espécies hospedeiras. Isto quer dizer que, se a análise se restringisse a somente estes táxons, nós não identificaríamos nenhum valor de divergência próximo ou menor àqueles de genes nucleares destas espécies. No entanto, a divergência encontrada entre seqüências *gypsy* do grupo *willistoni* (com exceção de *D. willistoni* que será discutida adiante) e as do grupo *melanogaster* é bastante alta, chegando a representar de 2 a 3 vezes aquelas estimadas para genes nucleares. Estes resultados podem ser explicados por uma alta taxa evolutiva de *gypsy* nestas linhagens (talvez isto seja o esperado levando em consideração seu *status* de retrovírus). Uma hipótese alternativa envolveria a origem destas seqüências antes da diversificação de suas espécies hospedeiras. Uma origem anterior àquela das espécies reporta-nos para o fenômeno de polimorfismo ancestral. Levando em consideração esta hipótese, uma possível explicação para as divergências acima salientadas é que cada grupo adquiriu uma cópia distinta de um "pool" de elementos ancestrais. A mesma linhagem recebida pelo grupo *melanogaster* deve ter sido adquirida por *D. subobscura*, a julgar pela baixa taxa de divergência entre esta espécie e *D. melanogaster*. No entanto, alguns pontos permanecem não resolvidos. Se *D. subobscura* adquiriu a mesma linhagem de *D. melanogaster* porque sua divergência não é equivalente, quando comparações são traçadas com *D. simulans*? Além disso, a divergência entre seqüências *gypsy* de *D. simulans* e *D. melanogaster* também apresenta-se muito maior do que a observada para seus genes nucleares, e teríamos que também supor polimorfismo ancestral antes da divergência destas duas espécies. Diferentes taxas evolutivas, nos diferentes ramos da filogenia do elemento, podem explicar os padrões obtidos. No entanto, não

podemos descartar uma possível transmissão horizontal entre *gypsy* de *D. subobscura* e *D. melanogaster* como o que, provavelmente, ocorreu entre *D. nebulosa* e *D. paulistorum*.

Se *gypsy* foi transmitido de modo vertical no subgênero *Sophophora* e o grau de divergência observado for compatível com altas taxas evolutivas, isto reporta-nos para uma possível atividade destes elementos. Pelas nossas análises, não temos condições de avaliar o potencial para transposição destas seqüências. No entanto, a relação existente entre taxa de substituição sinônima em sítios sinônimos (ds) e taxa de substituição não-sinônima em sítio não sinônimo demonstram, para *D. paulistorum* e *D. nebulosa*, valores sugestivos de neutralidade destas seqüências. O tempo necessário para um acúmulo indiscriminado nos sítios sinônimos e não sinônimos, a ponto de não mais identificarmos qualquer sinal de seleção purificadora, nos é, entretanto, desconhecido.

A análise do agrupamento G demonstra a presença de *D. neocardini* junto às espécies do grupo *willistoni*. A divergência encontrada entre esta espécie e *D. nebulosa* é de 0,09 (distância Kimura). Se observarmos os valores das Tabelas 7 e 8, verificaremos que esta divergência é extremamente baixa, principalmente, se levarmos em consideração que estas são espécies de subgêneros diferentes. Uma seleção tão forte quanto aquela encontrada para regiões de rDNA deveria ser evocada para explicar estes dados. No entanto, assim como *D. nebulosa* e *D. paulistorum*, as mudanças ocorridas nas seqüências *gypsy* de *D. neocardini* mostraram-se neutras no teste realizado. Assim, a homologia encontrada entre estas espécies distantes sugere que

transmissão horizontal deva ter ocorrido num tempo recente, antes da provável inativação destas seqüências de *gypsy*.

O possível padrão de herança vertical detectado para o subgênero *Sophophora* não é tão claro para os outros ramos da filogenia. Um quadro ainda mais complexo desenha-se quando analisamos os dados para outras espécies. As comparações traçadas com os valores de divergência nucleares indicam que as seqüências *gypsy* de muitas espécies provavelmente invadiram recentemente seus genomas. Por outro lado, os dados de *D. maculifrons* e *D. ornatifrons* parecem ser compatíveis com a sua história evolutiva. O mesmo parece ser o observado para algumas espécies do grupo *tripunctata*. Algumas relações não puderam ser checadas por falta de informação quanto à taxas de divergência de genes nucleares, como para *D. polymorpha* e *D. annulimana*.

Segundo critérios revisados por Capy et al. (1998), elementos transponíveis com divergência de mais de 30% são classificados como subfamílias distintas, enquanto entre famílias é esperada uma diferenciação maior, em torno de 50% ou mais. Lohe et al. (1995) estabelecem um limiar maior de diferenciação para elementos pertencentes a uma mesma subfamília de *mariner* (cerca de 50% de divergência nucleotídica) e Clark e Kidwell (1997) classificaram como subfamílias distintas de *P*, elementos divergentes em 40% de seus nucleotídeos. Enquanto isso, Jordan e McDonald (1998) classificaram linhagens de *copia* com divergências entre 42,5% e 58,5% como famílias distintas deste elemento. Blesa et al. (2001) já seguem critérios onde divergências de 38% são consideradas indicativas de famílias diferentes do elemento *bilbo*. As distâncias obtidas entre os agrupamentos estabelecidos

(A a G) por nossos dados, nos permite inferir que múltiplas subfamílias de *gypsy* existem no genoma das diferentes espécies investigadas. Pinsker et al. (2001) chamam à atenção para o fato de que devemos ter cautela ao estabelecer subfamílias e famílias de elementos transponíveis quando somente uma região do elemento é investigada. No entanto, Alberola e De Frutos (1996) analisaram diferentes regiões de *gypsy* completo de *D. melanogaster*, *D. subobscura* e *D. virilis* e detectaram percentagens de divergência muito semelhantes ao longo das três ORFs neste retroelemento. Este dado é interessante, no sentido de validar nossas estimativas de divergência de uma região de *gypsy* para o elemento como um todo. Assim, nossos dados nos permitem inferir a existência de diferentes subfamílias de *gypsy* nas espécies investigadas.

Os agrupamentos de A a G podem representar diferentes subfamílias à medida em que suas percentagens de divergência variam entre 23% e 40,8% (Tabela 5). Segundo os critérios de Capy et al. (1998) o agrupamento F estaria no limiar entre representar uma subfamília ou uma família distinta de *gypsy* (divergências deste grupo com os demais variam entre 30% e 36,8%), enquanto o agrupamento G representaria uma família distinta de *gypsy*, apresentando de 34,9 a 44,3% de diferença nucleotídica em relação aos demais grupos. No entanto, como ressaltado anteriormente, estes critérios variam entre autores e nossa classificação encontra-se dentro desta margem de variação. Por outro lado, a distância encontrada entre o agrupamento H e os demais varia entre 52,3 e 62,1 e representa uma divergência muito grande para o enquadramento deste grupo como uma subfamília distinta de *gypsy*. Neste caso, preferimos classificar os elementos *gypsy* encontrados no genoma de *D. willistoni* como uma família distinta de elementos "like-*gypsy*". Análises

mais detalhadas de outras regiões deste elemento fazem-se necessárias para avaliar uma possível homologia desta seqüência com outras famílias de elementos transponíveis aparentadas com *gypsy*, já descritas na literatura.

Uma pesquisa TBLASTN no GeneBank (NCBI) por seqüências homólogas à região seqüenciada em nosso trabalho, para *D. willistoni*, foi realizada e os resultados apontaram para uma falta de homologia significativa (mais do que 30 pares de nucleotídeos seguidos) desta seqüência com qualquer outra no banco de dados usado. A grande divergência de seqüências amplificadas do genoma de *D. willistoni* torna-se ainda mais interessante à luz da conservação desta região entre populações distantes. Tal constatação contradiz uma alta taxa evolutiva desta linhagem de *gypsy* neste genoma em particular. O que parece mais plausível é uma origem distante, juntamente com perdas deste tipo de seqüência em todas as outras linhagens que levam às demais espécies investigadas, inclusive em espécies tão próximas dela quanto *D. paulistorum*. Cabe lembrar que *D. willistoni* e *D. paulistorum* são espécies crípticas e que alguns trabalhos já detectaram a produção de híbridos interespecíficos em laboratório (revisões em Cordeiro e Winge 1995; Valiati 1999).

Nossos dados apontam para uma possível transmissão vertical de seqüências *gypsy* entre as espécies do subgênero *Sophophora* com uma única exceção listada a seguir, incluindo neste ramo uma linhagem divergente em *D. willistoni*, provavelmente oriunda de um polimorfismo ancestral muito antigo. Para explicar as demais relações estabelecidas na árvore, vários eventos de transmissão horizontal são propostos (Figura 11). Aqueles mais marcantes envolvem seqüências homólogas a *gypsy* de: 1) *D.*

Figura

11

no

anexo.

nebulosa e *D. neocardini*; 2) *D. nebulosa* e *D. paulistorum*; 3) *D. subobscura* e *D. busckii*; 4) *D. subobscura* e *D. hydei*; 5) *D. simulans* e *D. Zaprionus*; 6) *D. simulans* e *D. latifasciaeformis*; 7) *D. pallidipennis* e *D. bandeirantorum*; 8) *D. virilis* e *D. hydei*; 9) *D. mediopicta* e *D. zotti*. Todas comparações entre dados de divergência de *gypsy* e de genes nucleares para estes 9 casos são completamente inconsistentes com transmissão vertical. Os valores mostraram-se sempre menores para *gypsy*. Além disso, em muitas destas seqüências, o teste de seleção mostrou-se negativo, o que descarta a possibilidade de seleção muito forte produzindo esta grande conservação. Este é o caso para seqüências de *D. neocardini* e *D. nebulosa*. As diferenças encontradas nos sítios sinônimos e não-sinônimos entre estas duas espécies não apontaram para qualquer tendência esperada por força de seleção purificadora. O mesmo verificamos para mudanças ocorridas em *Z. indianus*, após a invasão do genoma desta espécie por *gypsy*, provavelmente vindo de *D. simulans*. Neutralidade também é observada para *D. latifasciaeformis*, *D. virilis* e *D. hydei*.

O quadro geral que se desenha é a existência de, pelo menos, duas linhagens de seqüências *gypsy* na origem do gênero *Drosophila*. Parece ter ocorrido a manutenção destas linhagens no genoma de espécies do subgênero *Sophophora*. Além disso, nossos dados corroboram para uma invasão recente dos genomas da maioria das espécies dos demais subgêneros, pelas seqüências *gypsy* provenientes das espécies do subgênero *Sophophora*.

Usando a taxa de substituição de 1,6% por milhão de anos estimada para genes nucleares com baixo uso preferencial de códons (Sharp e Li 1989), nós estimamos o tempo de divergência de seqüências de *gypsy* envolvidas nos nove casos de transmissão horizontal proposto neste trabalho. Esta abordagem já foi implementada por outros autores (Alberola e De Frutos 1993a; Silva e Kidwell 2000; Loreto et al. 2001) e baseia-se em estudos de outros elementos que se transpõem via um intermediário de RNA, tais como *R1* e *R2* de *Drosophila* e *SURL* de ouriço do mar, que apresentaram taxas evolutivas semelhantes às aquelas observadas para genes nucleares com baixo uso preferencial de códons (Shields et al. 1988; Springer et al. 1991; Eickbush et al. 1995).

Estimativas de tempo de divergência encontram-se na Figura 11, junto às setas que indicam os possíveis casos de transmissão horizontal. Estas estimativas dão-nos uma boa idéia do tempo em que provavelmente ocorreu a invasão dos genomas por *gypsy*. A invasão mais recente data de 1,25 MYA, o que é considerado muito recente em termos de tempo evolutivo. As estimativas mais antigas ficam em torno de 20 MYA. Se os elementos *gypsy* detectados no genoma das espécies investigadas têm se comportado como um retrovírus ativo, como mostra-nos o cenário da Figura 11, as taxas evolutivas calculadas podem superestimar os tempos de divergência, já que elementos retrotransponíveis ativos tendem a apresentar taxas superiores às aquelas de genes nucleares (Eickbush et al. 1995). Assim, pode ser que os eventos de transmissão horizontal sejam ainda mais recentes do que o estimado.

A invasão recente dos genomas de *Drosophila* por *gypsy* parece ser corroborada pelos valores obtidos de taxa de substituição não sinônima por

sítio não sinônimo (dn) sobre taxa de substituição sinônima por sítio sinônimo (ds) (dn/ds). O valor médio obtido para todas as comparações entre os 40 táxons utilizados na nossa análise foi 0,29 (variando entre 0,03 e 2,5). Este valor significa que, em média, estão ocorrendo cerca de 3,5 vezes mais substituições sinônimas em relação àquelas não sinônimas. Estes valores estão bem próximos daqueles de genes nucleares (Akashi, 1994). Além disso, os testes de seleção realizados demonstraram que, para a maioria das comparações, seleção purificadora está operando ou operou recentemente sobre as seqüências de *gypsy* (para um nível de significância de 0,05). Gonzales e Lessions (1999) encontraram valores médios muito semelhantes aos nossos para os retrotransposons da família *SURL* e justificaram os seus dados como indicativo de algum grau de domesticação ou transposição recente. Esta última alternativa parece adaptar-se mais aos nossos dados. Terzian et al. (2000) investigou *gypsy* em espécies do subgrupo *melanogaster* e também identificou seleção purificadora, relacionando tal fato com eventos de transposição recentes. Traçando uma análise comparativa entre *D. melanogaster*, *D. virilis* e *D. subobscura*, Alberola e De Frutos (1996) também apontaram valores muito baixos de dn/ds, indicativos de seleção operante sobre estas seqüências. Nossos dados de dn e ds são comparáveis àqueles observados para *Adh* (Silva e Kidwell 2000), o que nos permite inferir que estas seqüências estão sofrendo ou devem ter sofrido forte seleção num passado recente.

Um outro ponto interessante a ser discutido é a existência de seqüências muito divergentes no genoma de uma mesma espécie. Este fenômeno pode ser observado em *D. mediopicta* e *D. mediopunctata*. Para *gypsy*, já foi

reportada a co-existência de mais de um tipo de subfamília no genoma de *D. subobscura* (Alberola e De Frutos 1996), *D. erecta* (Terzian et al. 2000) e *D. hydei* (Hochstenbach et al. 1996; Vázquez-Manrique et al. 2000). O polimorfismo encontrado nesta última espécie não foi detectado em nossas análises. Isto pode ter ocorrido por um erro amostral (talvez, se analisássemos um maior número de clones, esta divergência seria detectada) ou, ainda, porque os iniciadores que utilizamos não foram capazes de amplificar a região referida. Assim, a existência de diferentes subfamílias no mesmo genoma pode ser mais freqüente do que nossos resultados demonstram. Diferentes subfamílias de elementos do grupo *gypsy* são sugeridas como existindo, também, em táxons ancestrais, o que explicaria o padrão evolutivo deste grupo de TEs nos mais diferentes ramos dos seres vivos (Marin e Lloréns 2000). O fenômeno de múltiplas linhagens em uma mesma espécie é observado para vários elementos transponíveis como *mariner* (Robertson e MacLeod 1993), *P* (Clark e Kidwell 1997), *Minos* (Arcà e Savakis 2000), *R1* (Gentile et al. 2001), entre outros.

A co-existência de diferentes subfamílias de um dado TE no genoma hospedeiro é comentada por alguns autores como indício de invasões independentes de um genoma pelo elemento transponível, provavelmente por via horizontal (Zelentsova et al. 1999; Silva e Kidwell 2000; Arcà e Savakis 2000). No entanto, polimorfismo ancestral não pode ser descartado como explicação, desde que a divergência da seqüência das diferentes subfamílias seja compatível com as divergências esperadas em comparações com seqüências obtidas de outras espécies. Este parece ser o caso para as seqüências divergentes de *gypsy* encontradas no genoma de *D.*

mediopunctata. A distância genética calculada entre a seqüência de *gypsy* do clone ctata45 e as seqüências de *D. maculifrons* é de 0,09 enquanto distâncias entre regiões de rDNA de *D. maculifrons* e *D. mediopictoides* (uma espécie do mesmo grupo de *D. mediopunctata*) apresentam-se em torno de 0,02. Levando em consideração a grande conservação desta região de rDNA, estes índices podem ser considerados compatíveis com a hipótese de transmissão vertical da seqüência em questão ao longo da divergência destas espécies. Estas considerações apontam para um polimorfismo ancestral antes da divergência de *D. mediopunctata* e *D. maculifrons*.

Uma análise conjunta de seqüências homólogas a elementos da família *gypsy/Ty3*, por Marin e Lloréns (2000), a partir de vários genomas seqüenciados, permitiram aos autores inferir a existência de aproximadamente 20 linhagens distintas no genoma de *Drosophila melanogaster*. Seqüências homólogas a *gypsy* foram investigadas nos diferentes filos de seres vivos e os agrupamentos formados desta análise demonstraram que elementos "gypsy-like" de vertebrados são mais relacionados com aqueles de fungos e plantas do que com elementos de insetos, sugerindo que, pelo menos, um evento de transmissão horizontal possa ter ocorrido ao longo da divergência dos filos ou, ainda, que polimorfismo ancestral deva ter existido (Miller et al. 1999). Portanto, estes fenômenos parecem ser comuns ao longo da história evolutiva de *gypsy*.

A hipótese de transmissão vertical, apoiada como a provável explicação para a existência de linhagens divergentes no genoma de *D. mediopunctata*, parece não se aplicar para as duas subfamílias de *gypsy* detectadas em *D. mediopicta*. Uma homologia de 99.5% entre a seqüência do clone picta34 com seqüências *gypsy* de *D. zotti* (uma espécie de outra

radiação) não parecem refletir ancestralidade comum. Embora não tenhamos dados de regiões gênicas comparáveis, taxas de divergência a este nível são observadas entre espécies do mesmo grupo como *D. ornatifrons* e *D. maculifrons* para regiões de rDNA ou entre *D. paulistorum* e *D. willistoni* (espécies crípticas) para Sod. Forte efeito seletivo operando sobre estas regiões não é observado no teste de seleção implementado que apontou neutralidade na comparação feita entre seqüências *gypsy* de *D. zotti* e do clone *picta*³⁴. O que nos parece mais coerente, em relação a estes dados, é que tenha ocorrido um evento de transmissão horizontal desta seqüência *gypsy* entre o genoma de *D. zotti* e de *D. mediopicta*.

Vários autores têm proposto que elementos transponíveis mantêm-se nos genomas hospedeiros principalmente por invasão via transmissão horizontal. Sua manutenção vertical levaria, incondicionalmente, à inativação e conseqüentemente a uma degradação por acúmulo de mutações (Pinsker et al. 2001). Silva e Kidwell (2000) ressaltaram que, no caso do elemento *P*, a transmissão horizontal seria o principal mecanismo, e talvez a única fonte de pressão seletiva sobre a transposase deste elemento. Neste sentido, somente TEs capazes de transposição sobreviveriam nos genomas das espécies, através de sucessivas ondas de invasão. Hagemann et al. (1996) são incisivos ao concluir sobre os vários casos de transmissão horizontal do elemento *mariner*. Os autores discutem que este elemento tenderia a tornar-se inativo e perdido no genoma e que a única saída para sua manutenção seria o mecanismo de transmissão horizontal. Jordan e McDonald (1998) também chamam à atenção para que os múltiplos e recentes eventos de transmissão horizontal na

história evolutiva de *copia* sejam os fenômenos responsáveis pela sua sobrevivência em uma ampla distribuição de espécies.

Casos convincentes de transmissão horizontal têm sido limitados, até há bem pouco tempo, a elementos de Classe II como *P. mariner* e *hobo* (Daniels et al. 1990; Simmons 1992; Clark et al. 1994; Lohe et al. 1995). Entretanto, evidências sobre a importância deste fenômeno para os retrotransposons vêm se acumulando nos últimos anos. Tal constatação é verdadeira para elementos da família *Ty1-copia*, *412*, *SURL* e *gypsy* (Alberola e De Frutos 1996; Flavell et al. 1997; Cizeron et al. 1998; Jordan e McDonald 1998; Gonzales e Lession 1999; Jordan et al. 1999; Terzian et al. 2000; Vázquez-Manrique et al. 2000).

Dados obtidos, a partir de análises por *Southern blot* investigando a presença de seqüências homólogas a *gypsy* nos mais diferentes genomas das espécies de *Drosophila*, já vinham apontando, desde os trabalhos de Stacey et al. (1986), para uma descontinuidade entre os sinais de hibridização com a sonda oriunda do genoma de *D. melanogaster* e as relações de parentesco entre as espécies (Loreto et al. 1998a). Nossos resultados, através desta metodologia, demonstraram uma considerável divergência das espécies investigadas quanto aos sítios de restrição existentes nos genomas de *D. melanogaster*, *D. virilis* e *D. subobscura*. No entanto, esta abordagem nos permitiu detectar um possível compartilhamento de fragmentos produzidos pela enzima *HindIII* entre *D. busckii* e *D. subobscura*. A similaridade encontrada nestes ensaios foi confirmada pela análise das seqüências homólogas a *gypsy* que agrupou estas duas espécies com uma percentagem de divergência de apenas 5% para a região investigada. Assim, indícios de um dos casos de

transmissão horizontal proposto neste trabalho pode ser detectado por *Southern blot*. O mesmo pode ser comentado para a conservação do sítio de restrição em *Z. indianus*, para a enzima *Bgl*III. Mediante a divergência geral obtida, esta similaridade torna-se alvo de interesse, já que as espécies em questão (*D. melanogaster* e *Z. indianus*) pertencem a gêneros distintos. Novamente, a relação filogenética próxima entre as seqüências *gypsy* destas duas espécies se fez presente na construção da árvore obtida, a partir dos dados de seqüência, e outro evento de transmissão horizontal pode ser inferido envolvendo o grupo *melanogaster* e *Z. indianus*.

Os fortes produtos de amplificação obtidos a partir da utilização de oligonucleotídeos, cuja construção foi baseada em seqüências de *D. melanogaster*, *D. virilis* e *D. subobscura*, para algumas espécies filogeneticamente distantes daquelas tomadas como referência, são compreendidos quando tomamos como fonte de comparação a filogenia obtida a partir das seqüências de *gypsy*. Assim, tanto os nossos resultados de *Southern blot* (quando utilizamos o elemento inteiro) como aqueles obtidos dos produtos amplificados por PCR, indicaram inconsistências nas similaridades entre os elementos e a filogenia de suas espécies hospedeiras e foram confirmados pela análise filogenética baseadas nas seqüências de *gypsy*.

A ocorrência de transmissão horizontal, entretanto, requer alguns pressupostos que envolvem sobreposição geográfica e ecológica entre as entidades envolvidas, além de um mecanismo ou vetor de transferência. Sobre a primeira premissa, dos nove casos de transmissão horizontal inferidos a partir de nossos resultados, 5 deles envolvem pelo menos uma espécie cosmopolita: *D. buskii*, *D. virilis*, *D. hydei*, *D. latifasciaeformis*, *D. simulans*. Os

demais casos envolvem espécies que diversificaram em regiões neotropicais e mantém, desde então, simpatria. Uma grande sobreposição de substratos para alimentação e ovoposição preferenciais destas espécies é bastante conhecida para o gênero *Drosophila* (revisões em Carson 1951 e 1965; Throckmorton 1975)

A problemática em torno do vetor responsável pela transferência horizontal vem sendo discutida por vários autores e envolve a possível atuação de vírus, ácaros parasíticos bem como bactérias simbiotes (Friesen e Nissen 1990; Houck et al. 1991; Kidwell 1993). O que faz com que o cenário evolutivo do retrotransposon *gypsy* seja tão fascinante é o fato de que os múltiplos eventos de transmissão horizontal aqui propostos não requerem qualquer tipo de vetor. Se *gypsy* é capaz de codificar a proteína do envelope e comportar-se como um retrovírus, bastaria que ondas infectivas ocorressem para explicar o padrão observado. Se este aspecto do seu ciclo de vida não é hoje identificado por alterações na ORF3, o mesmo não pode ser dito para um passado talvez mais recente do que possamos imaginar.

REFERÊNCIAS BIBLIOGRÁFICAS

- AKASHI, H. 1994. Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. *Genetics* **136**:927-935.
- ALBEROLA, T. M., and R. De FRUTOS. 1993a. Gypsy homologous sequences in *Drosophila subobscura* (gypsyDS). *J. Mol. Evol.* **36**:127-135.
- ALBEROLA, T. M., and R. De FRUTOS. 1993b. Distribution of gypsy sequences in *Drosophila* species of the obscura subgroup. *Hereditas* **118**:131-137.
- ALBEROLA, T. M., and R. De FRUTOS. 1996. Molecular structure of a gypsy element of *Drosophila subobscura* (gypsyDS) constituting a degenerate form of insect retroviruses. *Nucleic Acids Res.* **24**:914-923.
- ALBEROLA, T. M., L. BORI, and R. De FRUTOS. 1997. Structural analysis of *Drosophila subobscura* gypsy elements (gypsyDs). *Genetica* **100**:39-48.
- ARCÀ, B., and C. SAVAKIS. 2000. Distribution of the transposable element *Minos* in the genus *Drosophila*. *Genetica* **108**:263-267.
- AVEDISOV, S. N., and Y. V. ILYIN. 1994. Identification of spliced RNA species of *Drosophila melanogaster* gypsy retrotransposon: new evidence for retroviral nature of the gypsy element. *FEBS Lett.* **350**:147-150.
- BAYEV Jr., A. A., N. V. LYUBOMIRSKAYA, E. B. DZHUMAGALIEV, E. V. ANANIEV, I. G. AMIANTOVA, and Y. V. ILYIN. 1984. Structural organization of transposable element mdg4 from *Drosophila melanogaster* and a nucleotide sequence of its long terminal repeats. *Nucleic Acids Res.* **12**:3707-3723.
- BIÉMONT, C., and G. CIZERON. 1999. Distribution of transposable elements in *Drosophila* species. *Genetica* **105**:43-62.

- BLESA, D., M. GANDÍA, and M. J. MARTÍNEZ-SEBASTIÁN. 2001. Distribution of the *bilbo* non-LTR retrotransposon in Drosophilidae and its evolution in the *Drosophila obscura* species group. *Mol. Biol. Evol.* **18**:585-592.
- BUCHETON, A. 1995. The relationship between the *flamenco* gene and *gypsy* in *Drosophila*: how to tame a retrovirus. *Trends Genet.* **11**:349-353.
- CAPY, P., C. BAZIN, D. HIGUET, and T. LANGIN. 1998. Dynamics and evolution of transposable elements. Landes Bioscience, Austin, Tex.
- CARSON, H. L., and H. D. STALKER. 1951. Natural breeding sites for some wild species of *Drosophila* in the eastern United States. *Ecology* **32**:317-330.
- CARSON, H. L. 1965. Chromosomal morphism in geographically widespread species of *Drosophila*. Pp. 503-531 in *The genetics of colonizing species*, Academic Press, New York.
- CHALVET, F., C. DI FRANCO, A. TERRINONI, A. PELISSON, N. JUNAKOVIC, and A. BUCHETON. 1998. Potentially active copies of the *gypsy* retroelement are confined to the Y chromosome in some strains of *Drosophila melanogaster* possibly as the result of the female-specific effect of the *flamenco* gene. *J. Molec. Evol.* **46**:437-441.
- CIZERON, G., F. LEMEUNIER, C. LOEVENBRUCK, A. BREHM, and C. BIÉMONT. 1998. Distribution of the retrotransposable element *412* in *Drosophila* species. *Mol. Biol. Evol.* **15**:1589-1599.
- CLARK, J. B., W. P. MADDISON, and M. G. KIDWELL. 1994. Phylogenetic analysis supports horizontal transfer of *P* transposable elements. *Mol. Biol. Evol.* **11**:40-50.
- CLARK, J. B., and M. G. KIDWELL. 1997. A phylogenetic perspective on *P* transposable element evolution in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **94**:11428-11433.
- CORDEIRO, A. R., and H. WINGE. 1995. Levels of evolutionary divergence of *Drosophila willistoni* sibling species. In: *Genetics of natural populations: The*

continuing importance of T. Dobzhansky, L. Levine (ed.), Columbia University Press, New York.

DANIELS, S. B., A. CHOVNICK, and I. A. BOUSSY. 1990. Distribution of *hobo* transposable elements in the genus *Drosophila*. *Mol. Biol. Evol.* **7**:589-606.

De FRUTOS, R., K. R. PETERSON, and M. G. KIDWELL. 1992. Distribution of *Drosophila melanogaster* transposable element sequences in species of the *obscura* group. *Chromosoma* **101**:293-300.

DeSALLE, R. 1992. The phylogenetic relationships of flies in the Family Drosophilidae deduced from mtDNA sequences. *Mol. Phylogenet. Evol.* **1**:31-40.

DORSETT, D., G. A. VIGLIANTI, B. J. RUTLEDGE, and M. MEELSON. 1989. Alteration of *hsp82* gene expression by the *gypsy* transposon and suppressor genes in *Drosophila melanogaster*. *Genes Dev.* **3**:454-468.

DOWSETT, A. P., and M. W. YOUNG. 1982. Differing levels of dispersed repetitive DNA among closely related species of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **79**:4570-4574.

DRAKE, J. W. 1993. Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. USA* **90**:4171-4175.

EICKBUSH, D. G., W. C. LATHE III, M. P. FRANCINO, and T. H. EICKBUSH. 1995. *R1* and *R2* retrotransposable elements of *Drosophila* evolve at rates similar to those of nuclear genes. *Genetics* **139**:685-695.

FLAVELL, A. J., S. R. PEARCE, (J.S.) P. HESLOP-HARRISON, and A. KUMAR. 1997. The evolution of *Ty1-copia* group retrotransposons in eukaryote genomes. *Genetica* **100**:185-195.

FRIESEN, P. D., and M. S. NISSEN. 1990. Gene organization and transcription of TED, a lepidopteran retrotransposon integrated within the baculovirus genome. *Mol. Cell. Biol.* **10**:3067-3077.

- GENTILE, K. L., W. D. BURKE, and T. H. EICKBUSH. 2001. Multiple lineages of R1 retrotransposable elements can coexist in the rDNA loci of *Drosophila*. *Mol. Biol. Evol.* **18**:235-245.
- GEORGIEV, G. P., Y. V. ILYIN, V. G. CHMELIAUSKAITE, A. P. RYSKOV, D. A. KRAMEROV, K. G. SKRYABIN, A. S. KRAYEV, E. M. LUKANIDIN, and M. S. GRIGORYAN. 1981. Mobile dispersed genetic elements and other middle repetitive DNA sequences in the genomes of *Drosophila* and mouse: transcription and biological significance. *Cold Spring Harbor Symp. Quant. Biol.* **45**:641-654.
- GONZALEZ, P., and H. A. LESSIOS. 1999. Evolution of sea urchin retroviral-like (SURL) elements: evidence from 40 echinoid species. *Mol. Biol. Evol.* **16**:938-952.
- GRIMALDI, D. A. 1990. A phylogenetic, revised classification of genera in the Drosophilidae (Diptera). *Bull. Am. Mus. Nat. Hist.* **197**:1-139.
- HAGEMANN, S., E. HARING, and W. PINSKER. 1996. Repeated horizontal transfer of *P* transposons between *Scaptomyza pallida* and *Drosophila bifasciata*. *Genetica* **98**:43-51.
- HIGGINS, D. G., A. J. BLEASBY, and R. FUCHS. 1992. CLUSTAL V: improved software for multiple sequence alignment. *Computer Applications in the Biosciences (CABIOS)*. **8**:189-191.
- HOCHSTENBACH, R., H. HARHANGI, K. SCHOUREN, P. BINDELS, R. SUIJKERBUIJK, and W. HENNIG. 1996. Transcription of gypsy elements in a Y-chromosome male fertility gene of *D. hydei*. *Genetics* **142**:437-446.
- HOUCK, M. A., J. B. CLARK, K. R. PETERSON, and M. G. KIDWELL. 1991. Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. *Science* **253**:1125-1128.

- JORDAN, I. K., and J. F. McDONALD. 1998. Evolution of the *copia* retrotransposon in the *Drosophila melanogaster* species subgroup. *Mol. Biol. Evol.* **15**:1160-1171.
- JORDAN, I. K., L. V. MATYUNINA, and J. F. McDONALD. 1999. Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. *Proc. Natl. Acad. Sci. USA* **96**:12621-12625.
- KIDWELL, M. G. 1993. Lateral transfer in natural populations of eukaryotes. *Annu. Rev. Genet.* **27**:235-256.
- KIM, A., C. TERZIAN, P. SANTAMARIA, A. PÉLISSON, N. PRUD'HOMME, and A. BUCHETON. 1994. Retroviruses in invertebrates: the *gypsy* retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **91**:1285-1289.
- KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111-120.
- KUMAR, S., K. TAMURA, I. B. JAKOBSEN, and M. NEI. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*.
- KWIATOWSKI, J., D. SKARECKY, K. BAILEY, and F. J. AYALA. 1994. Phylogeny of *Drosophila* and related genera inferred from the nucleotide sequence of the Cu, Zn Sod gene. *J. Mol. Evol.* **38**:443-454.
- KWIATOWSKI, J., and F. J. AYALA. 1999. Phylogeny of *Drosophila* and related genera: conflict between molecular and anatomical analyses. *Mol. Phylogenet. Evol.* **13**:319-328.
- LAMBERTSSON, A., S. ANDERSSON, and T. JOHANSSON. 1989. Cloning and characterization of variable-sized *gypsy* mobile elements in *Drosophila melanogaster*. *Plasmid* **22**:22-31.
- LERAT, E., and P. CAPY. 1999. Retrotransposons and retroviruses: analysis of the envelope gene. *Mol. Biol. Evol.* **16**:1198-1207.

- LOHE, A. R., E. N. MORIYAMA, D. LIDHOLM, and D. L. HARTL. 1995. Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. *Mol. Biol. Evol.* **12**:62-72.
- LORETO, E. L. Da S., L. B. Da SILVA, A. ZAHA, and V. L. Da S. VALENTE. 1998a. Distribution of transposable elements in neotropical species of *Drosophila*. *Genetica* **101**:153-165.
- LORETO, E. L. Da S., A. ZAHA, and V. L. Da S. VALENTE. 1998b. Transposable elements in South American populations of *Drosophila simulans*. *Genet. Sel. Evol.* **30**:171-180.
- LORETO, E. L. Da S., V. L. Da S. VALENTE, A. ZAHA, J. C. SILVA, and M. G. KIDWELL. 2001. *Drosophila mediopunctata* P elements: a new example of horizontal transfer. *J. Heredity* **92**:375-381.
- MARÍN, I., and C. LLORÉNS. 2000. *Ty31Gypsy* retrotransposons: description of new *Arabidopsis thaliana* elements and evolutionary perspectives derived from comparative genomic data. *Mol. Biol. Evol.* **17**:1040-1049.
- MARQUES, E. K., M. NAPP, H. WINGE, and A. R. CORDEIRO. 1966. A corn meal, soybean flour, wheat germ medium for *Drosophila*. *D. I. S.* **41**:187.
- MARUYAMA, K., and D. L. HARTL. 1991. Evidence for interspecific transfer of the transposable element *mariner* between *Drosophila* and *Zaprionus*. *J. Mol. Evol.* **33**:514-524.
- MILLER, K., C. LYNCH, J. MARTIN, E. HERNIOU, and M. TRISTEM. 1999. Identification of multiple *gypsy* LTR-retrotransposon lineages in vertebrate genomes. *J. Mol. Evol.* **49**:358-366.
- MIZROKHI, L. J., and A. M. MAZO. 1991. Cloning and analysis of the mobile element *gypsy* from *D. virilis*. *Nucleic Acids Res.* **19**: 913-916.
- MONTCHAMP-MOREAU, C., S. RONSSERAY, M. JACQUES, M. LEHMANN, and D. ANXOLABÉHÈRE. 1993. Distribution and conservation of sequences

- homologous to the 1731 retrotransposon in *Drosophila*. *Mol. Biol. Evol.* **10**:791-803.
- NEI, M., and T. GOJOBORI. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**:418-426.
- NICHOLAS, K. B., and H. B. Jr. NICHOLAS. 1997. GeneDoc: A tool for editing and annotating multiple sequences alignments. WWW.psc.edu/biomed/genedoc.
- PELANDAKIS, M., and M. SOLIGNAC. 1993. Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. *J. Mol. Evol.* **37**:525-543.
- PÉLISSON, A., S. U. SONG, N. PRUD'HOMME, P. A. SMITH, A. BUCHETON, and V. G. CORCES. 1994. Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the *Drosophila flamenco* gene. *EMBO J.* **13**:4401-4411.
- PÉLISSON, A., L. TEYSSET, F. CHALVET, A. KIM, N. PRUD'HOMME, C. TERZIAN, and A. BUCHETON. 1997. About the origin of retroviruses and the co-evolution of the gypsy retrovirus with the *Drosophila flamenco* host gene. *Genetica* **100**:29-37.
- PINSKER, W., E. HARING, S. HAGEMANN, and W. J. MILLER. 2001. The evolutionary life history of *P* transposons: from horizontal invaders to domesticated neogenes. *Chromosoma* **110**:148-158.
- POWELL, J. R., and R. DeSALLE. 1995. *Drosophila* molecular phylogenies and their uses. *Evol. Biol.* **28**:87-138.
- PRUD'HOMME, N., M. GANS, M. MASSON, C. TERZIAN, and A. BUCHETON. 1995. *Flamenco*, a gene controlling the gypsy retrovirus of *Drosophila melanogaster*. *Genetics* **139**:697-711.

- ROBERTSON, H. M., and E. G. MacLEOD. 1993. Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. *Insect Mol. Biol.* **2**:125-139.
- RUSSO, C. A. M., N. TAKEZAKI, and M. NEI. 1995. Molecular phylogeny and divergence times of *Drosophilid* species. *Mol. Biol. Evol.* **12**:391-404.
- SASSI, A. 2000. Contribuição ao estudo dos elementos transponíveis em populações de *Drosophila willistoni*. Dissertação de Bacharelado, UFRGS, Brasil.
- SHARP, P. M., and W.-H. LI. 1989. On the rate of DNA sequence evolution in *Drosophila*. *J. Mol. Evol.* **28**:398-402.
- SHIELDS, D. C., P. M. SHARP, D. G. HIGGINS, and F. WRIGHT. 1988. Silent sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. *Mol. Biol. Evol.* **5**:704-716.
- SILVA, J. C., and M. G. KIDWELL. 2000. Horizontal transfer and selection in the evolution of *P* elements. *Mol. Biol. Evol.* **17**:1542-1557.
- SIMMONS, G. M. 1992. Horizontal transfer of hobo transposable elements within the *Drosophila melanogaster* species complex: evidence from DNA sequencing. *Mol. Biol. Evol.* **9**:1050-1060.
- SONG, S. U., T. GERASIMOVA, M. KURKULOS, J. D. BOEKE, and V. G. CORCES. 1994. An *env*-like protein encoded by a *Drosophila* retroelement: evidence that gypsy is an infectious retrovirus. *Genes Dev.* **8**:2046-2057.
- SPRINGER, M. S., E. H. DAVIDSON, and R. J. BRITTEN. 1991. Retroviral-like element in a marine invertebrate. *Proc. Natl. Acad. Sci. USA* **88**:8401-8404.
- STACEY, S. N., R. A. LANSMAN, H. W. BROCK, and T. A. GRIGLIATTI. 1986. Distribution and conservation of mobile elements in the genus *Drosophila*. *Mol. Biol. Evol.* **3**:522-534.

- TERZIAN, C., C. FERRAZ, J. DEMAILLE, and A. BUCHETON. 2000. Evolution of the gypsy endogenous retrovirus in the *Drosophila melanogaster* subgroup. *Mol. Biol. Evol.* **17**:908-914.
- THROCKMORTON, L. 1975. The phylogeny, ecology, and geography of *Drosophila*. Pp. 421-469 in R. C. KING, ed. *Handbook of genetics*. Vol. 3. Plenum, New York.
- VALIATI, V. H. 1999. Genética de populações marginais de *Drosophila willistoni* e *D. paulistorum* e avaliação de suas potencialidades para produção de híbridos interespecíficos. Tese de Doutorado, UFRGS, Brasil.
- VÁZQUEZ-MANRIQUE, R. P., M. HERNÁNDEZ, M. J. MARTÍNEZ-SEBASTIÁN, and R. DE FRUTOS. 2000. Evolution of gypsy endogenous retrovirus in the *Drosophila obscura* species group. *Mol. Biol. Evol.* **17**:1185-1193.
- VIEIRA, C., D. LEPETIT, S. DUMONT, and C. BIÉMONT. 1999. Wake up of transposable elements following *Drosophila simulans* worldwide colonization. *Mol. Biol. Evol.* **16**:1251-1255.
- WHEELER, M. R. 1981. The Drosophilidae: a taxonomic overview. Pp. 1-97 in M. ASHBURNER, H. L. CARSON, and J. N. THOMPSON, eds. *The genetics and biology of Drosophila*. Vol. 3a. Academic Press, New York.
- XIONG, Y., and T. H. EICKBUSH. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* **9**:3353-3362.
- ZELENTSOVA, H., H. POLUECTOVA, L. MNJOIAN, G. LYOZIN, V. VELEIKODVORSKAJA, L. ZHIVOTOVSKY, M. G. KIDWELL, and M. B. EVGEN'EV. 1999. Distribution and evolution of mobile elements in the *virilis* species group of *Drosophila*. *Chromosoma* **108**:443-456.

CAPÍTULO 3

Analyses of the gypsy transposable element in neotropical *Drosophila maculifrons* and *D. griseolineata* using *in situ* hybridization in polytene chromosomes and Southern blot.

Fabiana de Oliveira Herédia^a, Elgion Lúcio da Silva Loreto^b and Vera Lúcia da Silva Valente^{a*}

^aDepartamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul. Caixa Postal 15053. CEP 91501-970. Porto Alegre, RS, Brazil.

^b Departamento de Biologia, Universidade Federal de Santa Maria, RS, Brazil.

*Correspondence and reprints

e-mails: vera.gaiesky@ufrgs.br; vera.valente@bol.com.br

Running title: *In situ* localization and Southern blot analysis of the gypsy retrotransposon in *D. guarani* species group.

Trabalho aceito para publicação em Genetics, Selection and Evolution.

Abstract- A sequence homologous to the retrotransposon *gypsy* was localized in *Drosophila griseolineata* and *Drosophila maculifrons* population using *in situ* hybridization in polytene chromosomes. The copy number and integrity of these sequences was estimated by the Southern blot method. A comparison of the results obtained by those two methodologies was also made. The location of the *in situ* hybridization sites and the patterns of cleavage obtained using restriction endonucleases are shared by *D. maculifrons* and *D. griseolineata*, suggesting that the *gypsy* transposon invaded the genomes of these flies before the evolutionary divergence of these species.

Drosophila maculifrons* / *Drosophila griseolineata* / *in situ* hybridization / Southern blot / *gypsy

Résumé- Analyse de l'élément transposable *gypsy* chez les espèces neotropicales de *Drosophila maculifrons* et *D. griseolineata* par hybridation *in situ* des chromosomes polytènes et Southern blot. Une séquence homologue à l'élément transposable *gypsy* a été localisée chez les populations de *Drosophila griseolineata* et de *Drosophila maculifrons* par hybridation *in situ* des chromosomes polytènes. Le nombre de copies et l'intégrité de ces séquences a été estimé par le méthode de Southern blot et les résultats obtenues par les deux méthodes ont été comparés. La localisation des sites d'hybridation et les profils de coupure avec des enzymes de restriction se sont avérés identiques entre *D. maculifrons* et *D. griseolineata*, ce qui suggère que l'élément *gypsy* a envahi les génomes de ces mouches avant la divergence évolutive de ces espèces.

***Drosophila maculifrons* / *Drosophila griseolineata* / *in situ* hybridization / Southern blot / gypsy**

1. INTRODUCTION

The neotropical *Drosophila maculifrons* [12] (= *D. guaramunu*; [10]) and *D. griseolineata* [11] are members of the *guaramunu* subgroup of the *guarani* *Drosophila* group [19]. These two species were collected between latitudes 17° S and 33° S on the eastern slopes of the South America Andean Hills [24]. Within this environment they share common habitats, although *D. maculifrons* have been more widely detected and this species also appears to be more opportunistic from the ecological point of view, because it is capable of occupying a greater range of environments than *D. griseolineata* [14, 23, 27]. Both these species are chromosomally polymorphic in respect to paracentric inversions, although *D. maculifrons* has a higher degree of chromosomal polymorphism [8]. According to [24], these species are, chromosomally, very closely related. It is thought that *D. griseolineata* diverged from *D. maculifrons* [16] or that both diverged from a common ancestor [19]. They are homosequential with respect to the banding patterns of their polytene chromosomes. Salivary gland polytene nuclei of both species present five long euchromatic arms and one small arm, while brain ganglia metaphasic plates show 5 rods and 1 dot element [4]. We have been interested in the *D. guarani* group of flies because it is always well represented in our field collections, and

because this group is a representative sample of the great diversity found in the neotropics.

Transposable elements (TEs) are now known to be ubiquitous components of genomes and have played an important role in eukaryotic evolution [17]. A considerable number of studies have been done on the distribution of TEs in the *Drosophila*, as reviewed by [3], but the endemic neotropical species have been poorly analyzed. For example, only two studies have been done on TEs in the *guarani* group, i.e. the Southern blot study of [9] which showed that the *D. maculifrons* genome lacked *P* elements and the work of [20] who used the dot-blot method to analyze the genome of *D. maculifrons* and *D. griseolineata* for the presence or absence of *P*, *hobo* and *gypsy* TEs, finding only the *gypsy* transposable element in these species.

The *gypsy* retrovirus of *Drosophila* is a member of the *Ty3/gypsy* family of retroelements widely distributed among the Eukaryota [7]. *Gypsy* in *Drosophila melanogaster* are 7.5 kb long, including 482 bp of long terminal repeats (LTRs). *Gypsy* transpose using intermediary RNA, and is controlled by a *Drosophila* gene called *flamenco*, the restrictive alleles of which maintain the retrovirus in a repressed state [22]. Authors as [18] and [25] showed that *gypsy* is an infectious retrovirus. The wide distribution of *gypsy* among the various species of the genus *Drosophila* can be explained both by the presence of an ancestral *gypsy* sequence which was present before the first radiation of this species group and/or by the dissemination of these sequences by horizontal transfers in a way characteristic of an infectious retrovirus [1, 26].

This paper aim to contribute to clarify the role of transposable elements in the generation of variability in the *Drosophila* species which are highly represented in neotropical assemblies.

2. MATERIALS AND METHODS

2.1. *Drosophila* stocks

Adult *Drosophila maculifrons* and *D. griseolineata*, attracted using banana baits, were collected at sites on Morro Santana (Santana hill), Porto Alegre in the Southern Brazilian state of Rio Grande do Sul in May, 1997. They were reared as massal crosses using standard culture medium [21] at 17°C ±1°C and 60% relative humidity.

2.2. Probe

The *gypsy* probe was a pGGHS plasmid isolated from the genome of *D. melanogaster*. It was labelled with dATP-7-biotin (GIBCO/ BRL) for *in situ* hybridization, and with the *random prime* method using the kit Gene Image® (Amersham Pharmacia Biotech) for Southern blot analysis. The restriction map of the *D. melanogaster gypsy* with the cleavage sites of the enzymes used is shown in *Figure 1*.

2.3. Southern blots

Genomic DNA from adult flies of each species was extracted according to the method of [15] and 10 µg of DNA per *D. maculifrons* and *D. griseolineata* sample were digested with *Bam*HI, *Xba*I and *Kpn*I restriction enzymes (GIBCO/

BRL), fractionated in 0.8% agarose gel, transferred to a Hybond N+ (Amersham Pharmacia Biotech) membrane and hybridized according to the manufacturers protocol. The membranes were hybridized to the random prime labeled probe at 60 °C in a mixture containing 0.1% SDS, 5% dextran sulfate and a 20-fold dilution of liquid block in 5xSSC. The filter was washed twice at 60° C, firstly with 1x SSC and 0.1% SDS and then with 0.5 SSC and 1% SDS, in both cases agitating for 15 minutes. For detection, we used the CPD-Star kit (Amersham Pharmacia Biotech) method. The molecular length (in kb) of the detected fragments was determined using fragments of marker 1 kb Plus DNA ladder or *lambda* DNA cleaved with *HindIII*, (both from GIBCO/BRL) as controls.

2.4. *In situ* hybridization procedures

Polytene chromosome squashes of the salivary glands of third instar larvae were prepared according to [2]. Hybridization was carried out in a humid chamber at 37° C in 30% formamide for 40h [13], using 500 ng of labeled probe for each slide. Four individual *D. maculifrons* and three individual *D. griseolineata* were analyzed for gypsy chromosomal insertion sites.

Washes were performed at room temperature with 2xSSC for 10 min, 2xSSC pre-heated at 37° C for 30 min and 2xSSC at room temperature for 30 min. The BluGene® (GIBCO/BRL) non-radioactive detection system was used for identification of the hybridization sites. The slides were stained with an acetic lactic: orcein solution diluted 1:10 in 45% acetic acid, and mounted in Entellan® (Merck). They were analyzed microscopically with the aid of a *D. griseolineata* chromosomal photomap [16], the same map being used for *D. maculifrons* because these species are homosequential in regard to chromosomal order.

The chromosomal bands showing positive hybridization signals were photomicrographed and each observation was marked on the reference map as a dot, dots occurring in the same band in at least 30% of the chromosomes were considered as being reliable.

3. RESULTS AND DISCUSSION

3.1. *In situ* hybridization

All observed chromosomal sites hybridized with *gypsy* in *D. maculifrons* and in *D. griseolineata* lie in the euchromatin. In both the *D. maculifrons* and *D. griseolineata* slides analyzed, the third chromosome showed the highest number of positive hybridization signals with the *gypsy* probe. Consistent labels (with a frequency equal to or higher than 30%) were obtained in sections 47, 54, 57 and 59 of chromosome 3 and section 79 of chromosome 4 of *D. maculifrons*. The consistent labels observed in both species are represented as dots and stars on the chromosomes photomap of [16] (Figure 2). Fewer hybridization labels were found in *D. griseolineata* than in *D. maculifrons*.

Band 59 labeled in *D. maculifrons*, was also labeled in *D. griseolineata*, this shared insertion site possibly being very old, which suggests that an ancient *gypsy* element was probably inserted in this chromosomal region before the divergence of these two species. The antiquity of *gypsy* at this taxonomic level has previously been suggested by [5] and [6]. The aspect of the hybridizations is shown in Figure 3.

3.2. Southern blot

Hybridization between the *D. melanogaster gypsy* probe and the *D. maculifrons* and *D. griseolineata* genomic sequences produced weak signals. This result could have been due to the low homology between them, an hypothesis reinforced by the analyses of the labeled bands obtained by the restriction enzyme digestion of *D. maculifrons* and *D. griseolineata* DNA. Most of the fragments which were expected based on the restriction map of the *D. melanogaster gypsy* transposon were not observed under our samples (Figure 4).

Because the *Bam*HI enzyme does not recognize restriction sites in the *gypsy* region of *D. melanogaster*, *Bam*HI assays estimate the number of *gypsy* copies in the *D. melanogaster* genome. In our *Bam*HI assays, 12 bands were observed in *D. griseolineata* and 9 bands in *D. maculifrons*, but by comparing these findings with those obtained with the other restriction endonucleases it seems that the maximum number of copies does not exceed 6, the number seen after *Kpn*I cleavage. These findings suggests that in *D. maculifrons* and *D. griseolineata* some copies of *gypsy* may contain internal *Bam*HI cleavage sites.

In both *D. griseolineata* and *D. maculifrons* 4 bands were bigger than the *gypsy* bands of *D. melanogaster*. Only these 4 bands could be interpreted as possible complete *gypsy* elements without internal sites, the remaining bands possibly being explained either as deleted *gypsy* copies or as fragments resulting from enzymatic digestion. The deleted copies could be lying in the heterochromatin region, but we did not detect this in the *in situ* hybridization experiments. The number of fragments did not correspond to the number of defective copies found in the other Southern blot experiments with the other

enzymes, which supports the hypothesis of the existence of internal *Bam*HI sites in sequences homologous to *gypsy* in the genome of *D. griseolineata* and *D. maculifrons*.

Assays using *D. maculifrons* and *D. griseolineata* DNA cleaved with *Kpn*I to hybridize with the *gypsy* probe revealed the loss of internal *Kpn*I sites in both species, the expected 1.7 kb band not being found. If no other *Kpn*I site exists in the *gypsy* copies of *D. maculifrons* and *D. griseolineata*, it follows that the total number of bands, six, reflects the maximum number of copies contained in the genome of these species. Of those bands, only two in *D. griseolineata* and three in *D. maculifrons* were of a size comparable with that of complete elements devoid of internal restriction sites. We also observed coincident bands in the two species, one band being about 5.5 kb and the other about 4 kb. Such coincident bands could be due to them having the same chromosomal localization in *D. maculifrons* and *D. griseolineata* (these species having homosequential chromosomes) or to a new restriction *Kpn*I pattern that arose before the species diverged. Considering the coincident label between the chromosomes of the two species found by *in situ* hybridization, the first proposal appears to be more probable.

When the DNA of *D. maculifrons* and *D. griseolineata* was cleaved with *Xba*I, the hybridization experiment gave 4.5 kb bands, similar in size to that expected from their restriction map pattern. Some other expected fragments of 1.6 and 1.0 kb were not observed in *D. maculifrons* and *D. griseolineata*, once again pointing to divergence between the *D. melanogaster gypsy* sequence and the genomic sequences of *D. maculifrons* and *D. griseolineata*. We also observed bands of the similar sizes in the DNA of both species, one

band being about 4 kb and another about 8.5 kb, while a third band was of high molecular weight.

The Southern blot and *in situ* hybridization data allow us to conclude that the maximum number of copies of the *gypsy* retroelement in the genome of *D. maculifrons* and *D. griseolineata* is six, and that the DNA sequence of these two species appears to differ considerably from that of *D. melanogaster* in respect to the restriction patterns of the endonucleases employed in our analysis. It also appears that the *gypsy* transposon invaded the genome of the ancestral species that gave rise to *D. maculifrons* and *D. griseolineata* before the divergence of these two species. This hypothesis explains the sharing of the hybridized fragments found during analysis and the *in situ* hybridized sites in their polytene chromosomes of the two species.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. José A. B. Chies for French translation of the Abstract. The pGGHS plasmid was kindly provided by Dr. Dale Dorsett. This study was performed with fellowships and grants from the Brazilian agencies CNPq, FAPERGS, FINEP and PROPESQ-UFRGS.

REFERENCES

- [1] Alberola, T. M., and De Frutos R., *Gypsy* homologous sequences in *Drosophila subobscura* (*gypsyDS*), *J. Mol. Evol.* 36 (1993) 127-135.
- [2] Ashburner M., Patterns of puffing activity in the salivary gland chromosomes of *Drosophila*, *Chromosoma* 21 (1967) 398-428.
- [3] Biémont C. and Cizeron G., Distribution of transposable elements in *Drosophila* species, *Genetica* 105 (1999) 43-62.
- [4] Brcic D. Chromosomal variation in natural populations of *Drosophila guaramunu*, *Z. Indukt. Abstamm.-u. Vererblehre* 85 (1953) 1-11.
- [5] Britten R.J., Active *gypsy/Ty3* retrotransposons or retroviruses in *Caenorhabditis elegans*, *Proc. Nac. Acad. Sci. USA* 92 (1995) 599-601.
- [6] Bucheton A., The relationship between the *flamenco* gene and *gypsy* in *Drosophila*: how to tame a transposon. *Trends in Genetics* 11 (1995) 349-353.
- [7] Capy P., Bazin C., Higuët D. and Langin T., *Dynamics and Evolution of Transposable Elements*, Molecular Biology Intelligence Unit, Landes Bioscience, Texas, USA, 1998, 197 pp.
- [8] Da Cunha A.B., Brcic D. and Salzano F.M., A comparative study of chromosomal polymorphism in certain South American species of *Drosophila*, *Heredity* 7 (1953) 193-202.
- [9] Daniels S.B., Peterson K.R., Strausbaugh L.D., Kidwell M.G. and Chovnick A., Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species, *Genetics* 124 (1990) 339-355.
- [10] Dobzhansky Th. and Pavan C., Studies on Brazilian species of *Drosophila*, *Bolm. Fac. Filos. Cienc. S. Paulo* 36 (1943) 7-72.

[11] Duda O., Die costaricanischen Drosophiliden des Ungarischen National Museums zu Budapest, *Annl. hist. -nat. Mus. natn. hung.* 22 (1925) 149-229.

[12] Duda O., Die südamerikanischen Drosophiliden (Dipteren) unter Berücksichtigung auch der anderen neotropischen sowie der nearktischen Arten, *Arch. Naturgesch. Abt. A* 91 (1925) 1-228.

[13] Engels W.R., Preston C.R., Thompson P. and Egglestone N.B., "In situ" hybridization to *Drosophila* salivary chromosomes with biotinylated DNA probes and alkaline phosphatase, *Focus* 8 (1986) 6-8.

[14] Franck G. and Valente V.L.S., Study on the fluctuation in *Drosophila* populations of Bento Gonçalves, RS, Brazil, *Revta. bras. Biol.* 45 (1985) 133-141.

[15] Jowett T., Preparation of nucleic acids, in Roberts D.B. (Ed.), *Drosophila: a Practical Approach*, IRL Press, Washington, DC, 1986, pp. 275-286.

[16] Kastriasis C.D., The chromosomes of some species of the *guarani* group of *Drosophila*, *J. Hered.* 60 (1969) 51-57.

[17] Kidwell M.G. and Lisch D.R., Transposable elements and host genome evolution, *Trends in Ecology and Evolution* 15 (2000) 95-99.

[18] Kim A., Terzian C., Santamaria P., Pélisson A., Prud'homme N. and Bucheton A., Retrovirus in invertebrates: The gypsy retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*, *Proc. Natl. Acad. Sci. USA* 91 (1994) 1285-1289.

[19] King J.C., A comparative analysis of the chromosomes of the *guarani* group of *Drosophila*, *Evolution* 1 (1947) 48-62.

[20] Loreto E.L.S., Basso da Silva L., Zaha A. and Valente V.L.S., Distribution of transposable elements in neotropical species of *Drosophila*, *Genetica* 101 (1998) 153-165.

[21] Marques E.K., Napp M., Winge H. and Cordeiro A.R., A corn meal, soybean flour, wheat germ medium for *Drosophila*, *Drosophila Inf. Ser.* 41 (1966) 187.

[22] Pélisson A., Teyssset L., Chalvet F., Kim A., Prud'homme N., Terzian C. and Bucheton A., About the origin of retroviruses and the co-evolution of the gypsy retrovirus with the *Drosophila flamenco* host gene, *Genetica* 100 (1997) 29-37.

[23] Saavedra C.C.R., Callegari-Jacques S.M., Napp M. and Valente V.L.S., A descriptive and analytical study of four neotropical drosophilid communities, *J. Zool. Syst. Evol. Res.* 33 (1995) 62-74.

[24] Salzano F.M., Chromosomal polymorphism in two species of the guarani group of *Drosophila*, *Chromosoma* 7 (1955) 39-50.

[25] Song S. U., Gerasimova T., Kurkulos M., Boeke J. D. and Corces V. G., An env-like protein encoded by a *Drosophila* retroelement: evidence that gypsy is an infectious retrovirus, *Gen. Dev.* 8 (1994) 2046-2057.

[26] Stacey S.N., Lansman, R.A., Brock H.W. and Grigliatti T.A., Distribution and conservation of mobile elements in the genus *Drosophila*, *Mol. Biol. Evol.* 3 (1986) 522-534.

[27] Valente V.L.S. and Araújo A.M., Ecological aspects of *Drosophila* species in two contrasting environments in southern Brazil (Diptera, Drosophilidae), *Revta. bras. Ent.* 35 (1991) 237-253.

Figura 1 no anexo.

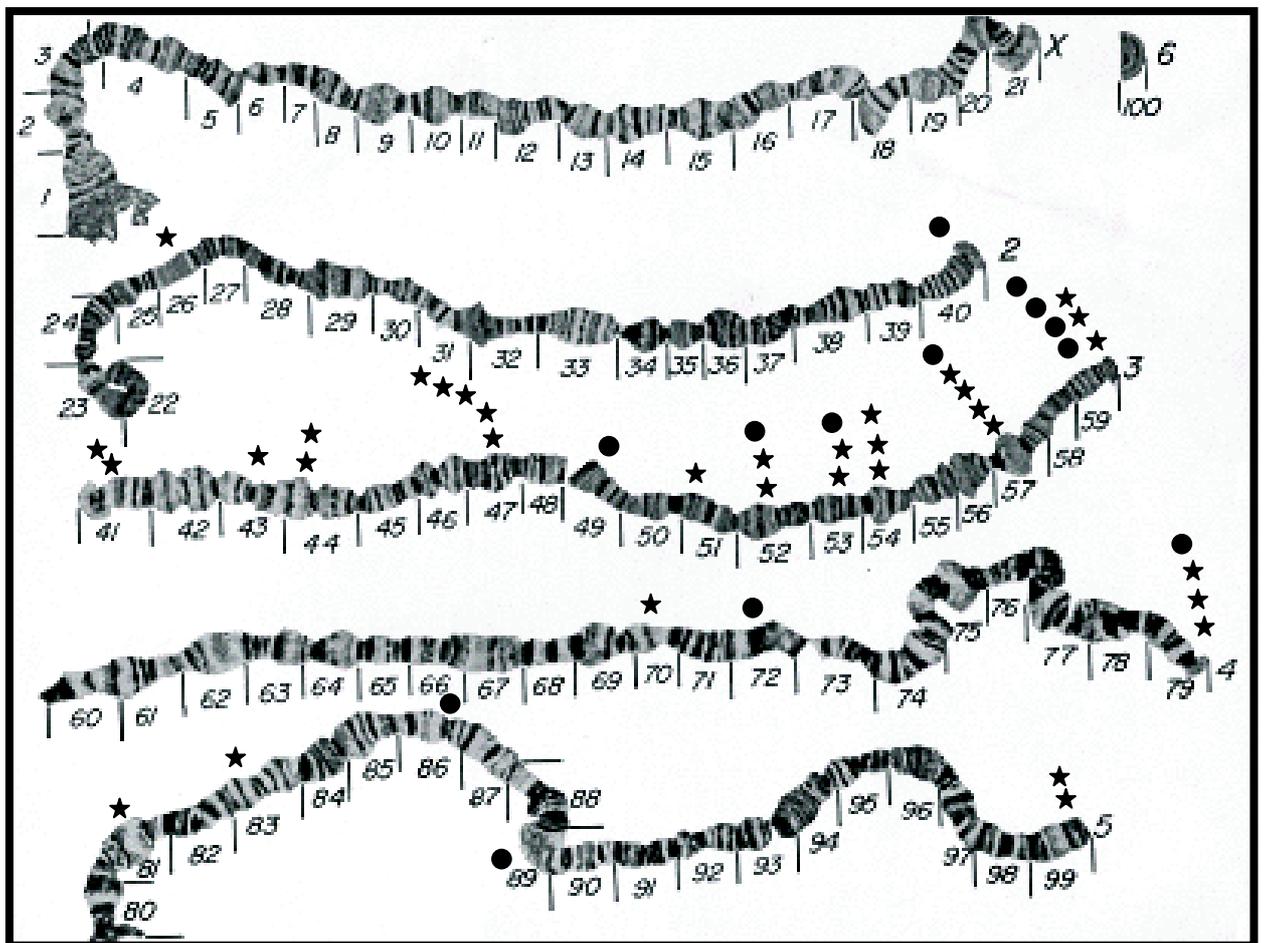


Figure 2.

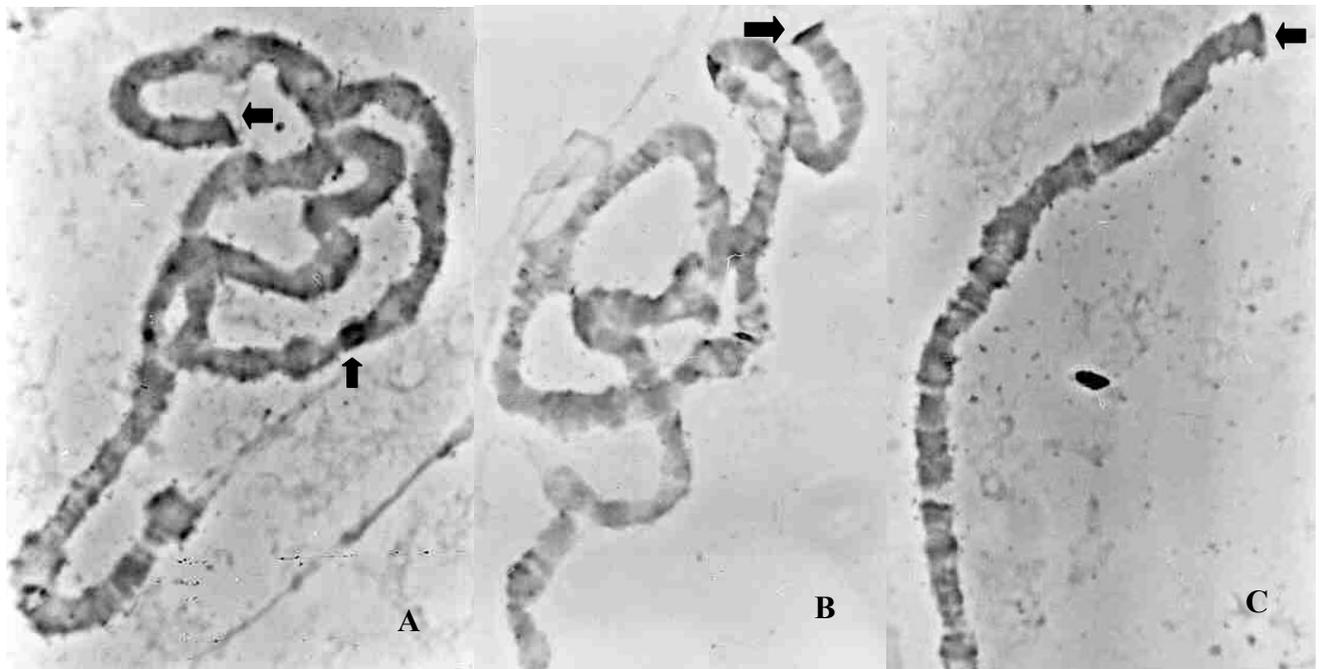


Figure 3.

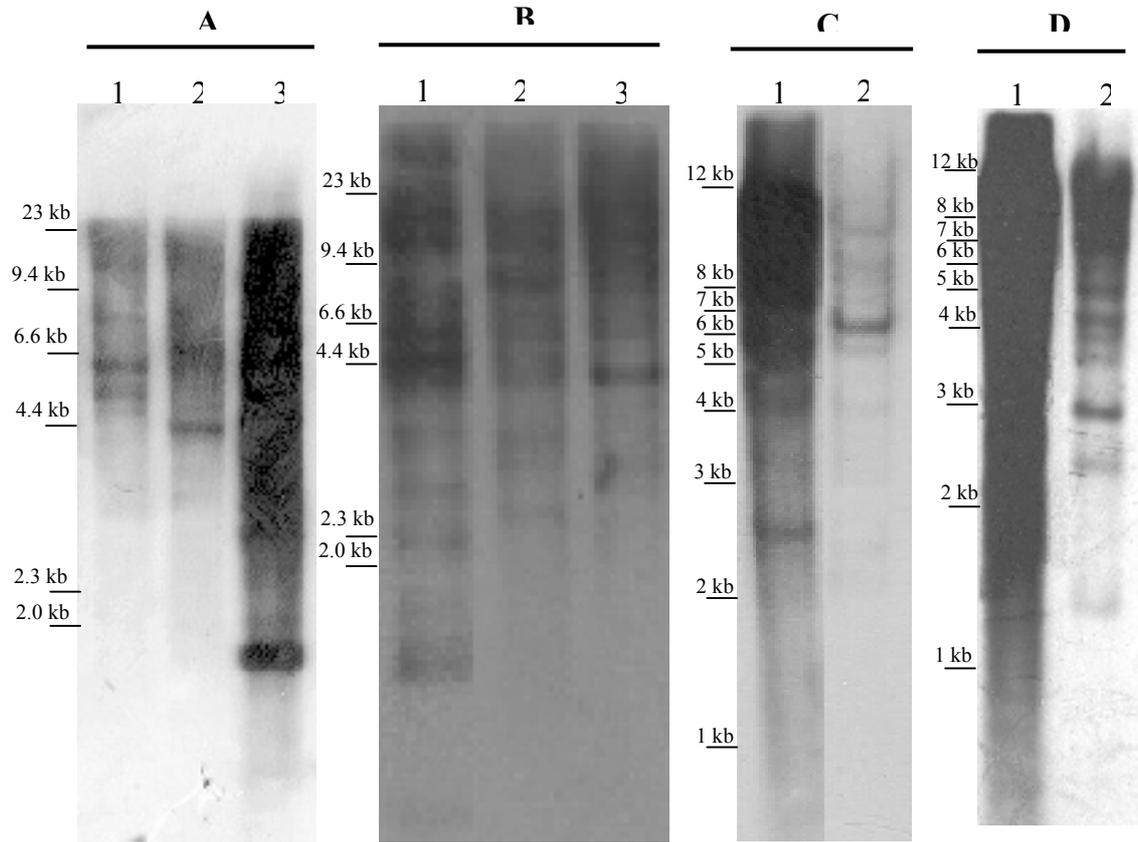


Figure 4.

Figure legends

Figure 1. Representation of a complete *gypsy* transposon of *D. melanogaster*.

Figure 2. Photomap of *D. griseolineata* polytene chromosomes, subdivided according to Kastritsis [16], in which there are represented the hybridization sites with the *gypsy* probe. Dots above the chromosomes correspond to the hybridization sites detected in *D. griseolineata* and stars correspond to those found in *D. maculifrons*.

Figure 3. Aspect of some positive *in situ* hybridization signals of *gypsy* in polytene chromosomes of: (A) and (B) *D. maculifrons* and (C) *D. griseolineata*. (A) Labels in sections 52 and 59 (third chromosome); (B) Label in section 79 (fourth chromosome); (C) Label in section 59 (third chromosome). x588.

Figure 4. Southern blot analysis of *gypsy* in *D. maculifrons*, *D. griseolineata* and *D. melanogaster* (positive control). (A) *KpnI* digested genomic DNA of (1) *D. maculifrons*, (2) *D. griseolineata* and (3) *D. melanogaster*; (B) *XbaI* digested genomic DNA of (1) *D. melanogaster*, (2) *D. griseolineata* and (3) *D. maculifrons*; (C) *BamHI* digested genomic DNA of (1) *D. melanogaster* and (2) *D. griseolineata*; (D) *BamHI* digested genomic DNA of (1) *D. melanogaster* and (2) *D. maculifrons*.

The molecular length (kb) of the fragments detected was determined by comparison with *HindIII* digested *lambda* DNA or with 1kb Plus DNA ladder (GIBCO).

CAPÍTULO 4

DISCUSSÃO GERAL E PERSPECTIVAS

Elementos transponíveis são seqüências egoístas, selecionadas para uma alta atividade transposicional ou necessárias para a sobrevivência de seu hospedeiro? Constituem o DNA lixo da célula ou uma grande quantidade de fragmentos domesticados com funções, talvez, ainda, desconhecidas?

Que os TEs representam uma grande força, que participa da evolução dos seus genomas hospedeiros, isto é bem documentado e aceito pela comunidade científica, mas este processo é acidental, do ponto de vista destes elementos, ou realmente estabelece-se uma co-evolução?

O presente trabalho é uma tentativa de entendimento destas questões. A complexa evolução do retrotransposon *gypsy*, apontada pelos resultados aqui obtidos, remete-nos para o quão instigantes são as respostas às novas questões aqui levantadas.

O que nos parece razoavelmente claro é a existência de múltiplas cópias de *gypsy*, evoluindo independentemente no genoma, levando a uma heterogeneidade de seqüências, traduzidas no complexo padrão de fragmentos hibridizados nos ensaios de *Southern blot*. A permanência destas diferentes cópias nos genomas hospedeiros pode culminar na divergência destas seqüências. Possivelmente, a maioria das cópias encontradas sejam *gypsy* degenerados, sem qualquer potencial infectivo ou possibilidade de

transposição. Este parece ser o destino de TEs num genoma. Se alguma função, estrutural ou funcional, puder ser aproveitada destas cópias degeneradas rumo ao desaparecimento, é esperado que isto ocorra e seja selecionado. Talvez deste processo tenha emergido a enorme gama de funções que hoje são relacionadas aos elementos transponíveis.

Mutações e mudanças regulatórias na expressão de genes por inserção de TEs em regiões codificantes ou próximas a elas, rearranjos cromossômicos, resultantes de quebras provocadas pela saída e entrada destes elementos ou recombinação entre cópias de uma mesma família de TE são algumas das conseqüências imediatas de sua presença no genoma (revisão em Capy et al. 1998; Deininger e Batzer 1999; Pickeral et al. 2000; Cáceres et al. 2001). Em uma escala evolutiva, atribui-se aos elementos transponíveis o surgimento de mecanismos de regulação e supressão de seqüências repetidas, como metilação, compartimentalização nuclear, interferência de RNA, entre outros (Fedoroff 2000; Plasterk e Ketting 2000). Surgimento de íntrons, expansão de genomas, perpetuação de telômeros, estruturação de regiões heterocromáticas, estabelecimento do ciclo sexual nos seres vivos, maior capacidade de resposta à seleção de caracteres quantitativos em ambientes estressantes, matéria-prima para a domesticação de funções e seqüências dos TEs pelos seus genomas hospedeiros, são algumas das especulações feitas em torno da importância dos elementos transponíveis (Lin et al. 1998; Fedoroff 2000; Kidwell e Lisch 2000; Kidwell e Lisch 2001; Petrov 2001; Wilder e Hollocher 2001; Wright e Finnegan 2001).

Como seqüências autônomas, TEs devem visar à sua expansão a nível genômico e populacional. Silva e Kidwell (2000) têm proposto que a principal e

talvez a única fonte de seleção realmente atuante sobre estas seqüências, seja o da transmissão horizontal. Este seria o mecanismo essencial para a sobrevivência dos elementos móveis. Neste sentido, a presença de elementos completos e/ou capazes de fazerem cópias de si mesmos no genoma traduz-se em hóspedes recentes, de um ponto de vista evolutivo.

Kidwell (1993), numa revisão sobre a relevância do fenômeno de transmissão horizontal para a relação TE/hospedeiro, salienta a raridade deste processo. No entanto, a autora reconhece a limitação dos dados disponíveis e não descarta sua significância evolutiva.

Certamente, problemas de paralogia, recombinação, variação nas taxas evolutivas e polimorfismo ancestral dificultam a identificação precisa de eventos de transferência laterais. Contudo, em muitos casos, a hipótese evolutiva de longe mais parcimoniosa encontra-se enredada a este fenômeno.

A disponibilidade de seqüências de genomas inteiros, em banco de dados público, possibilita uma enorme gama de análises comparativas e abre novas e ricas perspectivas para estudos evolutivos. Wolf et al. (1999) compararam o genoma de eubactérias com arqueobactérias e registraram evidências de transferência horizontal de genes entre estes grupos. Além disto, os autores detectaram vários casos de troca gênica recíproca entre bactérias parasitas intracelulares obrigatórias e seus hospedeiros eucarióticos, assim como entre bactérias distantemente relacionadas mas com nichos ecológicos sobrepostos.

Doolittle (1999) postula que uma nova era para o entendimento da biologia evolutiva está surgindo a partir de um panorama montado pelos

dados de seqüências genômicas agora disponíveis. O autor discute que a visão tradicional que postula uma única e universal árvore filogenética para explicar as relações entre todos os seres vivos tem recentemente sido questionada. Ao invés desta, uma outra teoria está se fundamentando: a que defende que as verdadeiras relações entre os seres vivos só podem ser representadas por uma rede complexa de relações. Tais relações, muitas vezes, estabelecem-se entre entidades não aparentadas. A rede referida representaria, neste contexto, múltiplos eventos de transmissão horizontal.

Incongruências entre seqüências gênicas de eubactérias, arqueobactéria e eucariotos têm sido reportadas. Estas constatações são baseadas em várias análises como conteúdo G + C, uso de códons preferenciais, bem como organização genética, e todos resultados apontam para a existência de um grande número de genes "estrangeiros", de recente chegada no genoma (revisão em Doolittle 1999; Dröge et al. 1998). Como exemplos deste fenômeno, temos os dados de Lawrence e Roth (1996) que concluem que 18% dos genes de *E. coli* foram introduzidos em seu genoma por transmissão horizontal nos últimos 100 milhões de anos. Adicionalmente, 24% do genoma de uma espécie de eubactéria hipertermófila parece terem sido providos por genes de uma arqueobactéria que convive com a primeira (Nelson et al. 1999).

Outros casos de transmissão horizontal envolvem genes de *E. coli* e células humanas, fungos e bactérias, assim como *Agrobacterium* e tabaco. Este último sistema tem sido, inclusive, descrito como um indício de plantas naturalmente transgênicas (revisão em Dröge et al. 1998).

Estes últimos trabalhos, que tratam da discussão do papel da transferência horizontal para a evolução dos seres vivos, vêm nos alertando para o fato deste fenômeno ser mais freqüente do que pensávamos. Neste contexto, os elementos transponíveis, com sua maquinaria própria e capacidade de inserção, figuram como ótimos vetores de troca genética.

Características ecológicas próprias das assembléias de Drosofilídeos, muitos deles compartilhando substratos de criação e de alimentação e convivendo intimamente com parasitas como ácaros, parasitóides como microhimenópteros e uma ampla gama de microorganismos podem oportunizar tais eventos de transferência horizontal. Houck et al. (1991) encontrando o elemento *P* no genoma de um ácaro parasita de *Drosophila* e sugerem o seu possível papel como vetor deste processo.

Se a magnitude da transferência lateral é verdadeira, conclusões acerca de filogenias moleculares, principalmente aquelas baseadas em moléculas individuais, necessitam ser cautelosamente examinadas. Doolittle (1999) chama à atenção para o fato de que muitos genes com funções metabólicas importantes na célula são “importados” em bactérias. Isto remete-nos para a falta de garantia de um ancestral comum explicando a ampla distribuição de um gene conservado. O autor não descarta as hipóteses alternativas à transmissão horizontal mas salienta que elas violam as regras de parcimônia, tão radicalmente defendidas pelos filogeneticistas.

Certamente, os diferentes grupos de organismos relacionados filogeneticamente continuam retendo suas relações de parentesco, a despeito da troca de seqüências gênicas por via lateral. Muitos genes podem

evoluir juntos por longos períodos evolutivos. Contudo, a existência e a extensão do fenômeno de transmissão horizontal não deve ser renegada.

Talvez, filogenias obtidas de seqüências de elementos transponíveis, como aquelas publicadas para *P. mariner*, assim como a do presente trabalho para *gypsy*, nos ensine um pouco sobre a extensão da “promiscuidade” dos genes.

Nossos resultados podem representrar, ainda, a “ponta de um iceberg”. Se um número maior de genomas fosse investigado, possivelmente, muitos outros eventos de transmissão horizontal pudessem ser inferidos, bem como a direção destes eventos. Além disso, dados de seqüências de genes hospedeiros de um maior número de espécies, principalmente daquelas neotropicais que, freqüentemente, não são representadas nas filogenias de espécies até hoje publicadas, ajudariam na construção de um panorama evolutivo mais claro.

Uma análise mais detalhada foi o que iniciamos, de forma particularizada com o grupo *guarani*, do qual investigamos o número e a localização genômica de *gypsy*, por *Southern blot* e hibridização “*in situ*” em *Drosophila maculifrons* e *D. grisolineata*. A expansão de estudos deste tipo, com espécies relacionadas, pode ajudar a esclarecer algumas destas questões.

Finalmente, existe um limite de invasão de um determinado genoma hospedeiro por um elemento específico? A garimpagem das diferentes subfamílias de *gypsy*, diagnosticadas em nossas análises, nos genomas das espécies estudadas, poderá também contribuir para o entendimento desta questão.

Quem sabe um dos grandes desafios da era genômica seja o de dimensionar o fenômeno de transmissão horizontal de seqüências gênicas entre os seres vivos. Não seria surpreendente se, em um futuro trabalho, pudéssemos traçar o perfil cronológico da mudança de pensamento científico acerca do assunto, como aquele que mostramos nas primeiras páginas desta tese, para a descoberta e importância dos elementos transponíveis.

RESUMO

A fim de contribuir para o entendimento da evolução de elementos transponíveis (TEs) no genoma de suas espécies hospedeiras, nós investigamos a distribuição, a representatividade e a conservação de seqüências homólogas ao retrotransposon *gypsy* de *Drosophila melanogaster* em 40 espécies, pertencentes ao gênero *Drosophila* e uma espécie incluída no gênero *Zaprionus* (*Z. indianus*).

O elemento *gypsy* pertence a uma superfamília de TEs extremamente relacionada aos retrovírus e caracteriza-se pela conservação da terceira ORF. Em algumas linhagens de *D. melanogaster*, o gene encontrado nesta ORF transcreve e codifica uma proteína similar àquela do envelope capacitando, assim, estes elementos móveis a formar partículas virais.

Nossos resultados não só confirmam a ampla distribuição de *gypsy* no Gênero *Drosophila*, mas ampliam seu estudo para várias espécies neotropicais e para o Gênero *Zaprionus*. Seqüências *gypsy*, nos mais diferentes genomas investigados, demonstram uma alta divergência em relação ao elemento de *D. melanogaster*. No entanto, a conservação de sítios de restrição entre espécies não relacionadas filogeneticamente, instigou-nos a uma análise mais minuciosa das relações entre estas seqüências. Para tanto, nós reportamos um estudo filogenético baseado em, aproximadamente, 500 pb do gene *env*. Nossas análises indicam grandes incongruências entre a filogenia traçada a partir do elemento *gypsy* e as relações de parentesco entre suas espécies hospedeiras. Os dados obtidos, a partir de seqüências do elemento *gypsy* de 23 espécies, permitem-nos inferir sobre um complexo padrão evolutivo que

parece incluir polimorfismo ancestral, transmissão vertical, taxas evolutivas desiguais em diferentes genomas e, principalmente, múltiplos casos de transmissão horizontal. Este último fenômeno pode ser corroborado e facilitado pelo potencial infectivo destes elementos móveis.

Finalmente, o padrão de transmissão vertical de *gypsy* para duas espécies do grupo *guarani*, inferido das análises de seqüência, foram confirmados por uma análise mais detalhada, através do método de hibridização *in situ*. Um sítio de inserção de *gypsy*, coincidente nos braços eucromáticos dos cromossomos politênicos de *D. maculifrons* e *D. griseolineata*, indica a antigüidade deste retroelemento em seus genomas.

SUMMARY

Aiming to contribute to the comprehension of the evolution of transposable elements (TEs) in the genome of the host species, we investigated the distribution, representativity and the conservation of sequences homologous to the retrotransposon *gypsy* of *Drosophila melanogaster* in 41 species of the Genus *Drosophila* and in one species of the Genus *Zaprionus* (*Z. indianus*).

The *gypsy* retroelement is member of a superfamily of TEs extremely related with the retroviruses, and is characterized by the conservation of the third ORF. In some strains of *D. melanogaster*, the gene found in this ORF is able to transcribe and to codify a protein similar to that of the envelope, allowing this transposable element to form viral particules.

Our results extended the knowledge about the wide distribution of *gypsy* in the Genus *Drosophila*, including several neotropical species until now not studied, and to the Genus *Zaprionus*. Sequences homologous to *gypsy* in the several different genomes, showed a high divergence respect to the *gypsy* element of *D. melanogaster*.

Nevertheless, the conservation of the restriction sites between species unrelated phylogenetically, instigated us to perform a further analysis of the relationship between these sequences. For this, we report a phylogenetic study based in approximately 500 bp of the gene *env*. Our analyses showed considerable discrepancy between the phylogeny of *gypsy* and the relationship of their host species. The data obtained of the *gypsy* homologous sequences in 23 *Drosophila* species allow to infer a complex evolutionary

pattern that include ancestral polymorphism, vertical transmission, unequal evolutionary rates in different genomes and, mainly, several cases of horizontal transmission. This last phenomenon could be facilitated by the infective potential of the gypsy elements.

Finally, the pattern of vertical transmission of gypsy for two species of the *guarani* group, inferred from the analyses of the sequences, were confirmed by a more accurate analysis, through the *in situ* hybridization method. One site of insertion of gypsy coincident in the euchromatic arms of the polytene chromosomes of *D. maculifrons* and *D. griseolineata* indicate the antiquity of this retroelement in their genomes.

REFERÊNCIAS BIBLIOGRÁFICAS

- AKASHI, H. 1994. Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. *Genetics* **136**:927-935.
- ALBEROLA, T. M., and R. De FRUTOS. 1993a. Gypsy homologous sequences in *Drosophila subobscura* (gypsyDS). *J. Mol. Evol.* **36**:127-135.
- ALBEROLA, T. M., and R. De FRUTOS. 1993b. Distribution of gypsy sequences in *Drosophila* species of the obscura subgroup. *Hereditas* **118**:131-137.
- ALBEROLA, T. M., and R. De FRUTOS. 1996. Molecular structure of a gypsy element of *Drosophila subobscura* (gypsyDS) constituting a degenerate form of insect retroviruses. *Nucleic Acids Res.* **24**:914-923.
- ALBEROLA, T. M., L. BORI, and R. De FRUTOS. 1997. Structural analysis of *Drosophila subobscura* gypsy elements (gypsyDs). *Genetica* **100**:39-48.
- ARCÀ, B., and C. SAVAKIS. 2000. Distribution of the transposable element *Minos* in the genus *Drosophila*. *Genetica* **108**:263-267.
- ARKHIPOVA, I., and M. MEELSON. 2000. Transposable elements in sexual and ancient asexual taxa. *Proc. Natl. Acad. Sci. USA* **97**:14473-14477.
- ASHBURNER, M. 1967. Patterns of puffing activity in the salivary gland chromosomes of *Drosophila*. *Chromosoma* **21**:398-428.
- AVEDISOV, S. N., and Y. V. ILYIN. 1994. Identification of spliced RNA species of *Drosophila melanogaster* gypsy retrotransposon: new evidence for retroviral nature of the gypsy element. *FEBS Lett.* **350**:147-150.
- BAYEV Jr., A. A., N. V. LYUBOMIRSKAYA, E. B. DZHUMAGALIEV, E. V. ANANIEV, I. G. AMIANTOVA, and Y. V. ILYIN. 1984. Structural organization of transposable element mdg4 from *Drosophila melanogaster* and a nucleotide sequence of its long terminal repeats. *Nucleic Acids Res.* **12**:3707-3723.

- BELL, A. C., A. G. WEST, and G. FELSENFELD. 2001. Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science* **291**:447-450.
- BENIT, L., N. De PARSEVAL, J. F. CASELLA, I. CALLEBAUT, A. CORDONNIER, and T. HEIDMANN. 1997. Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a gag coding sequence closely related to the *Fv1* restriction gene. *J. Virol.* **71**:5652-5657.
- BIÉMONT, C. 1992. Population genetics of transposable DNA elements: a *Drosophila* point of view. *Genetica* **86**: 67-84.
- BIÉMONT, C., and G. CIZERON. 1999. Distribution of transposable elements in *Drosophila* species. *Genetica* **105**:43-62.
- BIRD, A. P. 1995. Gene number, noise reduction and biological complexity. *Trends Genet.* **11**:94-100.
- BLESA, D., M. GANDÍA, and M. J. MARTÍNEZ-SEBASTIÁN. 2001. Distribution of the *bilbo* non-LTR retrotransposon in *Drosophilidae* and its evolution in the *Drosophila obscura* species group. *Mol. Biol. Evol.* **18**:585-592.
- BRITTEN, R. J. 1995. Active *gypsy*/*Ty3* retrotransposons or retroviruses in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **92**:599-601.
- BRITTEN, R. J. 1997. Mobile elements inserted in the distant past have taken on important functions. *Gene* **205**:177-182.
- BRNCIC, D. 1953. Chromosomal variation in natural populations of *Drosophila guaramunu*. *Z. Indukt. Abstamm.-u. Vererblehre.* **85**:1-11.
- BUCHETON, A. 1995. The relationship between the *flamenco* gene and *gypsy* in *Drosophila*: how to tame a retrovirus. *Trends Genet.* **11**:349-353.
- CÁCERES, M., J. M. RANZ, A. BARBADILLA, M. LONG, and A. RUIZ. 1999. Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **285**:415-418.

- CÁCERES, M., M. PUIG, and A. RUIZ. 2001. Molecular characterization of two natural hotspots in the *Drosophila buzzatii* genome induced by transposon insertions. *Genome Res.* **11**:1353-1364.
- CAI, H. N., and M. LEVINE. 1997. The gypsy insulator can function as a promoter-specific silencer in the *Drosophila* embryo. *EMBO J.* **16**:1732-1741.
- CAPY, P., D. ANXOLABÉHÈRE, and T. LANGIN. 1994. The strange phylogenies of transposable elements: are horizontal transfers the only explanation? *Trends Genet.* **10**:7-12.
- CAPY, P., R. VITALIS, T. LANGIN, D. HIGUET, and C. BAZIN. 1996. Relationships between transposable elements based upon the integrase-transposase domains: is there a common ancestor? *J. Mol. Evol.* **42**:359-368.
- CAPY, P., T. LANGIN, D. HIGUET, P. MAURER, and C. BAZIN. 1997. Do the integrases of LTR-retrotransposons and class II element transposases have a common ancestor? *Genetica* **100**:63-72.
- CAPY, P., C. BAZIN, D. HIGUET, and T. LANGIN. 1998. Dynamics and evolution of transposable elements. Landes Bioscience, Austin, Tex.
- CAPY, P., G. GASPERI, C. BIÉMONT, and C. BAZIN. 2000. Stress and transposable elements: co-evolution or useful parasites? *Heredity* **85**:101-106.
- CARARETO, C. M., W. KIM, M. F. WOJCIECHOWSKI, P. O'GRADY, A. V. PROKCHOROVA, J. C. SILVA, and M. G. KIDWELL. 1997. Testing transposable elements as genetic drive mechanisms using *Drosophila P* element constructs as a model system. *Genetica* **101**:13-33.
- CARSON, H. L., and H. D. STALKER. 1951. Natural breeding sites for some wild species of *Drosophila* in the eastern United States. *Ecology* **32**:317-330.
- CARSON, H. L. 1965. Chromosomal morphism in geographically widespread species of *Drosophila*. Pp. 503-531 in *The genetics of colonizing species*, Academic Press, New York.

- CHALVET, F., C. DI FRANCO, A. TERRINONI, A. PELISSON, N. JUNAKOVIC, and A. BUCHETON. 1998. Potentially active copies of the gypsy retroelement are confined to the Y chromosome in some strains of *Drosophila melanogaster* possibly as the result of the female-specific effect of the *flamenco* gene. *J. Molec. Evol.* **46**:437-441.
- CHALVET, F., L. TEYSSET, C. TERZIAN, N. PRUD'HOMME, P. SANTAMARIA, A. BUCHETON, and A. PÉLISSON. 1999. Proviral amplification of the Gypsy endogenous retrovirus of *Drosophila melanogaster* involves env-independent invasion of the female germline. *EMBO J.* **18**:2659-2669.
- CIZERON, G., F. LEMEUNIER, C. LOEVENBRUCK, A. BREHM, and C. BIÉMONT. 1998. Distribution of the retrotransposable element 412 in *Drosophila* species. *Mol. Biol. Evol.* **15**:1589-1599.
- CLARK, J. B., W. P. MADDISON, and M. G. KIDWELL. 1994. Phylogenetic analysis supports horizontal transfer of P transposable elements. *Mol. Biol. Evol.* **11**:40-50.
- CLARK, J. B., and M. G. KIDWELL. 1997. A phylogenetic perspective on P transposable element evolution in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **94**:11428-11433.
- CORDEIRO, A. R., and H. WINGE. 1995. Levels of evolutionary divergence of *Drosophila willistoni* sibling species. In: *Genetics of natural populations: The continuing importance of T. Dobzhansky, L. Levine (ed.)*, Columbia University Press, New York.
- DA CUNHA, A. B., D. BRNCIC, and F. M. SALZANO. 1953. A comparative study of chromosomal polymorphism in certain South American species of *Drosophila*. *Heredity* **7**:193-202.
- DANIELS, S. B., A. CHOVNICK, and I. A. BOUSSY. 1990a. Distribution of hobo transposable elements in the genus *Drosophila*. *Mol. Biol. Evol.* **7**:589-606.

- DANIELS, S. B., K. R. PETERSON, L. D. STRAUSBAUGH, M. G. KIDWELL, and A. CHOVNICK. 1990b. Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species. *Genetics* **124**:339-355.
- De FRUTOS, R., K. R. PETERSON, and M. G. KIDWELL. 1992. Distribution of *Drosophila melanogaster* transposable element sequences in species of the *obscura* group. *Chromosoma* **101**:293-300.
- DEININGER, P. L., and M. A. BATZER. 1999. Alu repeats and human disease. *Mol. Genet. Metab.* **67**:183-193.
- DEJ, K. J., T. GERASIMOVA, V. G. CORCES, and J. D. BOEKE. 1998. A hotspot for the *Drosophila gypsy* retroelement in the *ovo* locus. *Nucleic Acids Res.* **26**:4019-4024.
- DENNER, J. 1998. Immunosuppression by retroviruses: implications for xenotransplantation. *Ann. N. Y. Acad. Sci.* **862**:75-86.
- DeSALLE, R. 1992. The phylogenetic relationships of flies in the Family Drosophilidae deduced from mtDNA sequences. *Mol. Phylogenet. Evol.* **1**:31-40.
- DIMITRI, P., B. ARCA, L. BERGHELLA, and E. MEI. 1997. High genetic instability of heterochromatin after transposition of the LINE-like I factor in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **94**:8052-8057.
- DIMITRI, P., and N. JUNAKOVIC. 1999. Revising the selfish DNA hypothesis new evidence on accumulation of transposable elements in heterochromatin. *Trends Genet.* **15**:123-124.
- DOBZHANSKY, T., and C. PAVAN. 1943. Studies on Brazilian species of *Drosophila*. *Bolm. Fac. Filos. Cienc. S. Paulo* **36**:7-72.
- DOOLITTLE, W. F. 1999. Lateral genomics. *Trends Cell Biol.* **9**:M5-8.

- DORSETT, D., G. A. VIGLIANTI, B. J. RUTLEDGE, and M. MESELSON. 1989. Alteration of *hsp82* gene expression by the gypsy transposon and suppressor genes in *Drosophila melanogaster*. *Genes Dev.* **3**:454-468.
- DOWSETT, A. P., and M. W. YOUNG. 1982. Differing levels of dispersed repetitive DNA among closely related species of *Drosophila*. *Proc. Natl. Acad. Sci. USA* **79**:4570-4574.
- DRAKE, J. W. 1993. Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. USA* **90**:4171-4175.
- DRÖGE, M., A. PÜHLER, and W. SELBITSCHKA. 1998. Horizontal gene transfer as a biosafety issue: a natural phenomenon of public concern. *J. Biotechnol.* **64**:75-90.
- DUDA, O. 1925. Die costaricanischen Drosophiliden des Ungarischen National Museums zu Budapest. *Annls. hist. -nat. Mus. natn. hung.* **22**:149-229.
- DUDA, O. 1927. Die südamerikanischen Drosophiliden (Dipteren) unter Berücksichtigung auch der anderen neotropischen sowie der nearktischen Arten. *Arch. Naturgesch. Abt. A* **91**:1-228.
- DWYER, D. S. 1998. Assembly of exons from unitary transposable genetic elements: implications for the evolution of protein-protein interactions. *J. Theor. Biol.* **194**:11-27.
- EICKBUSH, D. G., W. C. LATHE III, M. P. FRANCINO, and T. H. EICKBUSH. 1995. *R1* and *R2* retrotransposable elements of *Drosophila* evolve at rates similar to those of nuclear genes. *Genetics* **139**:685-695.
- ENGELS, W. R., C. R. PRESTON, P. THOMPSON, and N. B. EGGLESTONE. 1986. "In situ" hybridization to *Drosophila* salivary chromosomes with biotinylated DNA probes and alkaline phosphatase. *Focus* **8**:6-8.
- FEDOROFF, N. 2000. Transposons and genome evolution in plants. *Proc. Natl. Acad. Sci. USA* **97**:7002-7007.

- FLAVELL, A. J., L. S. ALPHEY, S. J. ROSS, and A. J. LEIGH-BROWN. 1990. Complete reversions of a *gypsy* retrotransposon-induced *cut* locus mutation in *Drosophila melanogaster* involving *jockey* transposon insertions and flanking *gypsy* sequence deletions. *Mol. Gen. Genet.* **220**:181-185.
- FLAVELL, A. J. 1992. *Ty1-copia* group retrotransposons and the evolution of retroelements in eukaryotes. *Genetica* **86**:203-214.
- FLAVELL, A. J., S. R. PEARCE, (J.S.) P. HESLOP-HARRISON, and A. KUMAR. 1997. The evolution of *Ty1-copia* group retrotransposons in eukaryote genomes. *Genetica* **100**:185-195.
- FRANCK, G., and V. L. S. VALENTE. 1985. Study on the fluctuation in *Drosophila* populations of Bento Gonçalves, RS, Brazil. *Revta. bras. Biol.* **45**:133-141.
- FREUND, R., and M. MESELSON. 1984. Long terminal repeat nucleotide sequence and specific insertion of the *gypsy* transposon. *Proc. Natl. Acad. Sci. USA* **81**:4462-4464.
- FRIESEN, P. D., and M. S. NISSEN. 1990. Gene organization and transcription of TED, a lepidopteran retrotransposon integrated within the baculovirus genome. *Mol. Cell. Biol.* **10**:3067-3077.
- GAUDIERI, S., C. LEELAYUWAT, D. C. TOWNEND, J. K. KULSKI, and R. L. DAWKINS. 1997. Genomic characterization of the region between HLA-B and TNF: implications for the evolution of multicopy gene families. *J. Mol. Evol.* **44**:S147-154.
- GENTILE, K. L., W. D. BURKE, and T. H. EICKBUSH. 2001. Multiple lineages of R1 retrotransposable elements can coexist in the rDNA loci of *Drosophila*. *Mol. Biol. Evol.* **18**:235-245.
- GEORGIEV, G. P., Y. V. ILYIN, V. G. CHMELIAUSKAITE, A. P. RYSKOV, D. A. KRAMEROV, K. G. SKRYABIN, A. S. KRAYEV, E. M. LUKANIDIN, and M. S. GRIGORYAN. 1981. Mobile dispersed genetic elements and other middle repetitive DNA sequences in the genomes of *Drosophila* and mouse:

transcription and biological significance. Cold Spring Harbor Symp. Quant. Biol. **45**:641-654.

GINZBURG, L. R., P. M. BINGHAM, and S. YOO. 1984. On the theory of speciation induced by transposable elements. Genetics **107**:331-341.

GONZALEZ, P., and H. A. LESSIOS. 1999. Evolution of sea urchin retroviral-like (SURL) elements: evidence from 40 echinoid species. Mol. Biol. Evol. **16**:938-952.

GRANDBASTIEN, M.-A. 1998. Activation of plant retrotransposons under stress conditions. Trends Plant Sci. **3**:181-187.

GREGORY, T. R., and P. D. HEBERT. 1999. The modulation of DNA content: proximate causes and ultimate consequences. Genome Res. **9**:317-324.

GRIMALDI, D. A. 1990. A phylogenetic, revised classification of genera in the Drosophilidae (Diptera). Bull. Am. Mus. Nat. Hist. **197**:1-139.

HAGEMANN, S., E. HARING, and W. PINSKER. 1996. Repeated horizontal transfer of *P* transposons between *Scaptomyza pallida* and *Drosophila bifasciata*. Genetica **98**:43-51.

HANDLER, A. M. 2001. A current perspective on insect gene transformation. Insect Biochem. Mol. Biol. **31**:111-128.

HARTL, D. L., E. R. LOZOVSKAYA, D. I. NURMINSKY, and A. R. LOHE. 1997. What restricts the activity of *mariner*-like transposable elements? Trends Genet. **13**:197-201.

HIGGINS, D. G., A. J. BLEASBY, and R. FUCHS. 1992. CLUSTAL V: improved software for multiple sequence alignment. Computer Applications in the Biosciences (CABIOS). **8**:189-191.

HOCHSTENBACH, R., H. HARHANGI, K. SCHOUREN, P. BINDELS, R. SUIJKERBUIJK, and W. HENNIG. 1996. Transcription of gypsy elements in a Y-chromosome male fertility gene of *D. hydei*. Genetics **142**:437-446.

- HOUCK, M. A., J. B. CLARK, K. R. PETERSON, and M. G. KIDWELL. 1991. Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. *Science* **253**:1125-1128.
- HURST, G. D. D., and M. SCHILTHUIZEN. 1998. Selfish genetic elements and speciation. *Heredity* **80**:2-8.
- HURST, G. D. D., and J. H. WERREN. 2001. The role of selfish genetic elements in eukaryotic evolution. *Nature* **2**:597-606.
- ILYIN, Y. V., N. V. LYUBOMIRSKAYA, and A. I. KIM. 1991. Retrotransposon Gypsy and genetic instability in *Drosophila* (review). *Genetica* **85**:13-22.
- JORDAN, I. K., and J. F. McDONALD. 1998. Evolution of the *copia* retrotransposon in the *Drosophila melanogaster* species subgroup. *Mol. Biol. Evol.* **15**:1160-1171.
- JORDAN, I. K., L. V. MATYUNINA, and J. F. McDONALD. 1999. Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. *Proc. Natl. Acad. Sci. USA* **96**:12621-12625.
- JOWETT, T. 1986. Preparation of nucleic acids. Pp. 275-286 in Roberts D.B. (Ed.), *Drosophila: a Practical Approach*, IRL Press, Washington, DC.
- JUNAKOVIC, N., C. D. FRANCO, P. BARSANTI, and G. PALUMBO. 1986. Transposition of copia-like nomadic elements can be induced by heat shock. *J. Mol. Evol.* **24**:89-93.
- JUNAKOVIC, N., A. TERRINONI, C. DI FRANCO, C. VIEIRA, and C. LOEVENBRUCK. 1998. Accumulation of transposable elements in the heterochromatin and on the Y chromosome of *Drosophila simulans* and *Drosophila melanogaster*. *J. Molec. Evol.* **46**:661-668.
- KASTRITSIS, C. D. 1969. The chromosomes of some species of the *guarani* group of *Drosophila*. *J. Hered.* **60**:51- 57.

- KIDWELL, M. G., and J. F. KIDWELL. 1976. Selection for male recombination in *Drosophila melanogaster*. *Genetics* **84**:333-351.
- KIDWELL, M. G. 1993. Lateral transfer in natural populations of eukaryotes. *Annu. Rev. Genet.* **27**:235-256.
- KIDWELL, M. G., and D. LISCH. 1997. Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci. USA* **94**:7704-7711.
- KIDWELL, M. G., and D. R. LISCH. 2000. Transposable elements and host genome evolution. *Trends Ecol. Evol.* **15**:95-99.
- KIDWELL, M. G., and D. R. LISCH. 2001. Perspective: transposable elements, parasitic DNA, and genome evolution. *Evolution* **55**:1-24.
- KIM, A. I., E. S. BELYAEVA, and M. M. ASLANIAN. 1990. Autonomous transposition of *gypsy* mobile elements and genetic instability in *Drosophila melanogaster*. *Mol. Gen. Genet.* **224**:303-308.
- KIM, A., C. TERZIAN, P. SANTAMARIA, A. PÉLISSON, N. PRUD'HOMME, and A. BUCHETON. 1994. Retroviruses in invertebrates: the *gypsy* retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **91**:1285-1289.
- KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111-120.
- KING, J. C. 1947. A comparative analysis of the chromosomes of the *guarani* group of *Drosophila*. *Evolution* **1**:48-62.
- KUMAR, A., S. R. PEARCE, K. McLEAN, G. HARRISON, J. S.(P.) HESLOP-HARRISON, R. WAUGH, and A. J. FLAVELL. 1997. The *Ty1-copia* group of retrotransposons in plants: genomic organisation, evolution, and use as molecular markers. *Genetica* **100**:205-217.

- KUMAR, S., K. TAMURA, I. B. JAKOBSEN, and M. NEI. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*.
- KUZIN, A. B., N. V. LYUBOMIRSKAYA, B. M. KHUDAIBERGENOVA, Y. V. ILYIN, and A. I. KIM. 1994. Precise excision of the retrotransposon *gypsy* from the *forked* and *cut* loci in a genetically unstable *D. melanogaster* strain. *Nucleic Acids Res.* **22**:4641-4645.
- KWIATOWSKI, J., D. SKARECKY, K. BAILEY, and F. J. AYALA. 1994. Phylogeny of *Drosophila* and related genera inferred from the nucleotide sequence of the Cu, Zn Sod gene. *J. Mol. Evol.* **38**:443-454.
- KWIATOWSKI, J., and F. J. AYALA. 1999. Phylogeny of *Drosophila* and related genera: conflict between molecular and anatomical analyses. *Mol. Phylogenet. Evol.* **13**:319-328.
- LAMBERTSSON, A., S. ANDERSSON, and T. JOHANSSON. 1989. Cloning and characterization of variable-sized *gypsy* mobile elements in *Drosophila melanogaster*. *Plasmid* **22**:22-31.
- LAWRENCE, J. G., and D. L. HARTL. 1992. Inference of horizontal genetic transfer from molecular data: an approach using the bootstrap. *Genetics* **131**:753-760.
- LAWRENCE, J. G., and H. OCHMAN. 1998. Molecular archaeology of the *Escherichia coli* genome. *Proc. Natl. Acad. Sci. USA* **95**:9413-9417.
- LÉCHER, P., A. BUCHETON, and A. PÉLISSON. 1997. Expression of the *Drosophila* retrovirus *gypsy* as ultrastructurally detectable particles in the ovaries of flies carrying a permissive *flamenco* allele. *J. Gen. Virol.* **78**:2379-2388.
- LERAT, E., and P. CAPY. 1999. Retrotransposons and retroviruses: analysis of the envelope gene. *Mol. Biol. Evol.* **16**:1198-1207.
- LERAT, E., F. BRUNET, C. BAZIN, and P. CAPY. 1999. Is the evolution of transposable elements modular? *Genetica* **107**:15-25.

- LI, T.-H., J. SPEAROW, C. M. RUBIN, and C. W. SCHMID. 1999. Physiological stresses increase mouse short interspersed element (SINE) RNA expression in vivo. *Gene* **239**:367-372.
- LIM, J. K. 1988. Intrachromosomal rearrangements mediated by *hobo* transposons in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **85**:9153-9157.
- LIN, Y. J., L. SEROUDE, and S. BENZER. 1998. Extended life-span and stress resistance in the *Drosophila* mutant *methuselah*. *Science* **282**:943-946.
- LOHE, A. R., E. N. MORIYAMA, D. LIDHOLM, and D. L. HARTL. 1995. Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. *Mol. Biol. Evol.* **12**:62-72.
- LORETO, E. L. Da S., L. B. Da SILVA, A. ZAHA, and V. L. Da S. VALENTE. 1998a. Distribution of transposable elements in neotropical species of *Drosophila*. *Genetica* **101**:153-165.
- LORETO, E. L. Da S., A. ZAHA, and V. L. Da S. VALENTE. 1998b. Transposable elements in South American populations of *Drosophila simulans*. *Genet. Sel. Evol.* **30**:171-180.
- LORETO, E. L. Da S., V. L. Da S. VALENTE, A. ZAHA, J. C. SILVA, and M. G. KIDWELL. 2001. *Drosophila mediopunctata* P elements: a new example of horizontal transfer. *J. Heredity* **92**:375-381.
- LYUBOMIRSKAYA, N. V., S. N. AVEDISOV, S. A. SURKOV, and Y. V. ILYIN. 1993. Two *Drosophila* retrotransposon *gypsy* subfamilies differ in ability to produce new DNA copies via reverse transcription in *Drosophila* cultured cells. *Nucleic Acids Res.* **21**:3265-3268.
- MARÍN, I., and C. LLORÉNS. 2000. *Ty31Gypsy* retrotransposons: description of new *Arabidopsis thaliana* elements and evolutionary perspectives derived from comparative genomic data. *Mol. Biol. Evol.* **17**:1040-1049.

- MARQUES, E. K., M. NAPP, H. WINGE, and A. R. CORDEIRO. 1966. A corn meal, soybean flour, wheat germ medium for *Drosophila*. D. I. S. **41**:187.
- MARUYAMA, K., and D. L. HARTL. 1991. Evidence for interspecific transfer of the transposable element mariner between *Drosophila* and *Zaprionus*. J. Mol. Evol. **33**:514-524.
- McCLINTOCK, B. 1987. The discovery and characterization of transposable elements. Garland, New York.
- MILLER, K., C. LYNCH, J. MARTIN, E. HERNIOU, and M. TRISTEM. 1999. Identification of multiple gypsy LTR-retrotransposon lineages in vertebrate genomes. J. Mol. Evol. **49**:358-366.
- MIZROKHI, L. J., and A. M. MAZO. 1991. Cloning and analysis of the mobile element gypsy from *D. virilis*. Nucleic Acids Res. **19**: 913-916.
- MODOLELL, J., W. BENDER, and M. MESELSON. 1983. *Drosophila melanogaster* mutations suppressible by the suppressor of Hairy-wing are insertions of a 7.3-kilobase mobile element. Proc. Natl. Acad. Sci. USA **80**:1678-1682.
- MONTCHAMP-MOREAU, C., S. RONSSERAY, M. JACQUES, M. LEHMANN, and D. ANXOLABÉHÈRE. 1993. Distribution and conservation of sequences homologous to the 1731 retrotransposon in *Drosophila*. Mol. Biol. Evol. **10**:791-803.
- MOORE, J. K., and J. E. HABER. 1996. Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. Nature **383**:644-646.
- NEI, M., and T. GOJOBORI. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3**:418-426.
- NEKRUTENKO, A., and W.-H. LI. 2001. Transposable elements are found in a large number of human protein-coding genes. Trends Genet. **17**:619-621.

- NELSON, K. E. et al. 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**:323-329.
- NICHOLAS, K. B., and H. B. Jr. NICHOLAS. 1997. GeneDoc: A tool for editing and annotating multiple sequences alignments. WWW.psc.edu/biomed/genedoc.
- OKADA, N., and M. HAMADA. 1997. The 3' ends of tRNA-derived SINEs originated from the 3' ends of LINEs: a new example from the bovine genome. *J. Mol. Evol.* **44**:s52-s56.
- PARDUE, M.-L., and P. G. DeBARYSHE. 1999. Telomeres and telomerase: more than the end of the line. *Chromosoma* **108**:73-82.
- PEIFER, M., and W. BENDER. 1988. Sequences of the *gypsy* transposon of *Drosophila* necessary for its effects on adjacent genes. *Proc. Natl. Acad. Sci. USA* **85**:9650-9654.
- PELANDAKIS, M., and M. SOLIGNAC. 1993. Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. *J. Mol. Evol.* **37**:525-543.
- PÉLISSON, A., S. U. SONG, N. PRUD'HOMME, P. A. SMITH, A. BUCHETON, and V. G. CORCES. 1994. *Gypsy* transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the *Drosophila flamenco* gene. *EMBO J.* **13**:4401-4411.
- PÉLISSON, A., L. TEYSSET, F. CHALVET, A. KIM, N. PRUD'HOMME, C. TERZIAN, and A. BUCHETON. 1997. About the origin of retroviruses and the co-evolution of the *gypsy* retrovirus with the *Drosophila flamenco* host gene. *Genetica* **100**:29-37.
- PETROV, D. A. 2001. Evolution of genome size: new approaches to an old problem. *Trends Genet.* **17**:23-28.

- PICKERAL, O. K., W. MAKAOWSKI, M. S. BOGUSKI, and J. D. BOEKE. 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. *Genome Res.* **10**:411-415.
- PINSKER, W., E. HARING, S. HAGEMANN, and W. J. MILLER. 2001. The evolutionary life history of *P* transposons: from horizontal invaders to domesticated neogenes. *Chromosoma* **110**:148-158.
- PLASTERK, R. H. A., Z. IZSVÁK, and Z. IVICS. 1999. Resident aliens the Tc1/*mariner* superfamily of transposable elements. *Trends Genet.* **15**:326-332.
- PLASTERK, R. H. A., and R. F. KETTING. 2000. The silence of the genes. *Curr. Opin. Genet. Dev.* **10**:562-567.
- POWELL, J. R., and R. DeSALLE. 1995. *Drosophila* molecular phylogenies and their uses. *Evol. Biol.* **28**:87-138.
- PRUD'HOMME, N., M. GANS, M. MASSON, C. TERZIAN, and A. BUCHETON. 1995. *Flamenco*, a gene controlling the gypsy retrovirus of *Drosophila melanogaster*. *Genetics* **139**:697-711.
- REGNER, L. P., M. S. O. PEREIRA, C. E. V. ALONSO, E. ABDELHAY, and V. L. S. VALENTE. 1996. Genomic distribution of *P* elements in *Drosophila willistoni* and a search for their relationship with chromosomal inversions. *J. Heredity* **87**:191-198.
- ROBERTSON, H. M., and E. G. MacLEOD. 1993. Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. *Insect Mol. Biol.* **2**:125-139.
- ROBERTSON, H. M., and D. J. LAMPE. 1995. Recent horizontal transfer of a *mariner* transposable element among and between Diptera and Neuroptera. *Mol. Biol. Evol.* **12**:850-862.
- RUSSO, C. A. M., N. TAKEZAKI, and M. NEI. 1995. Molecular phylogeny and divergence times of Drosophilid species. *Mol. Biol. Evol.* **12**:391-404.

- SAAVEDRA, C. C. R., S. M. CALLEGARI-JACQUES, M. NAPP, and V. L. S. VALENTE. 1995. A descriptive and analytical study of four neotropical drosophilid communities. *J. Zool. Syst. Evol. Res.* **33**:62-74.
- SALZANO, F. M. 1955. Chromosomal polymorphism in two species of the guarani group of *Drosophila*. *Chromosoma* **7**:39-50.
- SASSI, A. 2000. Contribuição ao estudo dos elementos transponíveis em populações de *Drosophila willistoni*. Dissertação de Bacharelado, UFRGS, Brasil.
- SHARP, P. M., and W.-H. LI. 1989. On the rate of DNA sequence evolution in *Drosophila*. *J. Mol. Evol.* **28**:398-402.
- SHERMAN, A., A. DAWSON, C. MATHER, H. GILHOOLEY, Y. LI, R. MITCHELL, D. FINNEGAN, and H. SANG. 1998. Transposition of the *Drosophila* element *mariner* into the chicken germ line. *Nature Biotech.* **16**:1050-1053.
- SHIELDS, D. C., P. M. SHARP, D. G. HIGGINS, and F. WRIGHT. 1988. Silent sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. *Mol. Biol. Evol.* **5**:704-716.
- SHRIMPTON, A. E., T. F. C. MACKAY, and A. J. L. BROWN. 1990. Transposable element-induced response to artificial selection in *Drosophila melanogaster*: molecular analysis of selection lines. *Genetics* **125**:803-811.
- SILVA, J. C., and M. G. KIDWELL. 2000. Horizontal transfer and selection in the evolution of *P* elements. *Mol. Biol. Evol.* **17**:1542-1557.
- SIMMONS, G. M. 1992. Horizontal transfer of hobo transposable elements within the *Drosophila melanogaster* species complex: evidence from DNA sequencing. *Mol. Biol. Evol.* **9**:1050-1060.
- SMIT, A. F. A., and A. D. RIGGS. 1996. *Tiggers* and other DNA transposon fossils in the human genome. *Proc. Natl. Acad. Sci. USA* **93**:1443-1448.

- SMIT, A. F. A. 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.* **9**:657-663.
- SMITH, P. A., and V. G. CORCES. 1992. The suppressor of hairy-wing binding region is required for gypsy mutagenesis. *Mol. Gen. Genet.* **233**:65-70.
- SONG, S. U., T. GERASIMOVA, M. KURKULOS, J. D. BOEKE, and V. G. CORCES. 1994. An *env*-like protein encoded by a *Drosophila* retroelement: evidence that gypsy is an infectious retrovirus. *Genes Dev.* **8**:2046-2057.
- SPRINGER, M. S., E. H. DAVIDSON, and R. J. BRITTEN. 1991. Retroviral-like element in a marine invertebrate. *Proc. Natl. Acad. Sci. USA* **88**:8401-8404.
- STACEY, S. N., R. A. LANSMAN, H. W. BROCK, and T. A. GRIGLIATTI. 1986. Distribution and conservation of mobile elements in the genus *Drosophila*. *Mol. Biol. Evol.* **3**:522-534.
- STRAND, D. J., and J. F. McDONALD. 1985. *Copia* is transcriptionally responsive to environmental stress. *Nucleic Acids Res.* **13**:4401-4410.
- STUART-ROGERS, C., and A. J. FLAVELL. 2001. The evolution of Ty1-*copia* group retrotransposons in gymnosperms. *Mol. Biol. Evol.* **18**:155-163.
- TENG, S. C., B. KIM, and A. GABRIEL. 1996. Retrotransposon reverse-transcriptase-mediated repair of chromosomal breaks. *Nature* **383**:641-644.
- TERZIAN, C., C. FERRAZ, J. DEMAILLE, and A. BUCHETON. 2000. Evolution of the gypsy endogenous retrovirus in the *Drosophila melanogaster* subgroup. *Mol. Biol. Evol.* **17**:908-914.
- THROCKMORTON, L. 1975. The phylogeny, ecology, and geography of *Drosophila*. Pp. 421-469 in R. C. KING, ed. *Handbook of genetics*. Vol. 3. Plenum, New York.
- TSUBOTA, S. I., D. ROSENBERG, H. SZOSTAK, D. RUBIN, and P. SCHEDL. 1989. The cloning of the *Bar* region and the *B* breakpoint in *Drosophila melanogaster*: evidence for a transposon-induced rearrangement. *Genetics* **122**:881-890.

- VALENTE, V. L. S., and A. M. ARAÚJO. 1991. Ecological aspects of *Drosophila* species in two contrasting environments in southern Brazil (Diptera, Drosophilidae). *Revta. bras. Ent.* **35**:237-253.
- VALIATI, V. H. 1999. Genética de populações marginais de *Drosophila willistoni* e *D. paulistorum* e avaliação de suas potencialidades para produção de híbridos interespecíficos. Tese de Doutorado, UFRGS, Brasil.
- VÁZQUEZ-MANRIQUE, R. P., M. HERNÁNDEZ, M. J. MARTÍNEZ-SEBASTIÁN, and R. DE FRUTOS. 2000. Evolution of gypsy endogenous retrovirus in the *Drosophila obscura* species group. *Mol. Biol. Evol.* **17**:1185-1193.
- VIEIRA, C., D. LEPETIT, S. DUMONT, and C. BIÉMONT. 1999. Wake up of transposable elements following *Drosophila simulans* worldwide colonization. *Mol. Biol. Evol.* **16**:1251-1255.
- VILELA, C. R. 1999. Is *Zaprionus indianus* Gupta, 1970 (Diptera, Drosophilidae) currently colonizing the Neotropical region? *D. I. S.* **82**:37-39.
- VILLARREAL, L. P. 1997. On viruses, sex, and motherhood. *J. Virol.* **71**:859-865.
- WESSLER, S. 1996. Plant retrotransposons: turned on by stress. *Curr. Biol.* **6**:959-961.
- WHEELER, M. R. 1981. The Drosophilidae: a taxonomic overview. Pp. 1-97 in M. ASHBURNER, H. L. CARSON, and J. N. THOMPSON, eds. *The genetics and biology of Drosophila*. Vol. 3a. Academic Press, New York.
- WHITE, S. E., L. F. HABERA, and S. R. WESSLER. 1994. Retrotransposons in the flanking regions of normal plant genes: a role for copia-like elements in the evolution of gene structure and expression. *Proc. Natl. Acad. Sci. USA* **91**:11792-11796.
- WILDER, J., and H. HOLLOCHER. 2001. Mobile elements and the genesis of microsatellites in dipterans. *Mol. Biol. Evol.* **18**:384-392.

- WOLF, Y. I., L. ARAVIND, and E. V. KOONIN. 1999. Rickettsiae and Chlamydiae: evidence of horizontal gene transfer and gene exchange. *Trends Genet.* **15**:173-175.
- WRIGHT, S., and D. FINNEGAN. 2001. Genome evolution: sex and the transposable element. *Curr. Biol.* **11**:R296-R299.
- XIONG, Y., and T. H. EICKBUSH. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* **9**:3353-3362.
- YANG, Z., D. BOFFELLI, N. BOONMARK, K. SCHWARTZ, and R. LAWN. 1998. Apolipoprotein(a) gene enhancer resides within a LINE element. *J. Biol. Chem.* **273**:891-897.
- ZELENTSOVA, H., H. POLUECTOVA, L. MNJOIAN, G. LYOZIN, V. VELEIKODVORSKAJA, L. ZHIVOTOVSKY, M. G. KIDWELL, and M. B. EVGEN'EV. 1999. Distribution and evolution of mobile elements in the *virilis* species group of *Drosophila*. *Chromosoma* **108**:443-456.

ANEXOS

CAPÍTULO 1

NÍVEIS E CRITÉRIOS DE CLASSIFICAÇÃO

CLASSE Intermediário da transposição	CLASSE I Retrotransposons intermediário de RNA	CLASSE II Transposons intermediário de DNA	CLASSE III Mecanismo desconhecido de transposição
SUBCLASSE presença ou ausência de assinaturas, motivos ou seqüências	DDE Integrase/LTRs sim não SUBCLASSE I: Retrotransposons com LTRs SUBCLASSE II: Retrotransposons sem LTRs	DDE/DDD transposase sim não SUBCLASSE I SUBCLASSE II	
SUPERFAMÍLIA nível de similaridade, existência de motivos e assinaturas em comum; ordem de genes	Organização da ORF2 PR-RT-RH-IN / PR-IN-RT-RH Iniciador de t-RNA Env funcional sim / não	Transcriptase reversa sim / não	Tamanho das repetições terminais curtas / longas
FAMÍLIA nível de similaridade em motivos, assinaturas e seqüência nucleotídica 30 - 50% de similaridade	Gypsy-like Ty3 Tf1 Ty1-copia Ty1 TART/Het-A Jockey ALU S1 B1 B2	Mariner-Tc1 Pogo-Fot1 mariner Impala TBE1 ISa IS4 IS6 ISb IS1 IS91 IS2 hAT Hobo Ac Tam3 Slide P P MuDR MuDR CACTA Spm Tam1 Tgm1 Tx Tx1D	MITEs Folkback Tourist Stowaway Castaway Foldback TU
SUBFAMÍLIA nível de similaridade 50 - 70% de similaridade			

Figura 1. Esquema geral da classificação dos elementos transponíveis. Adaptado de Capy et al. (1998). A posição de *gypsy* encontra-se destacada.

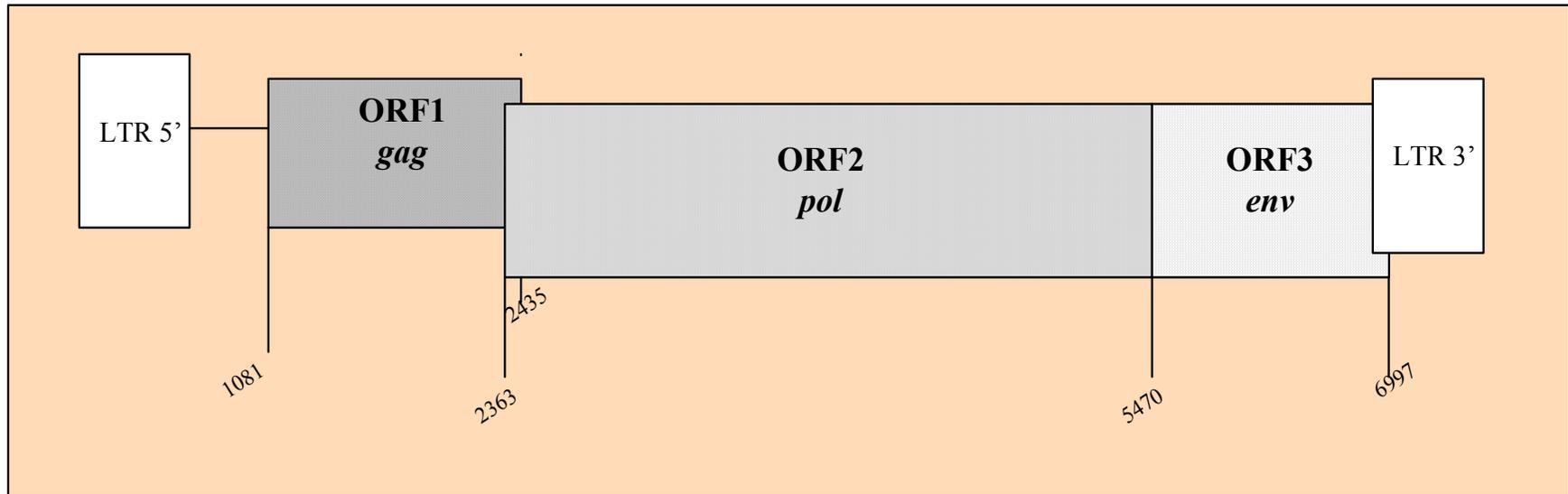


Figura 2. Esquema da estrutura do retroelemento gypsy de *D. melanogaster*. Os números indicam a posição dos nucleotídeos que limitam as ORFs.

CAPÍTULO 2

Tabela 1. Espécies utilizadas neste trabalho.

GÊNERO	SUBGÊNERO	GRUPO	ESPÉCIE	PROCEDÊNCIA E ANO
<i>Drosophila</i>	<i>Drosophila</i>	<i>guarani</i>	D. maculifrons	Morro Santana/RS- 1995
			<i>D. griseolineata</i>	Morro Santana/RS- 1995
			<i>D. ornatifrons</i>	Bento Gonçalves/RS- 1996
		<i>cardini</i>	<i>D. polymorpha</i>	Florianópolis/SC- 1997 (FLO)
				Morro Santana/RS- 1998 (MS) *
				Parque Estadual Florestal do Turvo/RS- 1998 (TU) *
			<i>D. cardinoides</i>	Morro Santana/RS- 1996
			<i>D. neocardini</i>	Florianópolis/SC- 1997
		<i>Immigrans</i>	<i>D. immigrans</i>	Bagé/RS-1995 (BA)
				Florianópolis/SC- 1998 (FLO) *
		<i>pallidipennis</i>		Morro Santana/RS- 1998 (MS) *
				Itapuã/RS-1995
		<i>tripunctata</i>	<i>D. pallidipennis</i>	Itapuã/RS-1995
			<i>D. bandeirantorum</i>	Morro Santana/RS-1996 (MS)
				Parque Estadual Florestal do Turvo/RS- 1998 (TU) *
			<i>D. angustibucca</i>	Itapuã/RS-1995
			<i>D. mediopunctata</i>	Morro Santana/RS-1996
			<i>D. mediosignata</i>	Sapucaia/RS-1996
			<i>D. mediopicta</i>	Itapuã/RS-1998
			<i>D. mediotriata</i>	Porto Alegre/RS-1996
			<i>D. virilis</i>	Bowling Green Center ** #
			<i>virilis</i> <i>repleta</i>	<i>D. hydei</i>
		<i>D. zotti</i>		Morro Santana/RS-1995
<i>D. mercatorum</i>	Dois Lajeados/RS-1995 (DL)			
	Morro Santana/RS- 1998 (MS) *			
<i>Sophophora</i>	<i>bromelioides</i>	<i>D. bromelioides</i>	Florianópolis/SC- 1999	
		<i>D. annulimana</i>	Florianópolis/SC-2000	
	<i>willistoni</i>	<i>D. willistoni</i>	Morro Santana/RS-1996	
			Belém/PA- 1997 (PA)***	
		<i>D. nebulosa</i>	Morro Santana/RS-1995	
	<i>D. fumipennis</i>	Florianópolis/SC-1998		
	<i>D. paulistorum</i>	Morro Santana/RS-1996		

		<i>D. bocainensis</i>	Turvo/RS-1998
		<i>D. capricorni</i>	Florianópolis/SC- 1997
		<i>D. equinoxialis</i>	Tefé/AM **
		<i>D. tropicalis</i>	Palma/GO **
		<i>D. sucinea</i>	México- 1998
	<i>saltans</i>	<i>D. insularis</i>	St. Kitts/Pequenas Antilhas **
		<i>D. prosaltans</i>	Morro Santana/RS-1995
		<i>D. sturtevanti</i>	Maquiné/RS-1995 (MA) Genetics Foundation ** (COL) * Mirassol ** (BRA) *
			Santana do Riacho/MG- 1995 (I27) *
			S. J. do Rio Preto/SP- 1997 (RP) *
			Azapan/México- 1998 (APA) *
			Matlapa/México- 1998 (MAT) *
	<i>melanogaster</i>	<i>D. kikkawai</i>	Porto Alegre/RS **
		<i>D. melanogaster</i>	Harwich ** #
		<i>D. ananassae</i>	Florianópolis/SC-1999
		<i>D. malerkotliana</i>	Florianópolis/SC-1997
		<i>D. simulans</i>	Florianópolis/SC-1997 (FLO) Morro Santana/RS- 1998 (MS) *
			Eldorado do Sul/RS- 1998 (ELD) *
	<i>Scaptodrosophila</i>	<i>D. lebanonensis</i>	Bowling Green Center **
		<i>D. galloi</i>	São J. do Rio Preto/SP **
		<i>D. latifasciaeformis</i>	S. J. Rio Preto/SP-1999
	<i>Dorsilopha</i>	<i>D. busckii</i>	Santa Maria/RS-1996 (SM) *
<i>Zaprionus</i>	<i>Zaprionus</i>	<i>armatus</i>	Florianópolis/SC- 1998 (FLO) *
		<i>Z. indianus</i>	Florianópolis/SC-2000

* Populações analisadas apenas por *Southern blot* para a enzima *HindIII*.

** Estoques mantidos há mais de 7 anos em laboratório.

*** População utilizada apenas na análise filogenética.

Espécies utilizadas como controle.

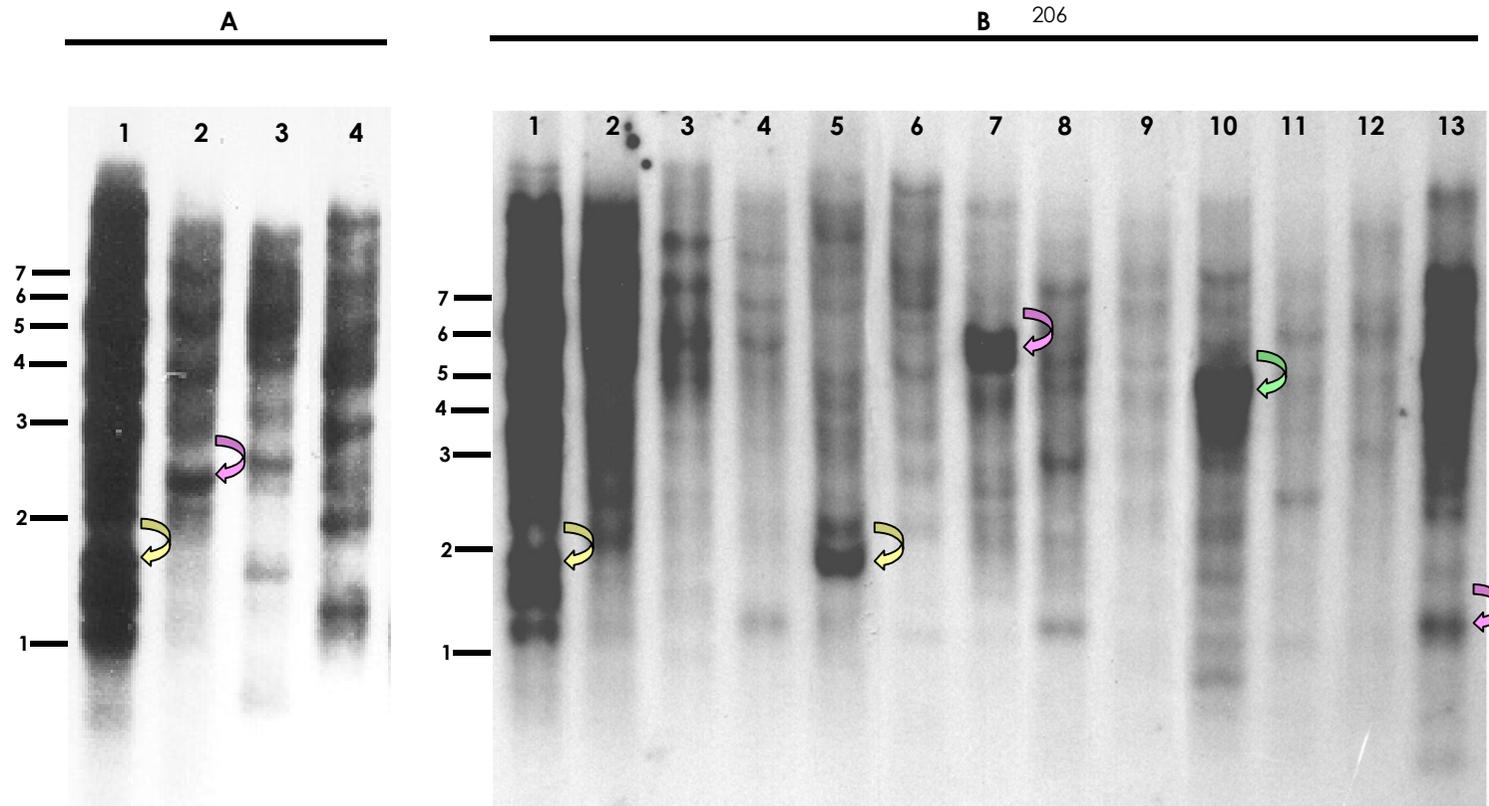


Figura 2. Southern blot de várias espécies do gênero *Drosophila*. O DNA genômico destas espécies foi digerido com a enzima de restrição *Hind*III e hibridizado com a sonda do elemento *gypsy* de *D. melanogaster*. A numeração representa as seguintes espécies em A: 1) *D. melanogaster* - Harwich; 2) *D. griseolineata*; 3) *D. kikkawai*; 4) *D. maculifrons*; em B: 1) *D. melanogaster* - Harwich; 2) *D. cardinoides*; 3) *D. pallidipennis*; 4) *D. mediopunctata*; 5) *D. mediotriata*; 6) *D. nebulosa*; 7) *D. tropicalis*; 8) *D. ananassae*; 9) *D. galloi*; 10) *D. virilis*; 11) *D. malerkotliana*; 12) *D. capricorni*; 13) *D. latifasciaeformis*. Setas amarelas indicam a presença do fragmento esperado de, aproximadamente, 1,7 kb, de acordo com o mapa de restrição de *D. melanogaster*. Setas rosas salientam fragmentos considerados de sinal predominantemente forte. Setas verdes indicam a presença do fragmento esperado de, aproximadamente, 4,9 kb, de acordo com o mapa de restrição de *D. virilis*. Os traços indicam a posição dos fragmentos do marcador de peso molecular utilizado, em kb.

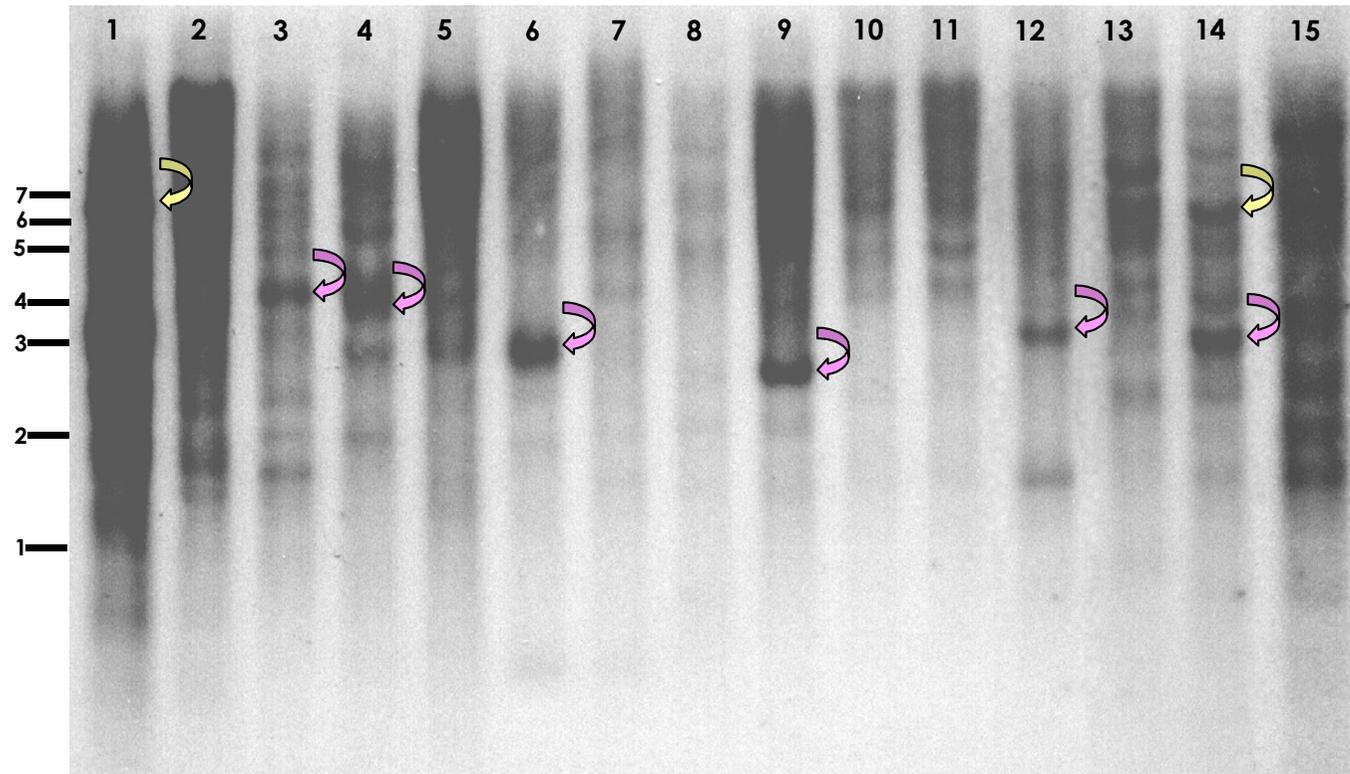


Figura 3. Southern blot de várias espécies do gênero *Drosophila*. O DNA genômico destas espécies foi digerido com a enzima de restrição *Bgl*III e hibridizado com a sonda do elemento *gypsy* de *D. melanogaster*. A numeração representa as seguintes espécies: 1) *D. melanogaster* - Harwich; 2) *D. cardinoides*; 3) *D. pallidipennis*; 4) *D. mediopunctata*; 5) *D. mediotriata*; 6) *D. nebulosa*; 7) *D. tropicalis*; 8) *D. sucinea*; 9) *D. ananassae*; 10) *D. lebanonensis*; 11) *D. galloi*; 12) *D. malerkotliana*; 13) *D. polymorpha*; 14) *D. griseolineata*; 15) *D. neocardini*. Setas amarelas indicam a presença do fragmento esperado de, aproximadamente, 7 kb, de acordo com o mapa de restrição de *D. melanogaster*. Setas rosas salientam fragmentos considerados de sinal predominantemente forte. Os traços indicam a posição dos fragmentos do marcador de peso molecular utilizado, em kb.

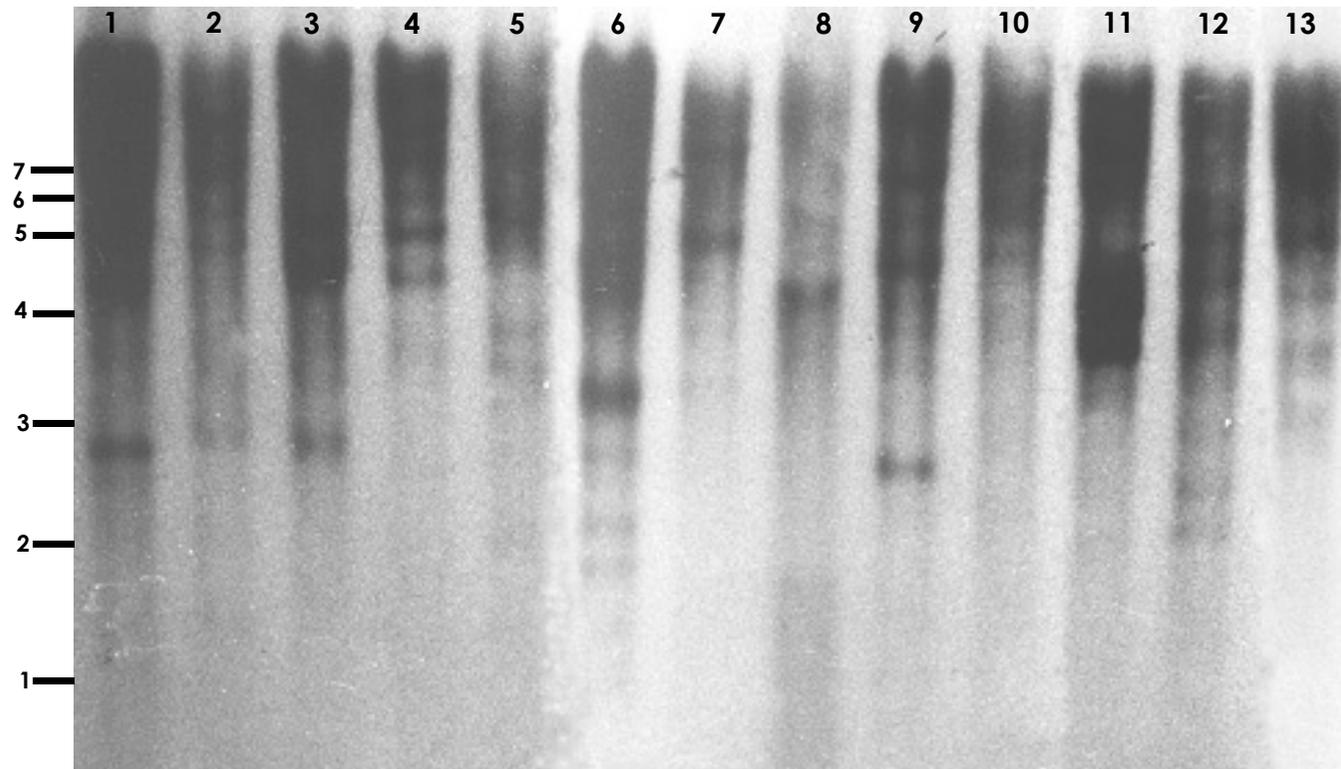


Figura 4. Southern blot de várias espécies do gênero *Drosophila*. O DNA genômico destas espécies foi digerido com a enzima de restrição *Bam*HI e hibridizado com a sonda do elemento *gypsy* de *D. melanogaster*. A numeração representa as seguintes espécies: 1) *D. melanogaster* - Harwich; 2) *D. polymorpha*; 3) *D. neocardini*; 4) *D. cardinoides*; 5) *D. immigrans*; 6) *D. bandeirantorum*; 7) *D. mediopunctata*; 8) *D. mediosignata*; 9) *D. mediopicta*; 10) *mediostriata*; 11) *D. tropicalis*; 12) *D. insularis*; 13) *D. bocainensis*. Os traços indicam a posição dos fragmentos do marcador de peso molecular utilizado, em kb.

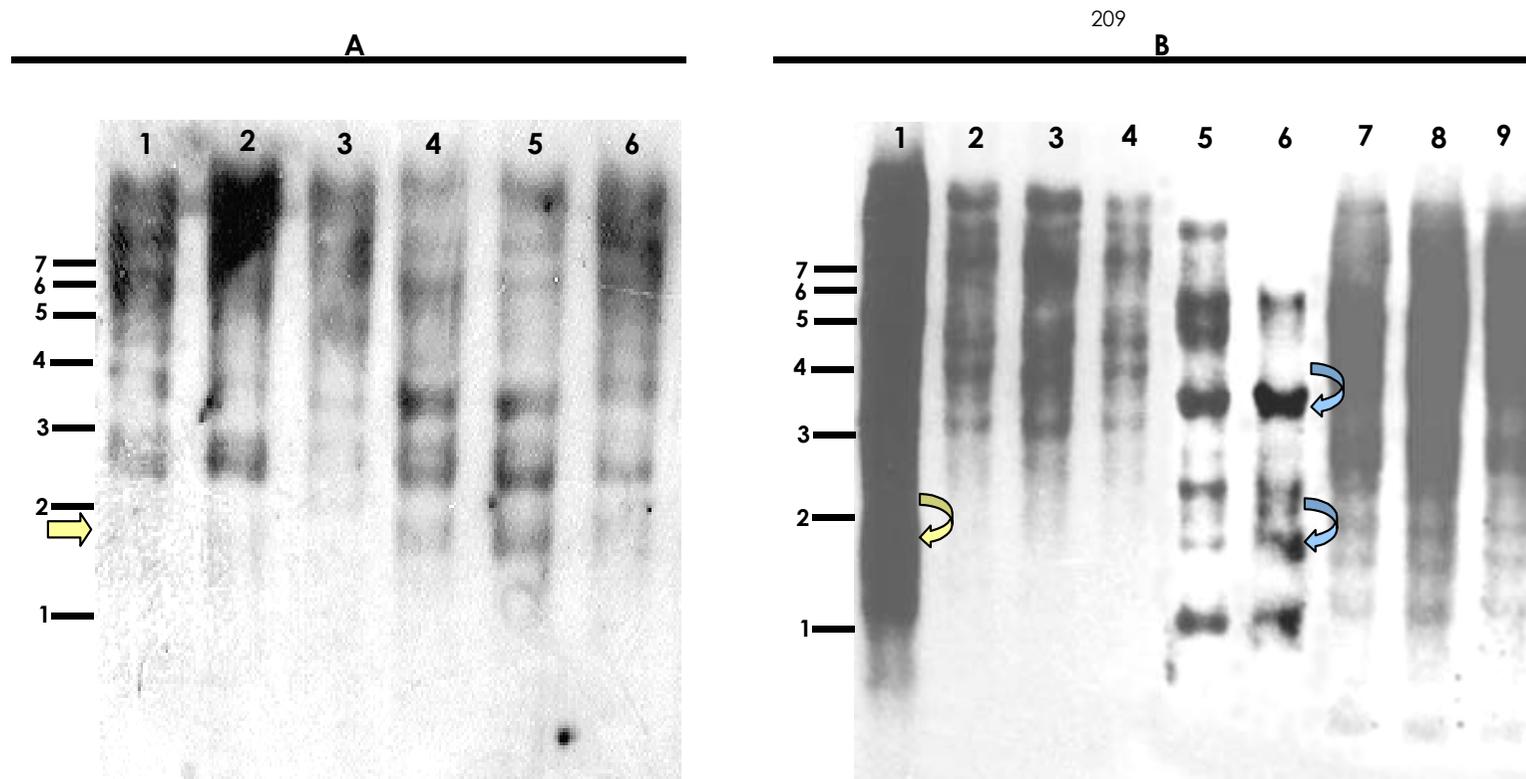


Figura 5. Southern blot de diferentes populações de espécies do gênero *Drosophila*. O DNA genômico destas populações foi digerido com a enzima de restrição *HindIII* e hibridizado com a sonda do elemento *gypsy* de *D. melanogaster*. A - populações de *D. sturtevantii*: 1) COL; 2) BRA; 3) I27; 4) RP; 5) APA E 6) MAT. B - 1) *D. melanogaster* - Harwich; 2-4) *D. polymorpha*, populações MS, SC e TU, respectivamente; 5-6) *D. busckii*, populações FLO e SM, respectivamente; 7-9) *D. mercatorum*, populações FLO, DL e MS, respectivamente. Setas amarelas indicam a presença do fragmento esperado de, aproximadamente, 1,7 kb, de acordo com o mapa de restrição de *D. melanogaster*. Setas azuis indicam a possível presença dos fragmentos esperados de, aproximadamente, 3,4 e 1,5 kb, de acordo com o mapa de restrição de *D. subobscura*. Os traços indicam a posição dos fragmentos do marcador de peso molecular utilizado, em kb.

Tabela 7. Análise Comparativa de distâncias genéticas (Kimura dois parâmetros) entre seqüências de genes nucleares (Sod) e seqüências do retroelemento *gypsy* entre espécies de *Drosophila* e *Zaprionus*

	1		2		3		4		5		6		7		8		9		10		11		
	<i>D. melanogaster</i>	<i>D. simulans</i>	<i>D. willistoni</i>	<i>D. subobscura</i>	<i>D. hydei</i>	<i>Zaprionus</i>	<i>Scaptodrosophila</i>	<i>D. virilis</i>	<i>D. busckii</i>	<i>D. paulistorum</i>	<i>D. nebulosa</i>	Sod	gyp	Sod	gyp	Sod	gyp	Sod	gyp	Sod	gyp	Sod	Gyp
1																							
2	0,03	0,12																					
3	0,27	1,04	0,26	1,17																			
4	0,22	0,26	0,20	0,37	0,27	0,96																	
5	0,31	0,23	0,33	0,36	0,28	1,04	0,36*	0,08*															
6	0,32	0,11	0,33*	0,07*	0,31	1,12	0,36	0,32	0,25	0,34													
7	0,32	0,13	0,30*	0,10*	0,32	1,15	0,34	0,32	0,33	0,30	0,32	0,14											
8	0,24	0,23	0,24	0,36	0,29	1,00	0,27	0,10	0,19*	0,06*	0,30	0,33	0,33	0,31									
9	0,28	0,21	0,27	0,32	0,26	0,97	0,27*	0,05*	0,24	0,09	0,24	0,30	0,33	0,30	0,22	0,09							
10	0,26	0,66	0,26	0,76	0,05	1,02	0,26	0,60	0,31	0,57	0,30	0,69	0,32	0,71	0,27	0,60	0,27	0,60					
11	0,24	0,59	0,22	0,66	0,17	1,12	0,28	0,60	0,27	0,54	0,29	0,58	0,37	0,63	0,28	0,56	0,26	0,58	0,18	0,11			

* Comparações importantes para as inferências dos casos de transmissão horizontal

Tabela 8. Análise Comparativa de distâncias genéticas (Kimura dois parâmetros) entre seqüências de genes nucleares (rDNA) e seqüências do retroelemento *gypsy* entre espécies de *Drosophila* e *Zaprionus*

	1		2		3		4		5		6		7		8		9		10		11	
	<i>D. melanogaster</i> rDNA	<i>D. melanogaster</i> gyp	<i>D. simulans</i> rDNA	<i>D. simulans</i> gyp	<i>D. willistoni</i> rDNA	<i>D. willistoni</i> gyp	<i>Zaprionus</i> rDNA	<i>Zaprionus</i> gyp	<i>D. busckii</i> rDNA	<i>D. busckii</i> gyp	<i>D. nebulosa</i> rDNA	<i>D. nebulosa</i> gyp	<i>D. maculifrons</i> rDNA	<i>D. maculifrons</i> gyp	<i>D. mediopictoides</i> rDNA	<i>D. mediopictoides</i> gyp	<i>D. pallidipennis</i> rDNA	<i>D. pallidipennis</i> gyp	<i>D. ornatifrons</i> rDNA	<i>D. ornatifrons</i> gyp	<i>D. latifasciaeformis</i> rDNA	<i>D. latifasciaeformis</i> gyp
1																						
2	0,01	0,12																				
3	0,07	1,04	0,08	1,17																		
4	0,07	0,11	0,08*	0,07*	0,08	1,12																
5	0,10	0,21	0,10	0,32	0,08	0,97	0,06	0,30														
6	0,06	0,59	0,07	0,66	0,04	1,12	0,08	0,58	0,08	0,58												
7	0,10	0,35	0,10	0,44	0,09	0,97	0,08	0,41	0,10	0,40	0,08	0,46										
8	0,10	*a	0,10	*b	0,08	*c	0,08	*d	0,08	*e	0,08	*f	0,02	*g								
9	0,10	0,23	0,11	0,30	0,09	1,11	0,08	0,26	0,10	0,11	0,09	0,46	0,03	0,38	0,04	*h						
10	0,09	0,34	0,10	0,45	0,10	0,91	0,05	0,40	0,09	0,35	0,09	0,48	0,04	0,13	0,05	*i	0,05	0,35				
11	0,10	0,38	0,10*	0,18*	0,08	1,14	0,12	0,14	0,12	0,30	0,07	0,63	0,11	0,41	0,10	*j	0,12	0,29	0,13	0,40		

*a- 0,21 (*D. bandeirantorum*); 0,39 (*D. mediosignata*); 0,34 (*D. mediopunctata*); 0,25 (*D. mediopicta*); 0,33 (clone ctata45); 0,40 (clone picta34)

*b- 0,31; 0,47; 0,42; 0,30; 0,41; 0,49

*c- 1,11; 1,33; 1,45; 0,91; 0,91; 1,00

*d- 0,28; 0,42; 0,39; 0,29; 0,39; 0,49

*e- 0,08; 0,35; 0,34; 0,25; 0,43; 0,39

*f- 0,54; 0,61; 0,66; 0,69; 0,53; 0,56

*g- 0,38; 0,53; 0,52; 0,38; 0,09; 0,18

h- 0,05; 0,35; 0,35; 0,32; 0,40; 0,38

*i- 0,34; 0,51; 0,52; 0,39; 0,16; 0,20

*j- 0,31; 0,41; 0,37; 0,28; 0,37; 0,45

As comparações realizadas para os ítems de a até j, referidos na tabela, seguem a mesma ordem especificada em a nos parênteses (valores acima).

* Comparações importantes para as inferências dos casos de transmissão horizontal

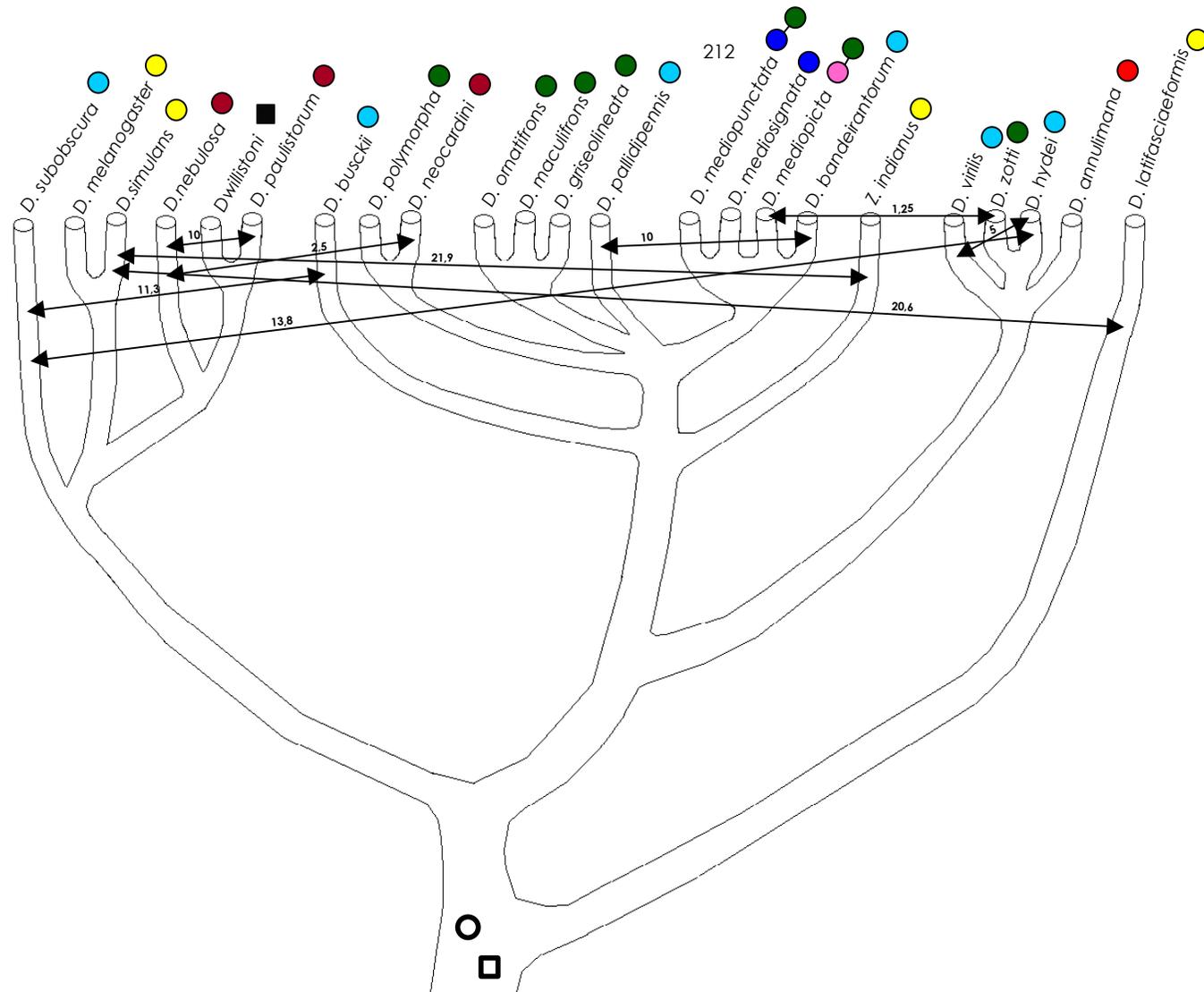


Figura 11. Esquema de uma proposta de cenário evolutivo para os múltiplos casos de transmissão horizontal, sugeridos neste trabalho, para o retroelemento *gypsy*. Círculos e Retângulos abertos correspondem às famílias ancestrais. Círculos e Retângulos fechados significam famílias atuais. As diferentes cores dos círculos representam as diferentes subfamílias propostas (subfamília A - azul claro; B - azul escuro; C - amarelo; D - rosa; E - vermelho; F - verde e G - marron). Círculos unidos representam subfamílias coexistindo na mesma espécie. As setas indicam os eventos de transmissão horizontal sugeridos. Os números acima das setas são as estimativas de tempo, em milhões de anos, da ocorrência destes diferentes eventos.

CAPÍTULO 3

