

LUCIANE MADUREIRA DE ALMEIDA

DISTRIBUIÇÃO GENÔMICA E RELAÇÕES
FILOGENÉTICAS DOS ELEMENTOS TRANSPONÍVEIS
COPIA, MICROPIA E L1 NOS EM ESPÉCIES DO GRUPO
REPLETA DE DROSOPHILA



LUCIANE MADUREIRA DE ALMEIDA

DISTRIBUIÇÃO GENÔMICA E RELAÇÕES
FILOGENÉTICAS DOS ELEMENTOS TRANSPONÍVEIS
COPIA, MICROPIA E MINOS EM ESPÉCIES DO GRUPO
REPLETA DE DROSOPHILA

**Tese apresentada para
obtenção do Título de Doutor
em Genética.**

**São José do Rio Preto - SP
2004**

UNESP



UNIVERSIDADE ESTADUAL PAULISTA

Campus de São José do Rio Preto

INSTITUTO DE BIOCIÊNCIAS, LETRAS E CIÊNCIAS EXATAS

Programa de Pós-Graduação em Genética

LUCIANE MADUREIRA DE ALMEIDA

DISTRIBUIÇÃO GENÔMICA E RELAÇÕES
FILOGENÉTICAS DOS ELEMENTOS TRANSPONÍVEIS
COPIA, MICROPIA E MINOS EM ESPÉCIES DO GRUPO
REPLETA DE DROSOPHILA



t. 1295/04

Tese apresentada para
obtenção do Título de
Doutor em Genética

Orientadora: Profa. Dra. Cláudia Márcia Aparecida Carareto

São José do Rio Preto, 2004



575

A447d

20204000

21/10/2004

Doação

R\$ 27,00

Autor

ND: 111/04 -

Proc. 047/04

Almeida, Luciane Madureira.

Distribuição genômica e relações filogenéticas dos elementos transponíveis *copia*, *micropia* e *Minos* em espécies do grupo *repleta* de *Drosophila* - São José do Rio Preto : [s.n.], 2004

134 f. : il. ; 30 cm.

Orientador: Cláudia Márcia Aparecida Carareto

Tese (doutorado) - Universidade Estadual Paulista. Instituto de Biociências, Letras e Ciências Exatas

1. *Drosophila*. 2. *Drosophila* - Grupo *repleta*. 3. *Drosophila* - Elementos transponíveis. 4. *Drosophila* - Elementos transponíveis - *Copia*. 5. *Drosophila* - Elementos transponíveis - *Micropia*. 6. *Drosophila* - Elementos transponíveis - *Minos*. 7. *Drosophila* - Elementos transponíveis - Transferência horizontal. I. Carareto, Cláudia Márcia Aparecida. II. Universidade Estadual Paulista. Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU - 575



Agradecimentos

A Profa. Dra. Cláudia Márcia Carareto, minha orientadora há vários anos, pela amizade, incentivo, pela orientação e por ter me mostrado o caminho pelo qual pretendo continuar seguindo.

Ao Prof. Dr. Carlos Roberto Ceron, por suas sugestões sempre valiosas e pela amizade.

A Profa. Dra. Paula Rahal, pela ajuda sincera, incentivo, amizade e por ter disponibilizado o sequenciador automático para realização do meu projeto. Aos alunos Fabian e Jemalva por ajudarem no seqüenciamento dos elementos transponíveis.

A Prof. Dra. Lillian Madi-Ravazi e ao sempre amigo Felipe Rafael Torres, por ter cedido e identificado importantes estoques de linhagens utilizados nessa pesquisa, pela amizade e grande incentivo.

Aos Prof. Dr. Elgion da Silva Loreto; Prof. Dr. Fábio de Melo Sene; Profs. Dra. Lillian Madi-Ravazzi e Prof. Dr. Louis Bernard Klaczko, por fazerem parte da correção deste trabalho.

O presente trabalho foi realizado no Departamento de Biologia do Instituto de Biociências, Letras e Ciências Exatas de São José do Rio Preto, SP, da Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP, sob a orientação da Profa. Dra. Cláudia Márcia Aparecida Carareto, com o auxílio financeiro da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

Aos técnicos Sebastião, Paulo e Josué, a secretária Simone pelo apoio e amizade.

Ao pessoal da Seção de Pós-graduação, especialmente a Rosana e Silvia, pela disposição e simpatia.

Agradecimentos

À **Profa. Dra. Cláudia Márcia Carareto**, minha orientadora há vários anos, pela amizade, incentivo, pela orientação e por ter me mostrado o caminho pelo qual pretendo continuar seguindo.

Ao **Prof. Dr. Carlos Roberto Ceron**, por suas sugestões sempre valiosas e pela amizade.

À **Profa. Dra. Paula Rahal**, pela ajuda sincera, incentivo, amizade e por ter disponibilizado o sequenciador automático para realização do meu projeto. Aos alunos **Fabian e Janaina** por ajudarem no seqüenciamento dos elementos transponíveis.

À **Prof. Dra. Lilian Madi-Ravazi** e ao sempre amigo **Felipe Rafael Torres**, por ter cedido e identificado importantes estoques de linhagens utilizados nessa pesquisa, pela amizade e grande incentivo.

Aos **Prof. Dr. Élgion da Silva Loreto; Prof. Dr. Fábio de Melo Sene; Profa. Dra. Lilian Madi-Ravazzi e Prof. Dr. Louis Bernard Klaczko**, por fazerem parte da correção deste trabalho.

À **Profa. Dra. Hermione Elly Melara de Campos Bicudo**, pelo incentivo e amizade.

Aos demais professores do Departamento de Biologia pela atenção e colaboração sempre que precisei.

Aos técnicos **Sebastião, Paulo e Josué**, à secretária **Simone** pelo apoio e amizade.

Ao pessoal da **Seção de Pós-graduação**, especialmente a **Rosana e Silvia**, pela disposição e simpatia.

A Fapesp, pelo incentivo financeiro.

Aos colegas da pós-graduação, Luciana, Priscilla, Andréa, Patrícia, Uliana, Rosa, Alan, Amanda, pela amizade e companheirismo.

*Eu te desejo não parar tão cedo
Pois toda idade tem prazer e medo
E para os que erram feio e bastante
Que você consiga ser tolerante*

*Quando você ficar triste
Que seja por um dia e não o ano inteiro
E que você descubra que rir é bom
Mas rir de tudo é desespero*

*Desejo que você tenha a quem amar
E quando estiver bem cansado*

Ainda exista amor para recomeçar!!!!!!

Barão Vermelho

Ao Matheuzinho e ao Teli, basta um sorriso para vocês me encham de ânimo e felicidade.

Na vida como na ciência cometemos erros que nos desanimam, nos cansam e nos deixam menos confiantes. O amor que recebo da minha família e o meu amor pela ciência é que me permitiram recomeçar diante de cada derrota.

Agradeço em especial ao meu marido, Marcelo, por me acompanhar sempre de perto em tudo que faço, pelas sugestões e conselhos sempre valiosos, pelo incentivo e por todo amor que recebo. Sem sua ajuda eu nunca chegaria até aqui. Obrigada!

*"Hoje eu preciso te abraçar
Sentir seu cheiro no meu corpo limpo
Para esquecer de todas as coisas e dormir ao teu lado"*

Dedico este trabalho

A Deus, pela vida

*Ao meu marido Marcelo e
aos meus Pais Lázaro e Iraci*

À **Fapesp**, pelo incentivo financeiro.

Aos colegas da pós-graduação, **Luciana, Priscila, Andréa, Patrícia, Ulises, Rosa, Alan, Amanda**, pela amizade e companheirismo.

Aos companheiros de laboratório, **Flávia, Nathália, Juliana, Luís Gustavo, Fabrício, Márcia e Fernanda** pela troca de conhecimentos, pela ajuda em experimentos, e pela agradável companhia.

Aos amigos, **Felipe, Flávia, Paula, Daniela, Tatiane, Flavinha, Juliana, Nathália e Karina** pelos momentos de descontração e risadas no horário de almoço e a qualquer hora, pelo apoio, pela sincera amizade.

Aos preciosos e amados irmãos, **Heloise, Rogério e Heliane** e aos cunhadinhos, **Flávio, Tatiana e Arthur** por me darem carinho e apoio nesse e em todos outros momentos da minha vida.

Ao **Matheuzinho** e ao **Toli**, basta um sorriso para vocês me encherem de ânimo e felicidade.

Aos amados pais, **Lázaro e Iraci**, pelo apoio incondicional, pelo amor, pelos conselhos, por tudo que sempre fizeram e ainda fazem por mim. Vocês são realmente especiais!!!!

Agradeço em especial ao meu marido, **Marcelo**, por me acompanhar sempre de perto em tudo que faço, pelas sugestões e conselhos sempre valiosos, pelo incentivo e por todo amor que recebo. Sem sua ajuda eu nunca chegaria até aqui. Obrigada!

*"Hoje eu preciso te abraçar
Sentir seu cheiro de roupa limpa
Para esquecer de meus anseios e dormir em paz"*

Jota Quest

A Deus, pela vida.



SUMÁRIO

I. Introdução	
1.1. Definição e classificação dos elementos transponíveis	03
1.2. Origem, proliferação e extinção dos elementos transponíveis	06
1.3. Regulação dos elementos transponíveis	07
1.4. Localização dos elementos transponíveis no genoma hospedeiro	08
1.5. Consequências e implicações evolutivas da presença dos elementos transponíveis no genoma hospedeiro	09
1.6. Caracterização dos elementos transponíveis caracterizados neste estudo	10
2. O gênero <i>Drosophila</i>	13
2.1. O grupo <i>repleta</i>	13
2.2. Elementos transponíveis no grupo <i>repleta</i>	15
II. Objetivos	16
III. Resultados	18
Capítulo 1. Transferência horizontal de elementos transponíveis: impacto na evolução dos elementos e de seus hospedeiros	19
Capítulo 2. Identificação de duas subfamílias do elemento transponível <i>micropia</i> em populações de <i>Drosophila</i> do grupo <i>repleta</i>	49
Capítulo 3. Distribuição descontínua e relações filogenéticas entre retrotransposons <i>copia</i> em espécies de <i>Drosophila</i> do grupo <i>repleta</i>	61
Capítulo 4. Evidência da transferência horizontal do elemento transponível <i>Minos</i> entre espécies de <i>Drosophila</i>	83
IV. Discussão	111
V. Conclusões	121
VI. Referências	122
VII. Resumo	133
VIII. Abstract	134



O conceito da hereditariedade surgiu em 1866 com os experimentos realizados com ervilhas realizados por Gregor Mendel. Apesar de Mendel ter proposto um modelo matemático preciso para a transmissão das unidades hereditárias, seu trabalho não teve o reconhecimento da comunidade científica de sua época. Apenas em 1900, os resultados de Mendel foram simultaneamente redescobertos pelos botânicos Hugo De Vries, Carl Correns e Erich von Tschermak-Seysenegg que, por meio de suas próprias experiências, mostraram que existem regras para transmissão dos fatores genéticos entre os genitores e a sua prole. Em 1909, após várias denominações, W. Johannsen propôs o nome de genes para esses fatores. A teoria que qualifica o gene como unidade discreta do cromossomo foi desenvolvida em 1910 por T. H. Morgan, a partir de estudos com *Drosophila melanogaster*. Essas teorias geraram o conceito de estabilidade genômica, ou seja, que os genes por conterem informações necessárias à vida, deveriam estar protegidos de processos que causam mudanças. Atualmente, no entanto, admite-se que parte do genoma é instável e que justamente parte dessa instabilidade gera uma rápida adaptação dos organismos ao ambiente. Um dos principais responsáveis pela instabilidade genômica são os elementos transponíveis.

A caracterização genética dos elementos transponíveis (TEs, do inglês *Transposable Element*) começou em 1976 com os estudos citogenéticos em milho desenvolvidos por Barbara McClintock. Portanto, a descoberta de McClintock não foi reconhecida cientificamente naquela década, pois questionava o paradigma de estabilidade do genoma. Somente após a caracterização dos elementos transponíveis em leveduras, bactérias e *Drosophila* nas décadas de 60, 70 e início da década de 80 (Levinson, 1962; Kidwell et al., 1977; Green, 1977, 1984, 1986; Engels, 1979; Engels & Preston, 1979; Schaefer et al., 1979; Kidwell, 1979, 1981, 1983, 1984, 1985; O'Hare & Rubin, 1981), essas seqüências foram reconhecidas como importantes componentes dos genomas. Desde então, inúmeros trabalhos corroboram a incrível capacidade de mobilização e de promoção de instabilidade genômica desses elementos genéticos. Atualmente, pesquisadores de várias partes do mundo questionam

I. INTRODUÇÃO

e tentar entender os efeitos da presença dos TEs nos genomas hospedeiros. Estudos recentes mostram que o papel dos elementos transponíveis pode variar de simples expansões do genoma (Kidwell, 2002; Vieira *et al.*, 2002) a antigens de imunidade específica (Agrawal *et al.*, 1995) ou mantenedores dos telômeros (Nakamura & Cech, 1998). Porém, muito da sua influência parece ser irreversível como, por exemplo, nos eventos de mistura de exons e nos rearranjos cromossômicos (Moran *et al.*, 1999).

Devido à sua natureza invasiva e aparente ausência de manifestações fenotípicas benéficas relacionadas à sua presença, os TEs foram inicialmente classificados como parasitas genômicos.

O conceito da hereditariedade surgiu em 1866 com os experimentos realizados com ervilhas realizados por Gregor Mendel. Apesar de Mendel ter proposto um modelo matemático preciso para a transmissão das unidades hereditárias, seu trabalho não teve o reconhecimento da comunidade científica de sua época. Apenas em 1900, os resultados de Mendel foram simultaneamente redescobertos pelos botânicos Hugo De Vries, Carl Correns e Erich von Tschermak-Seysenegg que, por meio de seus próprios experimentos, mostraram que existem regras para transmissão dos fatores genéticos entre os genitores e a sua prole. Em 1909, após várias denominações, W. Johannsen propôs o nome de genes para esses fatores. A teoria que qualifica o gene como unidade discreta do cromossomo foi desenvolvida em 1910 por T. H. Morgan, a partir de estudos com *Drosophila melanogaster*. Essas teorias geraram o conceito de estabilidade genômica, ou seja, que os genes por conterem informações necessárias à vida, deveriam estar protegidos de processos que causam mudanças. Atualmente, no entanto, admite-se que parte do genoma é instável e que justamente parte dessa instabilidade garante uma rápida adaptação dos organismos ao ambiente. Um dos principais responsáveis pela instabilidade genômica são os elementos transponíveis.

A caracterização genética dos elementos transponíveis (TEs, do Inglês *Transposable Elements*) iniciou-se na década de 40, com os estudos citogenéticos em milho desenvolvidos por Barbara McClintock. Entretanto, a descoberta de McClintock não foi reconhecida cientificamente naquela década, pois questionava o paradigma de estabilidade do genoma. Somente após a caracterização dos elementos transponíveis em leveduras, bactérias e *Drosophila* nas décadas de 60, 70 e início da década de 80 (Levitan, 1962; Kidwell *et al.*, 1977; Green, 1977, 1984, 1986; Engels, 1979; Engels & Preston, 1979; Schaefer *et al.*, 1979, Kidwell, 1979, 1981, 1983, 1984, 1985; O'Hare & Rubin, 1983), essas seqüências foram reconhecidas como importantes componentes dos genomas. Desde então, inúmeros trabalhos corroboram a incrível capacidade de mobilização e de promoção de instabilidade genômica desses elementos genéticos. Atualmente, pesquisadores de várias partes do mundo questionam



e tentam entender os efeitos da presença dos TEs nos genomas hospedeiros. Estudos recentes mostram que o papel dos elementos transponíveis pode variar de simples expansores do genoma (Kidwell, 2002; Vieira *et al.*, 2002) a antígenos de imunidade específica (Agrawal *et al.*, 1998) ou mantenedores dos telômeros (Nakamura & Cech, 1998). Porém, muito da sua influência parece ser irreconhecível como, por exemplo, nos eventos de mistura de exons e nos rearranjos cromossômicos (Moran *et al.*, 1999).

Devido à sua natureza invasiva e aparente ausência de manifestações fenotípicas benéficas relacionadas a sua presença, os TEs foram inicialmente classificados como parasitas genômicos, proposta conhecida como hipótese do Gene Egoísta (Doolittle & Sapienza, 1980; Orgel & Crick, 1980). O termo genoma hospedeiro deriva dessa hipótese: se os TEs são parasitas, os genomas que os abrigam podem ser classificados como hospedeiros. Atualmente outras hipóteses, como a de que os elementos transponíveis seriam também importantes simbiontes genômicos têm grande aceitação, porém a denominação de hospedeiro aos organismos que contém os TEs é ainda amplamente utilizada.

1.1 DEFINIÇÃO E CLASSIFICAÇÃO DOS ELEMENTOS TRANSPONÍVEIS

O que define e diferencia um elemento transponível das outras seqüências do genoma é sua capacidade de mobilização, sendo essa capacidade intrínseca ou não ao elemento (Hart *et al.*, 1992). Outras características são importantes para o reconhecimento dessas seqüências, tais como, possuírem de 10^3 a 10^4 pares de bases de comprimento, causarem duplicação do sítio-alvo de inserção; apresentarem polimorfismos de sítios de inserção e apresentarem variabilidade no número de cópias dentro e entre espécies (Capy *et al.*, 1998).

Apesar dos TEs estarem presentes em todos os genomas investigados, sua representatividade varia entre as espécies. Kidwell & Lisch (2000) relataram que os TEs representam mais de 50% do genoma de *Zea mays*, 15% de *Drosophila melanogaster*, mas apenas 5% do genoma de *Caenorhabditis elegans* e 2% de *Arabidopsis thaliana*. A razão dessa discrepância não está completamente entendida, mas o modo de reprodução do hospedeiro (Wright & Finnegan, 2001), condições ambientais e processo de colonização dessas seqüências (Arnault & Dufournel, 1994; Labrador & Fontdevila, 1994, Labrador *et al.*, 1999) parecem influenciar o número de TEs no genoma hospedeiro.

Dois tipos de seqüências móveis podem coexistir em um genoma, as chamadas cópias completas ou autônomas e as incompletas ou não-autônomas. O primeiro tipo codifica as

enzimas necessárias para a transposição, já, o segundo, só se transpõe quando outros elementos no genoma produzem as enzimas necessárias para sua movimentação. Como as seqüências móveis estão sujeitas a altas taxas mutacionais, a frequência de elementos não-autônomos nos hospedeiros é relativamente alta (Kidwell & Lisch, 2001).

A classificação dos TEs baseia-se nas características estruturais e moleculares da suas seqüências, podendo ser agrupados em classes, subclasses, superfamílias e subfamílias. O critério utilizado para o agrupamento em classes é o modo de transposição, sendo duas as classes principais. Elementos da Classe I utilizam a transcriptase reversa para se transporem via um intermediário de RNA. Elementos da Classe II se transpõem diretamente de DNA a DNA usando a enzima transposase por eles codificada (Finnegan, 1985).

O critério utilizado para a classificação em subclasses é a presença ou ausência de domínios e assinaturas nas proteínas codificadas pelos TEs (Capy *et al.*, 1998). A Classe I está dividida em duas subclasses: os retrotransposons, elementos que possuem longas repetições terminais (LTR, do Inglês *Long Terminal Repeats*) são agrupados na Subclasse I e, os retroposons, elementos que não as possuem, são agrupados na Subclasse II. A Classe II, que abriga os transposons, também é dividida em duas subclasses: na subclasse I são agrupados os elementos que possuem a assinatura DDE (D = ácido aspártico e E = ácido glutâmico) na transposase e na subclasse II são agrupados os elementos que não possuem essa assinatura.

As superfamílias são classificadas tanto pela organização da ORF2 (do Inglês, *Open Reading Frame*) como também pela presença ou ausência de motivos ou assinaturas (Capy *et al.*, 1998). Os retrotransposons são subdivididos nas superfamílias *gypsy-like*, *Ty3*, *Tf1* e *Ty1-copia*. As superfamílias *Gypsy-like*, *Ty3* e *Tf1* são diferenciadas de *Ty1-copia* pela organização da ORF2, sendo 5' Proteínase - Integrase - Transcriptase Reversa - RNase H 3' no primeiro grupo e 5' Proteínase - Transcriptase Reversa - RNase - Integrase H 3' no segundo. *Gypsy-like* e *Ty3* são diferenciadas da superfamília *Tf1* pelo fato da última não utilizar um iniciador de tRNA durante a transcrição. A superfamília *Gypsy-like* é diferenciada da *Ty3* pela presença do gene para o envelope viral na primeira e sua ausência na segunda. Por outro lado, os retroposons podem ser subdivididos nas superfamílias LINE (do Inglês, *Long Interspersed Nucleotide Elements*) e SINE (do Inglês, *Short Interspersed Nucleotide Elements*). Os LINES apresentam de 5 a 9 kb, usam promotores internos para a sua transcrição e codificam uma transcriptase reversa e uma RNase. Os SINEs variam de 75 a 500 pb, são formados pela transcrição reversa de transcritos produzidos pela RNA polimerase II e III e não codificam enzimas necessárias para sua transposição, sendo esta fornecida, provavelmente, *in trans* por

elementos LINEs presentes no genoma. As superfamílias dos elementos de DNA são caracterizadas de acordo com a identidade de motivos e seqüências de nucleotídeos. A subclasse I é subdividida nas superfamílias *mariner-Tc1* e *ISa* e a subclasse II nas superfamílias *ISb*, *hAT*, *P*, *MuDR*, *CACTA* e *Tx*.

As famílias são caracterizadas pelo nível de similaridade entre seqüências de um tipo de elemento (30 a 50% de similaridade em seus motivos e seqüências de aminoácidos). As principais famílias da Classe I são: *gypsy*, *Ted*, *Tom*, *17.6*, *297*, *Ty3*, *412*, *micropia*, *Ulysses*, *CFT1*, *Tf1*, *Ty1 copia*, *1731*, *maize*, *roo*, *Alu*, *S1*, *B1*, *B2*, e *TART-HetA*, *jockey*, *R1*, *R2*, *L1*. As principais famílias da Classe II são: *pogo-fot1*, *mariner*, *impala*, *TBE1*, *IS4*, *IS5*, *IS2*, *IS1*, *hobo*, *Ac*, *Tam3S*, *lide*, *P*, *MuDR*, *Spm*, *Tam1S*, *pm*, *Tgm1* e *Tx1D*. A classificação dos elementos transponíveis, de classe a família, está ilustrada na Figura 1.

O último nível de classificação é o de subfamília, que também se baseia no grau de similaridade das seqüências dos diferentes elementos transponíveis (50% a 70% de similaridade em nível de aminoácidos). Exemplos de subfamílias de elementos transponíveis são as subfamílias M, T e O do elemento *P* (Capy *et al.*, 1998) e as subfamílias *mellifera*, *mauritiana*, *cecropia* e *irritans* do elemento *mariner* (Lohe *et al.*, 1995), ambos TEs da Classe II.

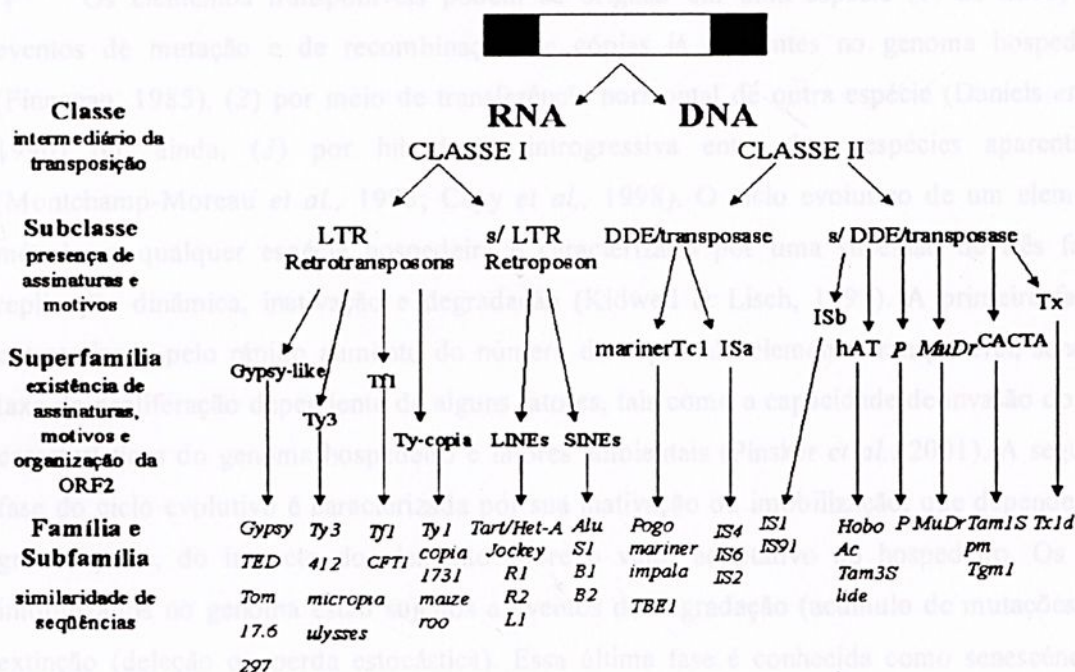


Figura 1. Classificação geral dos elementos transponíveis de classe a família. Modificado de Capy *et al.* (1998).

Além das duas classes bem documentadas, existe ainda, na opinião de alguns autores, uma terceira classe, que seria representada pelos elementos do tipo *MITE* e *Foldback*, cujo mecanismo de transposição ainda permanece desconhecido. A estrutura interna dos *MITEs* é semelhante aos elementos da Classe II, mas seu alto número de cópias no genoma e sua preferência insercional são semelhantes aos elementos da Classe I. Já, os elementos do tipo *Foldback* são caracterizados por conterem amplas repetições terminais que formam estruturas do tipo grampo (Capy *et al.*, 1998).

1.2 ORIGEM, PROLIFERAÇÃO E EXTINÇÃO DOS ELEMENTOS TRANSPONÍVEIS.

Para entender a dinâmica de um TE deve-se primeiro procurar conhecer como essas seqüências surgem, como são reguladas, quais os tipos de pressões às quais estão submetidas e como são eliminadas dos genomas hospedeiros. Apesar desses aspectos ainda não estarem completamente esclarecidos, um panorama geral vem sendo estabelecido.

Os elementos transponíveis podem se originar em uma espécie (1) *de novo*, por eventos de mutação e de recombinação de cópias já existentes no genoma hospedeiro (Finnegan, 1985), (2) por meio de transferência horizontal de outra espécie (Daniels *et al.*, 1990) ou, ainda, (3) por hibridação introgressiva entre duas espécies aparentadas (Montchamp-Moreau *et al.*, 1993; Capy *et al.*, 1998). O ciclo evolutivo de um elemento móvel em qualquer espécie hospedeira é caracterizado por uma sucessão de três fases: replicação dinâmica, inativação e degradação (Kidwell & Lisch, 1997). A primeira fase é caracterizada pelo rápido aumento do número de cópias do elemento transponível, sendo a taxa de proliferação dependente de alguns fatores, tais como a capacidade de invasão do TE, características do genoma hospedeiro e fatores ambientais (Pinsker *et al.*, 2001). A segunda fase do ciclo evolutivo é caracterizada por sua inativação ou imobilização, que depende, em grande parte, do impacto do elemento sobre o valor adaptativo do hospedeiro. Os TEs imobilizados no genoma estão sujeitos a eventos de degradação (acúmulo de mutações) ou extinção (deleção ou perda estocástica). Essa última fase é conhecida como senescência e pode durar milhares de anos (Pinsker *et al.*, 2001).

Quando a autonomia de um TE é perdida em uma espécie, os descendentes deste hospedeiro, por transferência vertical, não possuirão cópias ativas desse elemento no genoma,

a não ser que haja a ressurreição de TEs previamente silenciados através de recombinações ectópicas entre cópias inativas na espécie descendente (Jordan & McDonald, 1998b), o que reiniciaria o ciclo evolutivo do elemento. Além da ressurreição, outro mecanismo que evita a extinção dos TEs é a transferência horizontal. Assim, os TEs transmitidos horizontalmente entre espécies reprodutivamente isoladas reiniciam seu ciclo em um novo hospedeiro (Kidwell & Lisch, 1997). O evento de transferência horizontal, dessa forma, atuaria como um escape ao processo de inativação vertical (Flavell, 1999).

1.3 REGULAÇÃO DOS ELEMENTOS TRANSPONÍVEIS

A transposição não é um processo aleatório, ela pode ser controlada pelo próprio elemento, por sinais dependentes do hospedeiro como também por fatores externos (Grandbastien *et al.*, 1997). Foram propostos alguns modelos para descrever o dinamismo do TE dentro do hospedeiro. O primeiro é baseado na auto-regulação, ou seja, o próprio elemento restringe seu movimento de acordo com sua abundância no genoma (Charlesworth & Langley, 1986). O segundo modelo propõe que os elementos seriam eliminados pela ação da seleção natural, assumindo que os efeitos mutacionais das inserções são negativos (Charlesworth, 1985). De acordo com este modelo, o equilíbrio no número de cópias será atingindo quando $(\mu - v)/s$, onde μ é a taxa de transposição, v é a probabilidade de excisão e s é o coeficiente de seleção associado às inserções dos TEs.

Segundo Badge & Brookfield (1997), os dois modelos citados acima falham ao não incluírem a possibilidade de a abundância dos TEs no genoma ser controlada por fatores do hospedeiro. Por exemplo, a proteína IRBP (do Inglês, *Inverted Repeat Binding Protein*), produzida pelo hospedeiro, é essencial para que ocorra a transposição do elemento *P* em *Drosophila melanogaster*. Essa proteína se liga aos 16 pares de base centrais do invertido terminal desse TE e permite que a transposase seja transcrita (Beall & Rio, 1996). A presença dessa enzima age, assim, como fator limitante da transcrição. Outro exemplo clássico de fatores do hospedeiro que regulam a atividade do TE é o da proteína codificada pelo gene *flamenco* que parece reprimir não só a transposição, mas também a expressão do elemento *gypsy* em *Drosophila* (Pélisson *et al.*, 1994; Alberola *et al.*, 1997). Outros mecanismos epigenéticos do hospedeiro, tais como a metilação e o RNA de interferência (Hurst & Werren, 2001) parecem também regular a atividade dos TEs.

Ainda, a atividade dos TEs pode estimulada por fatores ambientais. Por exemplo, em plantas, a taxa de transposição dos TEs pode ser aumentada em decorrência do estresse gerado por fatores bióticos, como o ataque de parasitas e, por fatores abióticos como o estresse químico e físico (Grandbastien, 1998). Por outro lado, em insetos, o choque de temperatura ativa a transposição de elementos como *P*, *I* e *hobo*, causando fenômenos como o da disgenesia híbrida (Bucheton *et al.*, 1984, Yannopolous *et al.*, 1987). A teoria evolutiva que explica a resposta de TEs a fatores estressantes considera que tenha evoluído uma taxa de transposição alta nestas condições propicia a permanência de uma cópia ativa do elemento na prole do indivíduo estressado.

Independente do modelo de regulação utilizado, o que parece lógico é que existem mecanismos que contém a mobilização indiscriminada dos elementos de transposição, caso contrário, eles poderiam passar a se inserir descontroladamente em diversas posições do genoma, interrompendo o funcionamento de um número cada vez maior de genes necessários à atividade celular. Dessa forma, os TEs alcançam um equilíbrio entre a sua “sobrevivência” e a sobrevivência da célula da qual dependem.

1.4 LOCALIZAÇÃO DOS ELEMENTOS TRANSPONÍVEIS NO GENOMA HOSPEDEIRO

Além de poderem regular a sua transposição, os elementos transponíveis podem ainda apresentar preferências por certos sítios genômicos para sua inserção (Gaur & Li, 2000), ou seja, alguns TEs são altamente específicos durante a transposição. Por exemplo, os elementos *IS4* em *Echerichia coli* sempre se inserem na mesma região do *operon* da galactosidase e outros elementos, como o *IS1*, se inserem preferencialmente em regiões ricas em AT.

Alguns TEs se inserem preferencialmente em regiões distantes das seqüências gênicas do hospedeiro, tais como regiões intergênicas, heterocromatina ou, freqüentemente, dentro de outro elemento. Poucos são os elementos ativos específicos de heterocromatina; dentre eles, pode-se citar os elementos *Het-A*, *TART* e *micropia*, em *D. melanogaster*, *Tras* em *Bombyx mori* e *Ty5* em *Saccharomyces cerevisiae* (Capy *et al.*, 1998). Outros TEs se inserem preferencialmente dentro ou próximo de seqüências de cópia única e, ainda, existem aqueles que não têm preferência por sítios de inserção. Ainda, a distribuição das seqüências ativas e inativas parece não ser aleatória no genoma hospedeiro (Capy *et al.*, 1998). Cópias ativas dos elementos transponíveis são principalmente encontradas na eucromatina, já as cópias

defectivas podem estar presentes tanto na eucromatina quanto na heterocromatina (Biemont *et al.*, 1997).

1.5 CONSEQUÊNCIAS E IMPLICAÇÕES EVOLUTIVAS DA PRESENÇA DOS ELEMENTOS TRANSPONÍVEIS NO GENOMA HOSPEDEIRO

Apesar de décadas de estudos, ainda existem dificuldades em se estimar a influência dos elementos transponíveis na evolução dos genomas. As discussões sobre o assunto vão desde a classificação dessas seqüências como DNA lixo até como parasitas ou simbioses (Kidwell & Lisch, 2001). A última hipótese sugere que a presença dos TEs no genoma hospedeiro representaria uma vantagem evolutiva, à medida que a variabilidade genômica é aumentada a cada nova inserção. Dessa forma, os TEs são fonte geradora inesgotável de variabilidade genética, promovendo rearranjos e mutações que são variantes genéticas sobre as quais a seleção pode atuar (Bingham *et al.*, 1982).

O aumento da variabilidade não é o único efeito sobre genoma hospedeiro. Na verdade, um número cada vez maior de evidências sugere que a atividade dos elementos transponíveis pode influenciar o fenótipo dos hospedeiros de maneira sutil e imperceptível em curto tempo evolutivo (Kidwell & Lish, 2001). Deve-se lembrar que os TEs contêm uma preciosa maquinaria enzimática que pode ser, potencialmente, utilizada pelo hospedeiro. Este recrutamento de função, denominado de domesticação, não é detectado facilmente, isto porque, o processo de recrutamento de função geralmente acontece muito tempo após a invasão do TE no genoma hospedeiro. Essas funções são variadas, tais como a regulação da expressão gênica e a expressão de enzimas e de outras proteínas necessárias ao hospedeiro.

Um recente exemplo de domesticação de um retroelemento tem sido descrito em mamíferos. O gene murine *Fv1*, que proporciona imunidade à infecção causada pelo vírus da leucemia murine, é de fato, uma seqüência *gag-like* sobre controle celular. Essa seqüência *gag-like* teve sua função seqüestrada pelo hospedeiro do gênero *Mus* em um recente passado evolutivo (Best *et al.*, 1996). Outro espetacular exemplo de domesticação parece estar relacionado com o sistema imune *VDJ* em vertebrados (Agrawal *et al.*, 1998; Hiom *et al.*, 1998). O mecanismo de recombinação e a estrutura dos genes *RAG1* e *RAG2*, responsável pelo sistema *V*, *D* e *J*, é altamente similar à estrutura de um elemento transponível, uma vez que *RAG1* e *RAG2* possuem seqüências invertidas terminais e sinais flanqueando a região das ORFs, característicos de TEs. Em *Drosophila* também existem exemplos de elementos



transponíveis desempenhando função para seus hospedeiros. Os elementos *Het-A* e *TART*, em *D. melanogaster*, são seqüências móveis que têm atividade telomérica, se inserindo preferencialmente nas extremidades dos cromossomos e reparando as perdas de material genético que ocorrem nessas regiões a cada divisão celular (Greider, 1996; Pardue *et al.*, 1996). Um elemento *P* em *Drosophila* do subgrupo *montium* constitui um neogene que sintetiza uma proteína de 66 KDa, semelhante à proteína repressora da transposição produzida por elementos ativos mas, que devido a eventos de inserção e mistura (*exon shuffling*) passaram a desempenhar uma nova função no genoma hospedeiro (Nouaud & Anxolabéhère, 1997; Nouaud *et al.*, 2003).

Além dos efeitos citados, os TEs parecem estar relacionados, também, com a origem de microssatélites de DNA. Um exemplo é o dos retroposons de dípteros chamados de *mini-me* que atuam, tanto como agentes dispersos quanto como sítios de origem de microssatélites de DNA. Além dessa propriedade, os TEs também estão envolvidos em eventos de duplicação gênica, presentes em DNAs satélites e considerados agentes relacionados ao envelhecimento por causarem instabilidade genética (Nikitin & Shommokler Reis, 1997; Kidwell, 2002).

Atualmente os elementos transponíveis também desempenham importante papel na ciência moderna, sendo ótimas ferramentas da Biologia Molecular. Por exemplo, os TEs são amplamente utilizados como marcadores moleculares através dos quais podem ser estabelecidas relações filogenéticas entre as diferentes espécies hospedeiras (Capy *et al.*, 1998). Além disso, essas seqüências podem ser utilizadas como vetores da transferência de genes, sendo muito importantes em experimentos de mutagênese (Franz & Savakis, 1991; Klinaskis *et al.*, 2000; Drabek *et al.*, 2003).

1.6 CARACTERIZAÇÃO DOS ELEMENTOS TRANSPONÍVEIS ANALISADOS NESTE ESTUDO

A. Retrotransposon *micropia*

O elemento *micropia* pertence à família *gypsy-Ty3*, e foi primeiramente descrito em alças plumosas do cromossomo Y de *D. hydei* (Lankenau *et al.*, 1989). Esse retrotransposon recebeu o nome *micropia* por ter sido descoberto pela técnica de microclonagem e também por possuir seqüências similares ao retrotransposon *copia*. Seqüências homólogas a esse elemento já foram encontradas nos autossomos de algumas espécies do gênero *Drosophila*. As análises comparativas das seqüências nucleotídicas do elemento *micropia* demonstram que



existem seqüências divergentes dentro de um mesmo genoma. A divergência é provavelmente causada por deleções ou por inserções de outros elementos dentro da sua seqüência, tais como inserções do elemento *copia*.

Lankenau *et al.* (1994) identificaram diferenças entre o número de inserções de *micropia* entre machos e fêmeas de uma mesma espécie. Essa diferença é reflexo da maior incidência deste elemento no cromossomo Y. Sabe-se que em *Drosophila*, o cromossomo Y está relacionado com a fertilidade, portanto, o fato de existirem muitas seqüências *micropia* transcrionalmente ativas neste cromossomo levou aos autores sugerirem que esse elemento possua um importante papel na fertilidade de seu hospedeiro.

O elemento transponível *micropia* apresenta organização gênica típica dos retrotransposons, com aproximadamente 4,9 Kb de comprimento, sendo um dos poucos elementos transponíveis localizados em genes funcionais. Esse elemento é amplamente distribuído no gênero *Drosophila*, tendo sido detectado em espécies do grupo *melanogaster* (*D. melanogaster*, *D. simulans*, *D. birchii*, *D. yakuba*, *D. ananassae*), do grupo *willistoni* (*D. willistoni*), do grupo *immigrans* (*D. immigrans*), do grupo *funbris* (*D. funbris*), do grupo *mellanica* (*D. mellanica*), do grupo *repleta* (*D. repleta*, *D. neorepleta*, *D. paranaensis*, *D. mercatorum*, *D. buzzatii*, *D. mulleri*, *D. eohydei*) por Lankenau *et al.* (1993) e do grupo *saltans* (*D. saltans*) por Lankenau *et al.* (1993) e *D. austrosaltans*, *D. prosaltans*, *D. subsaltans*, *D. parasaltans*, *D. sturtevantii*, *D. dacunhai*, *D. milleri* e *D. neocordata* por Almeida *et al.* (2001).

A atividade do elemento *micropia* é complexa e pouco estudada, tendo apenas um interessante trabalho relacionado com sua expressão (Lankenau *et al.*, 1994). Neste, o autor relatou que seqüências *micropia* de *D. hydei* codificam três RNAs: um RNA *sense* de 5,0 kb e dois RNAs *antisense*, um de 1,0 kb e outro de 1,6 kb (provavelmente originado por inserção na seqüência do RNA *antisense* de 1,0 kb). A expressão dos RNAs *antisense* nessa espécie é trinta vezes maior que a do RNA *sense*, sendo provavelmente o primeiro repressor da expressão do segundo. Outro fato interessante descrito no trabalho é que sua expressão é tecido-específica, sendo produzidos RNAs *antisense* apenas nos testículos.

RNAs *antisense*, codificados pela fita 5'-3', foram descritos inicialmente em procariotos e até pouco tempo acreditava-se que eram restritos a este grupo de organismos. Entretanto, estudos recentes mostram que essa molécula também está amplamente distribuída em eucariotos. O RNA *antisense* é naturalmente transcrito no genoma e pode regular a expressão de seu transcrito complementar *sense* em qualquer etapa da sua transcrição ou tradução, provavelmente atuando na destruição das proteínas regulatórias ligadas ao RNA ou à

formação de RNA dupla fita (Vanhée-Brossollet & Vaquero, 1997). Os estudos atuais sobre RNA antisense são apenas a ponta de um iceberg que deve futuramente trazer novos conhecimentos sobre o ainda pouco entendido campo da expressão gênica.

B. Retrotransposon *copia*

O retrotransposon *copia* pertence à família *copia-Ty1*, e foi primeiramente sequenciado em *D. melanogaster* (Mount & Rubin, 1985). Este elemento tem 5,1 kb de comprimento e possui repetições terminais de 276 pb na mesma orientação do elemento. Ao longo de sua sequência existe uma longa cadeia aberta de leitura capaz de codificar uma poliproteína de 1.409 aminoácidos (Emori *et al.*, 1985). Essa poliproteína gera produtos similares aos genes *gag* e *pol* dos retrovírus. Esse TE é encontrado em representantes de todas as principais radiações de *Drosophila*, o que sugere que tenha surgido cedo na história dos drosofilídeos (Martin *et al.*, 1983).

Apesar de amplamente distribuído no gênero *Drosophila* existem indícios da ocorrência de transferência horizontal de seqüências *copia* entre *D. melanogaster* e *D. willistoni* (Jordan *et al.*, 1999). Os autores observaram maior similaridade entre as LTRs (longas repetições terminais) do elemento *copia* de *D. melanogaster* e *D. willistoni* (99%) do que entre as seqüências *copia* de *D. melanogaster* e de sua espécie críptica *D. simulans* (90%) o que, segundo os autores, é indicativo da ocorrência do evento de transferência horizontal.

C. Transposon *Minos*

O elemento transponível *Minos* pertence à família *Tcl-mariner* e foi primeiramente identificado no DNAr de *D. hydei* (Franz & Savakis, 1991). Atualmente vem recebendo grande atenção nas pesquisas de transformação gênica por não possuir mecanismos de regulação gênica complexos, nem requerer qualquer fator espécie-específico do hospedeiro, sendo assim facilmente manipulado. Estruturalmente, *Minos* apresenta 1,8 kb de comprimento, tendo repetições terminais invertidas de 254 pb. Ao longo da sua extensão existem duas cadeias abertas de leitura que codificam dois polipeptídeos, um de 153 e outro de 201 aminoácidos.

Existe apenas um trabalho relatando a distribuição filogenética desse elemento dentro do gênero *Drosophila* (Arcà & Savakis, 2000). Nesse trabalho foram utilizadas 28 espécies, sendo 19 espécies do subgênero *Sophophora*, 8 do subgênero *Drosophila* e uma do subgênero *Scaptodrosophila* e os resultados mostram uma distribuição descontínua desse elemento no gênero. Além da distribuição descontínua a alta similaridade de seqüências *Minos* entre *D.*

mojavensis e *D. saltans* levou os autores a sugerirem um evento de transferência horizontal para explicar a inconsistência entre a filogenia do elemento e a do hospedeiro. Dentro do grupo *repleta* foram identificadas seqüências homólogas a *Minos* nas espécies: *D. hydei*, *D. neohydei*, *D. mercatorum*, *D. arizonae*, *D. mojavensis* e *D. mulleri*, não havendo relatos sobre as demais espécies do grupo.

2. O gênero *Drosophila*

Uma simples busca aos bancos de dados mostrou mais de 3200 trabalhos, apenas durante o ano de 2003, utilizando *Drosophila* como modelo biológico. Essa popularidade, é devido à grande quantidade de informações acumuladas sobre a ecologia, biologia populacional, comportamento, desenvolvimento, genética e biologia molecular desse organismo. Além disso, suas características morfológicas e a fácil manutenção em estoques são ótimos atrativos para utilizar as espécies desse gênero como material biológico. Soma-se atualmente a essas vantagens o fato do genoma de *Drosophila melanogaster* ter sido completamente seqüenciado (Adams *et al.*, 2000; Celniker & Rubin, 2003).

O gênero *Drosophila* surgiu durante o Eoceno (Cenozóico, Terciário), há aproximadamente 60 milhões de anos, a partir de várias radiações dentro da família *Drosophilidae*. De acordo com Throckmorton (1975), esse gênero sofreu extensiva diversificação ao longo do tempo originando cinco grandes radiações, a mais precoce originou o subgênero *Scaptodrosophila*, sendo seguida pelas radiações *Sophophora* e *Drosophila*, essa com três ramos: radiações *virilis-repleta*, *immigrans* e *hirtodrosophila*. Devido a grande quantidade de espécies, os drosofilistas criaram um sistema de classificação que reconhece, além do gênero, grupos, complexos e *clusters* de espécies. Embora não reconhecida pelo Código Internacional de Nomenclatura Zoológica, essa classificação é amplamente utilizada em estudos envolvendo espécies do gênero.

2.1 O grupo *repleta*

A radiação *virilis-repleta* originou o grupo *repleta* no continente americano, que é constituído por 92 espécies (Tidon-Sklorz *et al.*, 1994; Tidon-Sklorz & Sene, 1995), sendo assim, um dos grupos com maior número de espécies. Esse grupo está subdividido em cinco subgrupos: *hydei*, *mercatorum*, *mulleri*, *repleta* e *fasciola*. Cada subgrupo, por sua vez, tem sido subdividido em complexos e *clusters*.

O conhecimento atual sobre a filogenia do grupo *repleta* está baseado em dados morfológicos (Throckmorton, 1982a; Vilela 1983), citológicos (Wasserman, 1982; 1992), aloenzimáticos (Zouros, 1973; Richardson *et al.*, 1975; Richardson & Smouse, 1976; Richardson *et al.*, 1977; Heed *et al.*, 1990) e moleculares (Sullivan *et al.*, 1990; Russo *et al.*, 1995; Spicer, 1995; 1996; Rodriguez-Trelles *et al.*, 2000; Durando *et al.*, 2000; Tartarenkov & Ayala, 2001). Apesar da grande quantidade de trabalhos, a filogenia do grupo ainda não foi completamente estabelecida.

As espécies do grupo *repleta* ocupam diferentes habitat, mas sem dúvida, a característica mais notável é a capacidade de utilizarem os tecidos de cactos em decomposição como nicho trófico. Essa adaptação tem permitido às espécies do grupo ocuparem desertos e zonas áridas do Novo Mundo e facilitado a exploração de nichos desocupados, permitindo uma rápida evolução e formação de novas espécies (Vilela & Sene, 1977; Ruiz & Fontdevila, 1981; Tidon-Sklorz & Sene, 1995).

O grau de utilização dos tecidos cactáceos é variado entre as espécies. Por exemplo, a maioria das espécies do subgrupo *mulleri* utiliza as cactáceas como local de criação, chegando a ser essa planta indispensável à vida de certas espécies, como *D. mojaviensis* que é cactofílica estrita. Já, as espécies do grupo *fasciola*, *mercatorum*, *repleta* e *hydei* são generalistas, alimentando-se de outros frutos vegetais (Pereira *et al.*, 1983).

Espécies do subgrupo *fasciola* habitam florestas e encontram-se predominantemente na América do Sul, enquanto que as espécies do subgrupo *mulleri* habitam as regiões de deserto, sendo comumente encontradas no México. As espécies do subgrupo *hydei* habitam o sudeste da América do Norte e o Noroeste da América do Sul, enquanto que as espécies do subgrupo *repleta* são encontradas nas Américas do Norte, do Sul e Central. O menor subgrupo, denominado *mercatorum*, habita especialmente o oeste da Índia, América do Sul e América Central (Vilela, 1983). As espécies *D. repleta*, *D. hydei*, *D. mercatorum*, *D. aldrichi* e *D. buzzatii* têm uma ampla distribuição pelo mundo, essa dispersão deve-se provavelmente ao relacionamento dessas espécies com a atividade humana.

O complexo *buzzatii*, pertencente ao subgrupo *mulleri*, é composto por três *clusters* alopátricos. De acordo com os dados de inversão o *cluster stalker* é o mais primitivo, ocorre nas ilhas do Caribe e na Flórida e contém as espécies *D. stalker* e *D. richardsoni*. O *cluster martensis*, o mais derivado, ocorre nos desertos da Venezuela e Colômbia e contém as espécies *D. martensis*, *D. venezolana*, *D. starmeri* e *D. uniseta* (Ruiz & Wasserman, 1993). O *cluster buzzatii* ocorre no Brasil, Argentina e Bolívia e contém as espécies *D. buzzatii*, *D. borborema*, *D. koepferae*, *D. serido*, *D. gouveai*, *D. antonietae* e *D. seriema* (Tidon-Sklorz &

Sene, 2001). Dentre essas, *D. serido* (Vilela & Sene, 1977), espécie descrita a partir de moscas coletadas em Milagres (Estado da Bahia), das quais deriva a linhagem estudada neste trabalho, é constituída por populações heterogêneas. A alta diversidade interpopulacional fundamentou a proposição de Sene *et al.* (1982) de elevar *D. serido* à categoria de superespécie. Análises de DNA mitocondrial (Manfrin *et al.*, 2001) sugeriram que a superespécie é monofilética em relação às demais espécies do *cluster*, mas é composta de dois clados. Um deles é restrito às regiões sul e central do Brasil e outro é distribuído nas regiões centrais, nordeste e costa Atlântica.

2.2 ELEMENTOS TRANSPONÍVEIS NO GRUPO *REPLETA*

Poucos são os trabalhos relativos à distribuição dos elementos transponíveis em populações de espécies do grupo *repleta* (revisão Biemont & Cizeron, 1999). Das 92 espécies pertencentes ao grupo, para apenas 33 são encontradas informações sobre a presença de elementos transponíveis em seu genoma. Dentre elas, três espécies pertencem ao subgrupo *mercatorum* (*D. mercatorum*, *D. paranaensis* e *D. peninsulares*), quatro ao subgrupo *hydei* (*D. heoydei*, *D. hydei*, *D. hydeioides* e *D. neohydei*), cinco ao subgrupo *repleta* (*D. fulvimaculata*, *D. melanopalpa*, *D. neorepleta*, *D. repleta* e *D. zotii*) e 21 são espécies do subgrupo *mulleri* (*D. arizonae*, *D. aldrichi*, *D. borborema*, *D. buzzatii*, *D. huaylasi*, *D. koepferae*, *D. mayaguana*, *D. mojavensis*, *D. martensis*, *D. mulleri*, *D. navojoa*, *D. nigrodumosa*, *D. richardsoni*, *D. serido*, *D. stalker*, *D. starmeri*, *D. straub*, *D. uniseta*, *D. venezolana*, *D. wheeleri* e *D. meridiana* (<http://flybase.bio.indiana.edu>). A diversidade de elementos relatados até o momento também é relativamente baixa, sendo identificados apenas nove elementos da Classe I (*412*, *bilbo*, *blanco*, *copia*, *gypsy*, *micropia*, *osvaldo*, *R1* e *R2*), quatro da Classe II (*Bari 1*, *gandalf*, *newton* e *Minos*) e dois da Classe III (*galileo*, *kepler*). Dentre esses elementos, os mais frequentes são *micropia*, *Bari1*, *copia*, *gandalf*, *gypsy* e *Minos*. Análises de seqüências e o estabelecimento de relações filogenéticas desses TEs não foram realizadas até o momento para a maioria dos elementos no grupo *repleta*.



Visto a escassez de informações sobre a ocorrência e polimorfismos de seqüências dos elementos transponíveis no grupo *repleta* de *Drosophila*, este trabalho foi realizado com o objetivo geral de avaliar a distribuição, o número de inserções no genoma hospedeiro e o polimorfismo das seqüências dos elementos transponíveis *microplia*, *copla* e *Minos* em 26 espécies do grupo *repleta*. Os objetivos específicos foram os seguintes:

1. Reunir informações da literatura sobre o impacto da transferência horizontal na sobrevivência dos elementos transponíveis e as possíveis consequências dessa transferência nos genomas das espécies hospedeiras.
2. Averiguar a distribuição, o número de inserções genômicas, o polimorfismo de seqüências, a atividade transcricional e as relações evolutivas das seqüências *microplia* em espécies de *Drosophila* do grupo *repleta*.
3. Avaliar a distribuição, o número de inserções genômicas, o polimorfismo de seqüências, as relações evolutivas e a ação da seleção natural sobre a região da LTR-U/LR do retrotransposon *copla* em espécies de *Drosophila* do grupo *repleta*.
4. Avaliar a distribuição, o número de inserções genômicas, o polimorfismo de seqüências e relações evolutivas das seqüências *Minos* em espécies de *Drosophila* do grupo *repleta*.

Os resultados obtidos foram apresentados na forma de artigos científicos. O Capítulo 1 é um artigo de revisão referente ao objetivo 1, o qual será submetido à publicação em revista nacional.

II. OBJETIVOS

O objetivo 2 é está apresentado na forma de artigo *in press* no periódico *Genetica* (Holanda). Os Capítulos 3 e 4 referem-se aos objetivos 3 e 4 e estão redigidos na forma de artigos a serem submetidos à publicação nos periódicos *Hereditas* (Inglaterra) e *Molecular Phylogenetics and Evolution* (Estados Unidos), respectivamente.



Visto a escassez de informações sobre a ocorrência e polimorfismos de seqüências dos elementos transponíveis no grupo *repleta* de *Drosophila*, este trabalho foi realizado com o objetivo geral de avaliar a distribuição, o número de inserções no genoma hospedeiro e o polimorfismo das seqüências dos elementos transponíveis *micropia*, *copia* e *Minos* em 26 espécies do grupo *repleta*. Os objetivos específicos foram os seguintes:

1. Reunir informações da literatura sobre o impacto da transferência horizontal na sobrevivência dos elementos transponíveis e as possíveis conseqüências dessa transferência nos genomas das espécies hospedeiras.

2. Averiguar a distribuição, o número de inserções genômicas, o polimorfismo de seqüências, a atividade transcricional e as relações evolutivas das seqüências *micropia* em espécies de *Drosophila* do grupo *repleta*.

3. Avaliar a distribuição, o número de inserções genômicas, o polimorfismo de seqüências, as relações evolutivas e a ação da seleção natural sobre a região da LTR-ULR do retrotransposon *copia* em espécies de *Drosophila* do grupo *repleta*.

4. Avaliar a distribuição, o número de inserções genômicas, o polimorfismo de seqüências e relações evolutivas das seqüências *Minos* em espécies de *Drosophila* do grupo *repleta*.

Os resultados obtidos foram apresentados na forma de artigos científicos. O Capítulo 1 é um artigo de revisão referente ao objetivo 1, o qual será submetido à publicação em revista nacional. O Capítulo 2 refere-se ao objetivo 2 e está apresentado na forma de artigo *in press* no periódico Genetica (Holanda). Os Capítulos 3 e 4 referem-se aos objetivos 3 e 4 e estão redigidos na forma de artigos a serem submetidos a publicação nos periódicos Heredity (Inglaterra) e Molecular Phylogenetics and Evolution (Estados Unidos), respectivamente.



TRANSFERÊNCIA HORIZONTAL DE ELEMENTOS
TRANSPONÍVEIS: IMPACTO NA EVOLUÇÃO DOS
ELEMENTOS E DE SEUS HOSPEDEIROS

III. RESULTADOS

Capítulo 1



Transferência horizontal de elementos transponíveis: impacto na evolução dos elementos e de seus hospedeiros.

Almeida, L.M.; Carareo, C.M.A.

Departamento de Biologia, Universidade Estadual Paulista - UNESP, 15054-000 São José do Rio Preto, SP, Brazil, e-mail: carareo@bta.ibilce.unesp.br. Phone:55-17-2312382, Fax:55-17-2312399. *Author for correspondence.

TRANSFERÊNCIA HORIZONTAL DE ELEMENTOS

TRANSPONÍVEIS: IMPACTO NA EVOLUÇÃO DOS ELEMENTOS E DE SEUS HOSPEDEIROS

Resumo: Inúmeros trabalhos, desde a descoberta dos TEs na década de 40, por Barbara McClintock, têm demonstrado a importância da transferência horizontal na evolução genômica desses elementos. Porém, muita permanece obscuro no campo das pesquisas sobre eventos de transferência horizontal desses elementos genômicos. Como sequências autônomas, os TEs precisam se multiplicar ao nível genômico e populacional para se manterem ao longo das gerações. Nesse contexto, a ideia de a transferência horizontal ser uma estratégia que habilita aos elementos transponíveis fugirem da inativação vertical e extinção vem crescendo no meio científico. Esta revisão tem como objetivo reunir informações da literatura sobre o impacto da transferência horizontal na sobrevivência dos elementos transponíveis e as possíveis consequências dessa transferência no genoma das espécies hospedeiras. Para melhor compreensão dos objetivos, foi realizada uma breve descrição de importantes características dos TEs relacionadas ao fenômeno da transferência horizontal, tais como, ciclo evolutivo, classificação, modos de transposição e incidência das diferentes classes de TEs no genoma hospedeiro.

Palavras chave: transferência horizontal, elementos transponíveis, evolução, genômica, transposição.

Capítulo 1



Transferência horizontal de elementos transponíveis: impacto na evolução dos elementos e de seus hospedeiros.

Almeida, L.M.; Carareto, C.M.A.

Departamento de Biologia, Universidade Estadual Paulista - UNESP, 15054-000 São José do Rio Preto, SP, Brazil, e-mail: carareto@bio.ibilce.unesp.br, Phone:55-17-2212382, Fax:+55-17-2212390. *Author for correspondence.

Resumo: Inúmeros trabalhos, desde a descoberta dos TEs na década de 40, por Barbara McClintock, corroboram a incrível capacidade de transposição e de promoção de instabilidade genômica desses elementos. Porém, muito permanece obscuro no campo das pesquisas sobre eventos de transferência horizontal dessas seqüências genéticas. Como seqüências autônomas, os TEs precisam se multiplicar ao nível genômico e populacional para se manterem ao longo das gerações. Nesse contexto, a idéia de a transferência horizontal ser uma estratégia que habilita aos elementos transponíveis fugirem da inativação vertical e extinção vem crescendo no meio científico. Esta revisão tem como objetivo reunir informações da literatura sobre o impacto da transferência horizontal na sobrevivência dos elementos transponíveis e as possíveis conseqüências dessa transferência no genoma das espécies hospedeiras. Para melhor compreensão dos objetivos, foi realizada uma breve descrição de importantes características dos TEs relacionadas ao fenômeno da transferência horizontal, tais como, ciclo evolutivo, classificação, modos de transposição e incidência das diferentes classes de TEs no genoma hospedeiro.

Palavras chave: transferência horizontal, elementos transponíveis, evolução genômica, transposição.





Figura 1. Árvore filogenética rotineiramente representada para as espécies de transferência horizontal entre as células.

I. Introdução (Doolittle (1999))

Antes de Darwin, o paradigma vigente sobre a origem das espécies não era científico, mas sim religioso: o criacionismo postulava ser Deus o criador de todos seres vivos, os quais permaneciam imutáveis ao longo do tempo. A partir de Charles Darwin, em 1859, com “A Origem das Espécies” passou-se a aceitar, particularmente no meio científico, que as forças seletivas modelam a evolução das espécies, originando-se o paradigma evolutivo que postula a existência de uma única e universal árvore filogenética para explicar as relações entre todos os seres vivos. O paradigma Darwiniano recebeu importantes contribuições de várias ciências que se desenvolveram no século XX, a Genética em especial, originando o Neodarwinismo, uma ampliação do Darwinismo clássico. Todavia, a visão Neodarwinista de transmissão do material genético com base unicamente na reprodução e na transferência vertical dos genes vem sendo ampliada atualmente. Complementando novamente o Darwinismo, outra teoria define que as relações entre os seres vivos são representadas por uma rede complexa de relações que, muitas vezes, se estabelecem entre espécies não aparentadas (Doolittle, 1999). Esse processo recebe o nome de transferência horizontal de genes (Figura 1).

Transferência horizontal de genes é conceituada como transferência assexual, por meio de vetores, da informação genética de um genoma para o outro, entre espécies distantemente aparentadas ou não, mas que compartilham o mesmo tempo e espaço geográfico. Esse fenômeno recebeu tal nome para ser diferenciado da transferência vertical usual de genes, segundo a qual a geração parental transfere a informação genética para a prole (Graur & Li, 2000).

Transferência horizontal de genes (Silva et al., 2004). Os elementos transportadores apresentam um aspecto peculiar nesse processo, pois, além de se transferirem horizontalmente entre as espécies por meio de outros vetores, como por exemplo, um vírus parenta (Block et al., 1991), podem também, eles próprios servir de vetores de transferência horizontal de outros

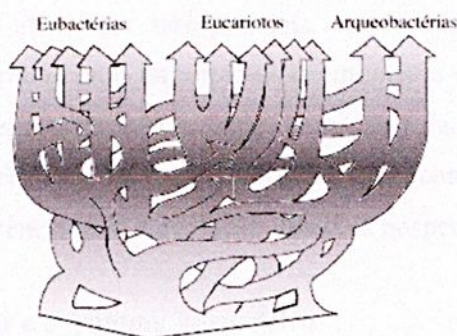


Figura 1. Árvore filogenética reticulada representando possíveis eventos de transferência horizontal entre os três domínios. De acordo com Doolittle (1990).

A transferência gênica horizontal já está bem documentada em organismos procarióticos (Groisman *et al.*, 1992; Lawrence *et al.* 1992; Coffey *et al.*, 1993; Mariani *et al.*, 1993). Também está bem documentada a transferência gênica horizontal, independente da direção, entre procariotos e eucariotos (Carlson & Chelm 1986; Doolittle *et al.*, 1989; Froman *et al.*, 1989; Sikorski *et al.*, 1990; Doolittle *et al.*, 1990; Smith & Doolittle 1992; Brok & Doolittle, 1992; Smith *et al.*, 1992; Nishikawa *et al.*, 1992). A transferência horizontal de genes entre eucariotos é, entretanto, uma questão polêmica. Apesar de existirem alguns estudos sobre esse assunto (Benveniste *et al.*, 1975; Scarpulla, 1984; Syvanen, 1986; Daniels *et al.*, 1990a; Maruyama & Hartl, 1991; Houck *et al.*, 1991; Clark *et al.*, 1994; Hagemann *et al.*, 1996), são relativamente escassas as evidências de que os genes podem se mover entre organismos eucariotos distintos. Esses trabalhos, além de tratarem da importância da transferência horizontal na evolução dos seres vivos, também alertam para o fato que esse fenômeno pode ser mais freqüente que previsto anteriormente.

* Os genes, em eucariotos, podem ultrapassar a barreira das espécies através da hibridização introgressiva ou da transferência horizontal por meio de vetores. A transferência gênica por hibridização ocorre principalmente entre espécies crípticas, enquanto que, a transferência por vetores não requer proximidade filogenética; exige apenas que as espécies sejam simpátricas. Os vetores mais comuns são insetos parasitas, vírus, bactérias simbiotes e elementos transponíveis (Silva *et al.*, 2004). Os elementos transponíveis apresentam um aspecto peculiar nesse processo, pois, além de se transferirem horizontalmente entre as espécies por meio de outros vetores, como por exemplo, um ácaro parasita (Houck *et al.*, 1991), podem também, eles mesmos serem os vetores da transferência horizontal de outros

genes. Isso ocorre quando o elemento tem um gene do seu hospedeiro inserido dentro da sua própria seqüência.

Nesse contexto, os elementos transponíveis, com sua maquinaria própria e sua capacidade de inserção, figuram como importantes componentes da troca genética horizontal. Em face dessa possibilidade, esta revisão enfocará o impacto da transferência horizontal dos elementos transponíveis não só na sua sobrevivência, como também, as possíveis conseqüências dessa transferência para a evolução dos seus hospedeiros.

II. Elementos transponíveis e o genoma hospedeiro

A caracterização genética dos elementos transponíveis iniciou-se na década de 40, com os estudos citogenéticos em milho desenvolvidos por Barbara McClintock. Entretanto, a descoberta de McClintock não foi reconhecida cientificamente naquela década, pois questionava o paradigma de estabilidade do genoma. Somente após a caracterização dos elementos transponíveis em leveduras, bactérias e *Drosophila* nas décadas de 60, 70 e início da década de 80 (Levitan, 1962; Kidwell *et al.*, 1977; Green, 1977, 1984, 1986; Engels, 1979; Engels & Preston, 1979; Schaefer *et al.*, 1979; Kidwell, 1979, 1981, 1983, 1984, 1985; O'Hare & Rubin, 1983), essas seqüências foram reconhecidas como importantes componentes dos genomas.

Os elementos transponíveis (TEs) são seqüências particulares de DNA, moderadamente repetidas, geralmente de 10^3 a 10^4 pares de base de comprimento, que têm a habilidade de mudar de posição dentro e entre seqüências gênicas. Essa capacidade pode ser intrínseca ou não ao elemento (Hartl *et al.*, 1992). O número de inserções dessas seqüências dentro do genoma hospedeiro pode variar desde poucas cópias até milhares. Cerca de 45% do genoma humano, por exemplo, é constituído de seqüências móveis, em contraposição a apenas 1,5% constituído pelos exons funcionais (International Human Genome Sequencing Consortium, 2001). Estas porcentagens geram discussões a respeito da verdadeira função dos TEs nos genomas hospedeiros.

A presença dos elementos transponíveis pode ter diferentes conseqüências evolutivas nos genomas hospedeiros. Uma dessas conseqüências é o aumento da variabilidade genética, fenômeno gerado a cada nova inserção do elemento (Kidwell & Lish, 1997; Kidwell, 2002). Segundo Vieira *et al.* (2002), além do aumento da variabilidade, cada nova inserção, também aumenta o tamanho do genoma, sendo um dos fatores responsáveis pelo paradoxo do valor C; por essa razão os TEs também são denominados expansores gênicos. Outra conseqüência é a



re-estruturação da arquitetura genômica, originada pelos rearranjos cromossômicos entre diferentes cópias dos TEs; esses rearranjos podem muitas vezes envolver também seqüências do DNA hospedeiro adjacente aos TEs (Capy *et al.*, 1998). A presença de TEs pode ainda estar relacionada com a regulação da expressão gênica do hospedeiro, uma vez que produtos gênicos codificados pelos TEs podem interagir com a maquinaria bioquímica do hospedeiro (Yang *et al.*, 1998). Além disso, os TEs podem estar relacionados com a expressão de novos genes, uma vez que suas seqüências podem ser recrutadas pelo genoma para exercerem novas funções; esse processo é denominado de domesticação (Kidwell & Lisch, 2001; Britten, 1997). Outra possível função para os elementos transponíveis é a de vetores naturais da transposição lateral de genes. Devido a essa característica, atualmente, os elementos transponíveis têm sido utilizados como ferramentas na manipulação genética e em experimentos de mutagênese (Franz & Savakis, 1991; Klinakis *et al.*, 2000; Shimizu *et al.*, 2000; Zagoraiou *et al.*, 2001; Kaminski *et al.*, 2002; Sasakura *et al.*, 2003; Drabek *et al.*, 2003; O' Brochta *et al.*, 2003).

III. Modos de transposição

Os mecanismos de transposição referem-se à natureza geral das reações através das quais um elemento de transposição se insere em um novo sítio genômico. Dois processos são descritos, o deslocamento da seqüência na forma de DNA ou na forma de RNA. Deve-se lembrar que apesar de todos os elementos existirem integrados ao genoma, na forma de DNA, alguns se deslocam através dos seus RNAs.

De acordo com o modo de transposição e suas características estruturais e moleculares, os elementos transponíveis podem ser classificados em classes, subclasses, superfamílias, famílias e subfamílias. As classes são subdivididas de acordo com o mecanismo de transposição do elemento, e as subclasses de acordo com a estrutura molecular dos seus produtos gênicos, ou seja, presença ou ausência de domínios e assinaturas. As superfamílias se diferenciam pela existência de domínios e assinaturas similares, enquanto que as famílias se caracterizam por apresentarem similaridade de 30 a 50% em seus motivos e seqüência de aminoácidos. Finalmente, as subfamílias agrupam os elementos que possuem similaridade entre 50 a 70% na seqüência de aminoácidos (Capy *et al.*, 1998).

Os elementos da Classe I são os TEs que se transpõem através de um RNA intermediário que é transcrito reversamente, pela transcriptase reversa, antes da nova inserção. Por outro lado, os elementos da Classe II se transpõem diretamente de DNA a DNA, usando a enzima transposase.



De acordo com Lewin (1994), o mecanismo de transposição via DNA pode ser conservativo ou replicativo. No primeiro, o elemento transponível é simplesmente removido de um dado local e inserido em outro, não aumentando dessa forma seu número de inserções no genoma. No segundo, o TE é duplicado antes de ser transportado para um novo local; dessa forma, uma cópia do elemento permanece no local e outra é inserida em um novo sítio, aumentando assim o número de inserções desse no genoma (Figura 2).

O processo de transposição via RNA é basicamente replicativo, uma vez que o RNA intermediário da transposição é transcrito reversamente em uma nova cópia de DNA. No caso dos retrotransposons com LTRs, essa molécula de DNA extracromossomal direciona-se para o núcleo da célula, sendo a enzima endonuclease por ele codificada responsável pela sua integração no genoma hospedeiro (Lewin, 1994). Entretanto, no caso dos retrotransposons sem LTRs, o RNA intermediário se ancora ao DNA-alvo antes de ser transcrito em DNA (Figura 2), dessa forma durante a transposição desses elementos não há moléculas de DNA livres no meio celular (Ostertag *et al.*, 2001). Os mecanismos utilizados na transposição estão estreitamente relacionados à capacidade do elemento ser transferido horizontalmente entre espécies (Silva *et al.*, 2004).

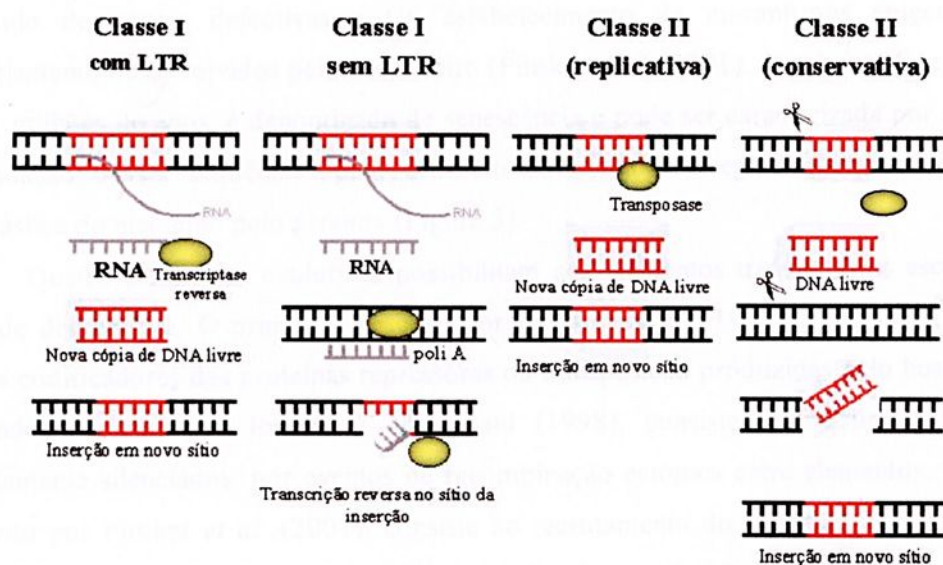


Figura 2. Mecanismos de transposição utilizados pelos elementos transponíveis. Elementos da Classe I (com LTRs e sem LTRs) utilizam a transcriptase reversa como principal enzima no processo de transposição. Elementos da Classe II (com transposição replicativa e conservativa) utilizam a transposase como principal enzima da transposição. Poli A refere-se a cauda de poliadenina do RNA mensageiro produzido pelo TE; transcriptase reversa, ✂ transposase utilizada para liberar o elemento da fita dupla de DNA durante o mecanismo denominado *cut and paste*.

IV. Ciclo de vida e inativação vertical dos elementos transponíveis

Existem três maneiras pelas quais uma família de elementos móveis pode se originar em uma espécie: (i) *de novo*, ou seja, por eventos de mutação e de recombinação de seqüências já presentes no genoma (Finnegan, 1985); (ii) por meio de transferência horizontal mediada por um ou mais vetores (Daniels *et al.*, 1990a) e (iii) por meio de hibridização introgressiva e poliespermia entre espécies aparentadas (Capy *et al.*, 1998). Posteriormente, os novos integrantes do genoma podem se espalhar por transmissão vertical a todas as populações a partir de linhagens ascendentes.

Após a origem, ocorre uma rápida propagação e aumento do número de cópias destas seqüências no genoma hospedeiro, levando a uma diversidade de sítios de inserção (Kaplan *et al.*, 1985). Nesse momento, as interações que se estabelecem entre o TEs e o genoma são muito importantes para a sobrevivência de ambos. A taxa com a qual esse fenômeno ocorre depende de vários fatores que regulam a transposição do elemento, como a sua capacidade de invasão, as características do genoma hospedeiro, fatores ambientais e ainda a interação de todos esses fatores (Pinsker *et al.*, 2001). A fase seguinte do ciclo de vida de um TE é caracterizada por dois eventos: (i) redução crescente da mobilidade do elemento devido ao acúmulo de cópias defectivas e (ii) estabelecimento de mecanismos epigenéticos de silenciamento desenvolvidos pelo hospedeiro (Pinsker *et al.*, 2001). A próxima fase, que pode durar milhões de anos, é denominada de senescência e pode ser caracterizada por eventos de degradação dessas seqüências pelo acúmulo de mutações, seguida pela provável perda estocástica do elemento pelo genoma (Figura 3).

Quatro estratégias evolutivas possibilitam aos elementos transponíveis escaparem da fase de degradação. O primeiro, descrito por Pelisson *et al.* (1994), consiste na perda dos alelos codificadores das proteínas repressoras da transposição produzidas pelo hospedeiro. O segundo, descrito por Jordan & McDonald (1998), consiste na ressurreição de TEs, previamente silenciados, por eventos de recombinação ectópica entre elementos. O terceiro, descrito por Pinsker *et al.* (2001), consiste no recrutamento do elemento transponível pelo genoma hospedeiro para o desenvolvimento de alguma função que beneficie a espécie, originando os agora denominados neogenes. Por último, a transferência horizontal, que consiste na invasão das seqüências móveis em um genoma ainda desprotegido (sem mecanismos de repressão da transposição) propicia aos TEs uma forma de escapar da inativação vertical (Flavell, 1999). Silva & Kidwell (2000) propuseram que a principal e

talvez, a única fonte de seleção, realmente atuante sobre os TEs, seja a transferência horizontal. Esse seria o mecanismo essencial para a sobrevivência dos elementos móveis.

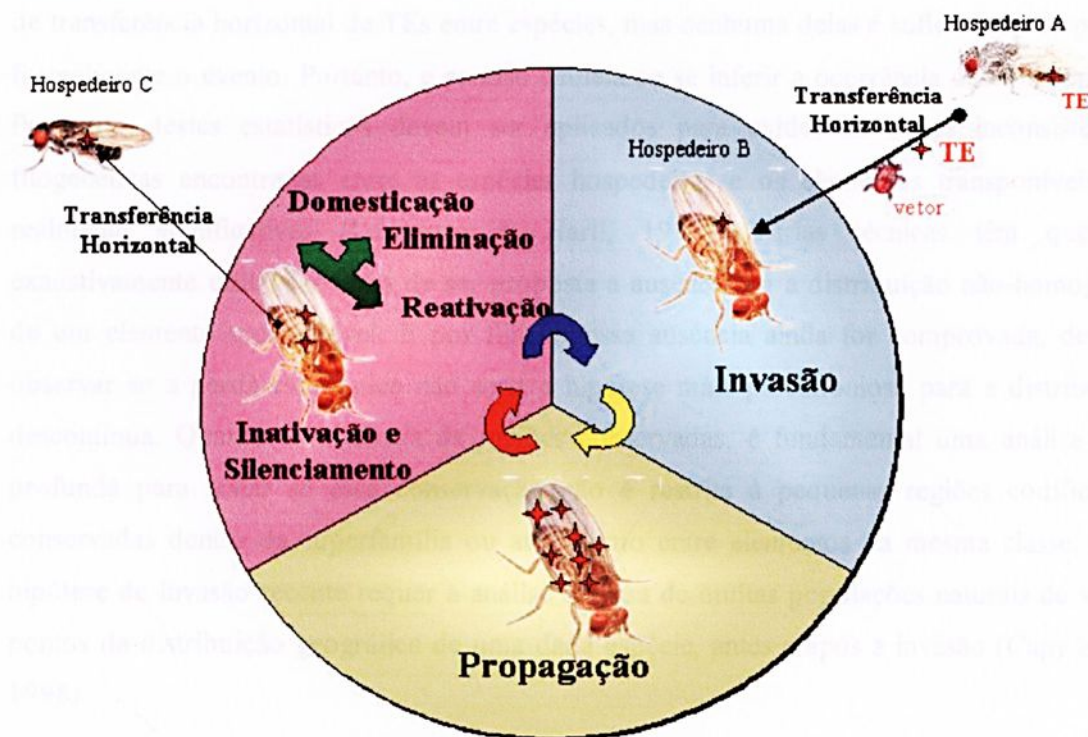


Figura 3. Modelo ilustrando as principais fases do ciclo evolutivo de um elemento transponível (TE †), usando *Drosophila* como espécie hospedeira. Primeira fase: invasão por transferência horizontal (TH →). Segunda fase: propagação e aumento do número de cópias dos TEs. Terceira fase: inativação e silenciamento dos TEs e possíveis destinos para estas seqüências (eliminação, domesticação, novo evento de transferência horizontal).

V. Quando a transferência horizontal pode ser sugerida

† A ocorrência de transferência gênica horizontal de TEs pode ser inferida quando uma das quatro situações seguintes pode ser observada: (i) alto grau de similaridade entre seqüências de TEs que ocorrem em diferentes espécies, impossível de ser conciliado com a filogenia das espécies hospedeiras (Silva & Kidwell, 2000); (ii) distribuição descontínua ou não homogênea do elemento em um grupo de espécies (Daniels *et al.*, 1990a); (iii) existência de motivos altamente conservados dentro da seqüência de diferentes TEs em espécies

distantemente relacionadas; ou ainda, (iv) invasão recente do elemento transponível em uma espécie hospedeira (Capy *et al.*, 1994). As quatro situações (Figura 3) sugerem a ocorrência de transferência horizontal de TEs entre espécies, mas nenhuma delas é suficiente para provar formalmente o evento. Portanto, é preciso cautela ao se inferir a ocorrência desse fenômeno. Para isto, testes estatísticos devem ser aplicados para evidenciar se as inconsistências filogenéticas encontradas entre as espécies hospedeiras e os elementos transponíveis são realmente significativas (Lawrence & Hartl, 1992). Várias técnicas têm que ser exaustivamente utilizadas antes de ser proposta a ausência ou a distribuição não-homogênea de um elemento transponível. E por fim, se essa ausência ainda for comprovada, deve-se observar se a perda estocástica não seria a hipótese mais parcimoniosa para a distribuição descontínua. Quanto à existência de regiões conservadas, é fundamental uma análise mais profunda para testar se essa conservação não é restrita à pequenas regiões codificantes conservadas dentro da superfamília ou até mesmo entre elementos da mesma classe. Já, a hipótese de invasão recente requer a análise extensa de muitas populações naturais de vários pontos da distribuição geográfica de uma dada espécie, antes e após a invasão (Capy *et al.*, 1998).

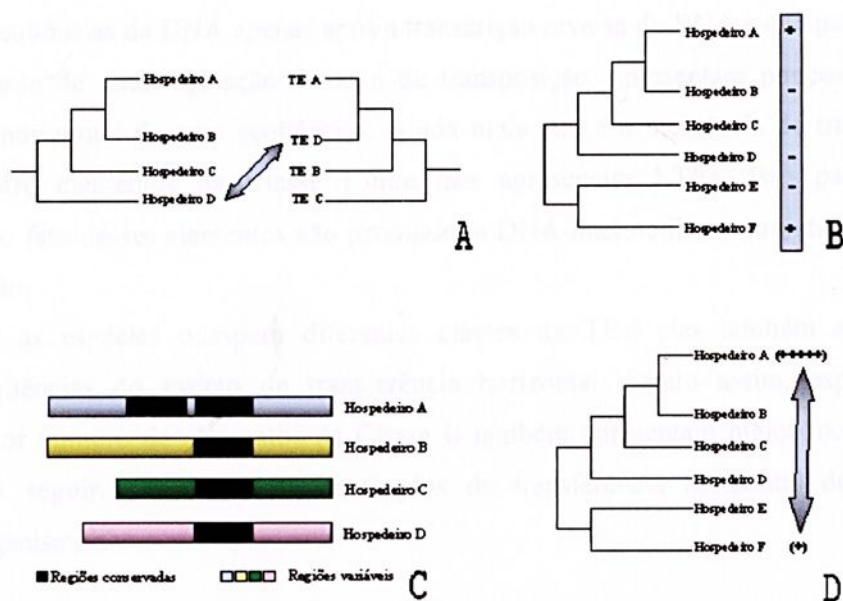


Figura 4. Eventos que sugerem a ocorrência de transferência horizontal. A: inconsistência entre a filogenia do hospedeiro e a do elemento transponível; B: distribuição descontínua do elemento transponível entre as espécies; C: regiões conservadas entre elementos transponíveis de espécies distantemente relacionadas e D: invasão recente de um elemento transponível. Modificado de Capy *et al.* (1998).

VI. Freqüência da transferência horizontal

Inúmeros relatos de transferência horizontal de TEs têm surgido nos últimos anos (Vásquez-Manrique *et al.*, 2000; Silva & Kidwell., 2000; Koga *et al.*, 2000; Zupunski *et al.* 2001; Lorcto *et al.*, 2001; Robertson *et al.*, 2002; Clark *et al.*, 2002), o que suporta a proposição de que esse fenômeno é mais freqüente do que se supunha anteriormente. Ainda, a freqüência desse evento não é a mesma para as diferentes classes de elementos transponíveis. Silva *et al.* (2004) argumentaram que a incidência do evento é bem maior entre os elementos da Classe II do que entre os da Classe I. Evidências mostram que o modo de transposição e a auto-regulação do número de cópias no genoma hospedeiro afetam sensivelmente a incidência do evento de transmissão horizontal (Luan *et al.*, 1993; Malik *et al.*, 1999; Silva *et al.*, 2004). Os elementos da Classe II se transpõem diretamente de DNA a DNA, sob ação da enzima transposase, e possuem a capacidade de regular sua própria taxa de transposição. De acordo com Silva *et al.* (2004), tais características parecem limitar seu número de cópias ao longo do tempo, sendo praticamente impossível evitar sua extinção no genoma hospedeiro. Desse modo, o fenômeno de transferência horizontal seria estritamente necessário para essa classe de elementos sobreviver ao longo da história evolutiva. Já os elementos da Classe I, que produzem as seqüências de DNA apenas após a transcrição reversa do RNA e que parecem não ter a capacidade de auto-regulação da taxa de transposição, apresentam poucos casos de transferência horizontal de suas seqüências. Ainda mais rara é a incidência de transferência horizontal entre elementos da Classe I que não apresentam LTRs. Isso parece estar relacionado ao fato desses elementos não produzirem DNA intermediário durante o processo de transposição.

Como as espécies possuem diferentes classes de TEs, elas também apresentam variadas freqüências do evento de transferência horizontal. Sendo assim, espécies que possuem maior número de elementos da Classe II também apresentam maior incidência do fenômeno. A seguir veremos alguns exemplos de transferência horizontal de TEs em diferentes organismos.

Transferência horizontal em bactérias

Em bactérias, eventos de transferência horizontal podem acontecer por meio de três mecanismos. O primeiro denominado de transformação, consiste na obtenção de seqüências de DNA livres no meio ambiente (não necessita de vetores para a transferência). O segundo



denominado de conjugação, consiste na transferência de DNA de um indivíduo para outro por meio de vetores tais como os plamídeos ou elementos transponíveis. O terceiro denominado de transdução, consiste na transferência do material genético por meio de um bacteriófago.

A transferência horizontal de elementos transponíveis pode acontecer em bactérias em alta frequência. Por exemplo, o elemento *Tn* está relacionado a vários eventos de transferência horizontal de genes envolvidos com resistência a antibióticos (Berg & Howe, 1989; Hall, 1997). Outro exemplo de transferência horizontal em bactérias pode ser ilustrado pela filogenia do elemento *IS1* (Insertion Sequence I) em diferentes espécies de *Escherichia*. Pela filogenia do gene *ompA*, a espécie *E. coli* está mais relacionada a *E. hermannii*, porém, a sequência *IS* de *E. coli* está mais relacionada à de *E. vulneris*. Esses dados levaram Lawrence *et al.* (1992) a sugerirem um evento de transferência horizontal das sequências de inserção *IS1* entre *E. coli* e *E. vulneris*.

Transferência horizontal em insetos.

Um dos exemplos mais interessantes e consistentes de divergência entre a filogenia de TEs e de sua espécie hospedeira é o caso do elemento *P* da subfamília canônica em espécies distantemente relacionadas de *Drosophila*. Análises moleculares de sequências *P* permitiram a identificação de dois tipos estruturais distintos, os completos e os defectivos. O elemento *P* completo possui 2.907 pb em seu comprimento e é estruturalmente composto por quatro *exons*, flanqueados por sequências invertidas terminais de 31 pb, e uma repetição subterminal invertida de 11 pb (O'Hare & Rubin, 1983).

A comparação entre todas as sequências *P* descritas nos drosofilídeos sugere que existam várias subfamílias que divergem aproximadamente 30% em suas sequências de DNA em relação ao elemento canônico de *D. melanogaster*. Até o momento, além da subfamília canônica, foram descritas com detalhes outras três subfamílias (tipo-M, tipo-O e tipo-T). O elemento *P* canônico espalhou-se rapidamente entre populações naturais de *D. melanogaster* nas últimas quatro décadas. Apesar de estar amplamente distribuído em *D. melanogaster*, está ausente em todas as outras espécies do subgrupo *melanogaster*, tais como *D. mauritiana*, *D. sechellia*, *D. simulans* e *D. yakuba*. Por outro lado, espécies dos subgrupos *willistoni* e *saltans*, distantemente relacionadas do grupo *melanogaster*, possuem tais sequências *P*. Daniels *et al.* (1990a), em um amplo estudo mostrou que o elemento *P* de *D. willistoni* é idêntico ao de *D. melanogaster*, com exceção de uma única substituição nucleotídica de A para G na posição 32. Essa semelhança é completamente inconsistente com os estimados 50 milhões de anos que separam a história evolutiva das duas espécies. Além da distribuição

descontínua e a alta similaridade de seqüências, outros dois argumentos fortalecem a hipótese de transferência horizontal. O primeiro é que linhagens antigas de *D. melanogaster*, coletadas na década de 40, e mantidas em laboratório até o presente, não contém seqüências *P* canônicas, o que corrobora a hipótese de invasão recente (Anxolabéhère *et al.*, 1984). O segundo é que *D. willistoni* e *D. melanogaster* compartilham o mesmo nicho ecológico (da Flórida ao Sul da América do Sul) onde, utilizando os mesmos sítios de ovoposição, os ovos de ambas as espécies são parasitados pelo ácaro *Proctolaelaps regalis*. Esse ácaro possui um aparelho bucal que se assemelha a uma agulha e, ao se alimentar dos ovos das diferentes espécies, pode ter servido como vetor da transferência horizontal de seqüências de TEs (Houck *et al.*, 1991).

Exemplos de inconsistências entre a filogenia do elemento *P* e a filogenia da sua espécie hospedeira não estão restritos a *D. melanogaster* e *D. willistoni*. Silva & Kidwell (2000) sugeriram a ocorrência de pelo menos 11 eventos de transferência horizontal para explicar a distribuição das seqüências *P* nos grupos *saltans* e *willistoni* de *Drosophila*. Haring *et al.* (2000) também sugeriram a ocorrência de pelo menos oito eventos de transferência para explicar a distribuição do elemento *P* do tipo O em espécies de *Drosophila*, *Scaptomyza* e *Lordiphosa*. Loreto *et al.* (2001) sugeriram um evento de transferência horizontal para explicar a presença de seqüências *P* altamente similares à seqüência *P* canônica em *D. mediopunctata* e Castro & Carareto (2004) propuseram a transferência de elementos *P* canônicos entre espécies do grupo *saltans*.

Também já foram encontradas fortes evidências de transferência horizontal do elemento *mariner* em espécies de insetos. O primeiro relato foi apresentado por Maryuama & Hartl (1991), entre *D. mauritiana* e *Zaprionus tuberculatus*. Dados moleculares mostraram que a seqüência *mariner* destas duas espécies distantemente relacionadas possuem uma similaridade de 97%, enquanto que a similaridade entre as seqüências *mariner* de *D. mauritiana* e a sua espécie irmã *D. tsacasi* foi de apenas 92%. Outro relato de transferência horizontal envolvendo o elemento *mariner* foi descrito por Robertson & Lampe (1995). Esses autores mostraram que existe alta similaridade de seqüências *mariner* (cerca de 78% em nível de DNA e 90% em nível de aminoácidos) entre abelhas (*Apis mellifera*), forficula (*Forficula auricularia*) e mosca doméstica (*Ceratitis capitata*). Também seria necessário pelo menos dois eventos de transferência horizontal para explicar a distribuição das diferentes subfamílias de *mariner* dentro do subgrupo *melanogaster* (Capy *et al.*, 1992; Brunet *et al.*, 1996). Outra ocorrência de inconsistência filogenética, unido a uma distribuição descontínua da seqüência

mariner-like, é relatada por Lohe *et al.* (1995) entre a espécie *D. erecta* e *Ctenocephalides felis* (pulga).

Além desses clássicos exemplos de transferência horizontal (*P* e *mariner*) entre insetos, existem ainda evidências de transferência lateral dos elementos transponíveis *Minos* e *hobo*, e dos retrotransposons *gypsy*, *copia* e *jockey* entre espécies de *Drosophila*. No caso do elemento transponível *Minos*, Arcà & Savakis (2000) mostraram existir maior similaridade de seqüências entre elementos *Minos* de *D. mojavensis* (grupo *repleta*) e de *D. saltans* (grupo *saltans*) do que entre *D. mojavensis* e *D. hydei* (ambas do grupo *repleta*). Essa alta similaridade entre *D. mojavensis* e *D. saltans* é discrepante quando comparada ao tempo de divergência entre as espécies (cerca de 50 milhões de anos). Diante de tais evidências, os autores sugeriram a ocorrência de transferência horizontal para explicar as incongruências entre a filogenia do elemento e a da espécie hospedeira.

Existem poucas informações sobre a distribuição filogenética do elemento *hobo* no gênero *Drosophila*. Até recentemente acreditava-se que sua ocorrência estivesse limitada ao grupo *melanogaster* (Daniels *et al.*, 1990b); entretanto, estudos mais recentes relatam a ocorrência de *hobo* em espécies dos grupos *saltans* (Castro, 2003) e *D. willistoni* (Loreto *et al.*, 1998) de *Drosophila*. Apesar desta distribuição restrita, eventos de transferência horizontal são sugeridos por Simmons (1992) entre as espécies *D. melanogaster*, *D. simulans* e *D. mauritiana*. O autor se baseia em duas observações para sugerir tal fenômeno. A primeira é a ausência do elemento *hobo* em antigas linhagens de laboratório de *D. melanogaster*, coletadas na década de 50, o que, do mesmo modo que aceito para o elemento *P*, é indicativo de introdução recente do elemento. A segunda refere-se à distribuição não-homogênea de *hobo* entre as espécies do subgrupo *melanogaster* (ausência destas seqüências em *D. teissieri*, *D. yakuba* e *D. erecta*) e à baixa divergência de seqüências entre *D. melanogaster*, *D. simulans* e *D. mauritiana*. Apesar de Simmons (1997) sugerir transferência horizontal como o evento mais parcimonioso, ele não descarta a possibilidade desses fatos serem explicados por eventos de transferência vertical somados a taxas de evolução desiguais e perda estocástica.

A ampla distribuição do retroelemento *gypsy* no gênero *Drosophila* foi atribuída, inicialmente, à existência de elementos ancestrais que foram transmitidos verticalmente ao longo da história evolutiva das espécies do gênero. Porém, análises mais detalhadas mostraram que existem várias seqüências *gypsy* evoluindo independentemente no genoma, levando a uma alta heterogeneidade de seqüências. As similaridades entre estas seqüências são, muitas vezes, discrepantes em relação ao tempo de divergência das espécies hospedeiras, e impossível de serem explicadas por meio de simples eventos de transferência vertical.

Alberola & DeFrutos (1993) sugeriram o evento de transferência horizontal para explicar a alta similaridade das seqüências encontradas entre *D. subobscura* (grupo *obscura*, subgênero *Sophophora*) e *D. virilis* (grupo *virilis*, subgênero *Drosophila*). Posteriormente, Vásquez-Manrique *et al.* (2000) sugeriram novamente o evento de transferência horizontal para explicar a incongruência entre a filogenia de *gypsy* e a filogenia dos hospedeiros *D. hydei* (grupo *repleta*, subgênero *Drosophila*) e *D. virilis* (grupo *virilis*, subgênero *Drosophila*), cujas seqüências se agrupam com as das espécies do distante subgrupo *affinis* (subgênero *Sophophora*). Heredia (2002), em um estudo mais detalhado, sugeriu a ocorrência de ao menos nove eventos de transferência horizontal para explicar a distribuição das diferentes subfamílias *gypsy* no gênero *Drosophila*. Essas teriam ocorrido dentro do mesmo subgênero, entre subgêneros ou mesmo entre gêneros diferentes. No subgênero *Sophophora* teriam ocorrido entre *D. nebulosa* (grupo *nebulosa*) e *D. neocardini* (grupo *cardini*), entre *D. nebulosa* (grupo *nebulosa*) e *D. paulistorum* (grupo *willistoni*) e, no subgênero *Drosophila*, entre *D. pallidipennis* (grupo *pallidipennis*) e *D. bandeirantorum* (grupo *tripunctata*), entre *D. virilis* (grupo *virilis*) e *D. hydei* (grupo *repleta*), entre *D. mediopicta* (grupo *tripunctata*) e *D. zotti* (grupo *repleta*). Os eventos entre diferentes subgêneros teriam ocorrido entre *D. subobscura* (grupo *obscura*, subgênero *Sophophora*) e *D. buskii* (grupo *buskii*, subgênero *Dorsilopha*); entre *D. subobscura* (grupo *obscura*, subgênero *Sophophora*) e *D. hydei* (grupo *repleta*, subgênero *Drosophila*) e, entre diferentes gêneros, entre *D. simulans* (subgênero *Sophophora*) e *S. latifasciaeformis* (gênero *Scaptodrosophila*) e entre *D. simulans* (subgênero *Sophophora*) e *Zaprionus indianus* (gênero *Zaprionus*).

Outro estudo, realizado por Jordan *et al.* (1999), mostrou que a similaridade das LTRs (longas repetições terminais) do elemento *copia* entre *D. melanogaster* e *D. willistoni* (99%) é superior à similaridade das seqüências *copia* de *D. melanogaster* e de sua espécie críptica *D. simulans* (90%). Para os autores, esse é um forte indício da ocorrência de transferência horizontal. Porém, Jordan *et al.* (1999) sugerem cautela e justificam que essa incongruência também poderia ser explicada por seqüências polimórficas transmitidas verticalmente ao longo da história evolutiva.

Mizrokhi & Maza (1990) analisaram a distribuição do retroelemento *jockey* nas espécies do gênero *Drosophila* e mostraram que estas seqüências estavam restritas às espécies do subgrupo *melanogaster* e a *D. funebris*. Um minucioso estudo destas seqüências revelou a existência de duas cadeias abertas de leitura altamente similares entre os elementos *jockey* de *D. melanogaster* e *D. funebris*. O alto grau de similaridade e a ausência de seqüências *jockey*

em espécies crípticas de *D. funebris* levou os autores a sugerirem transferência horizontal para explicar a história evolutiva destas seqüências no gênero *Drosophila*.

Transferência horizontal em plantas

A história evolutiva dos elementos transponíveis em plantas é preferencialmente explicada por eventos de transferência vertical associados à existência de polimorfismos de seqüências ancestrais e perda estocástica (Stuart-Rogers & Flavell, 2001). Ao contrário do observado em bactérias e em invertebrados, raros eventos de transferência horizontal vêm sendo sugeridos para explicar as inconsistências encontradas entre as filogenias dos elementos móveis e as das plantas. Dentre esses, destacam-se os eventos de transferência horizontal em *Arabidopsis thaliana*, *Zea mays* e *Gossypium*. As seqüências *Ta* (subfamílias *Ta1* a *Ta7*) de *A. thaliana* possuem maior similaridade com as seqüências *Tnt1* de tabaco do que com as outras subfamílias *Ta* presentes no genoma da planta (Grandbastien, 1992). O mesmo ocorre com *Z. mays*. As seqüências do elemento da subfamília *Maize1* possui maior similaridade com o elemento *Maize* de *A. thaliana* do que com a subfamília *Maize2* encontrada em *Z. mays*; por sua vez, *Maize2* apresenta maior similaridade com seqüências *Maize* de *Oriza sativa* (Voytas *et al.*, 1992). Finalmente, espécies diplóides e poliplóides de algodão, separadas há mais de 6 milhões de anos, apresentam seqüências *copia-like* mais semelhantes entre elas do que dentro da própria espécie (Vanderwiel *et al.*, 1993).

Transferência horizontal em vertebrados

Poucos casos de transferência horizontal de TEs são descritos em vertebrados, isto provavelmente é consequência de uma maior incidência de elementos da Classe I e, também ao fato desse tipo de pesquisa ser, ainda, pouco freqüente nessa classe.

Dentre os eventos de transferência horizontal já descritos em vertebrados destaca-se o do elemento *Bov-B LINE*. Esse TE, inicialmente descoberto em ruminantes (Lenstra *et al.*, 1993), foi recentemente identificado em algumas espécies de cobras (Kordis & Gubensek, 1997). O fato de sua distribuição ser descontínua em ambos os grupos (Squamata e Ruminantia), somado ao fato dessas seqüências apresentarem alta similaridade de nucleotídeos, levou Kordis & Gubensek (1998, 1999) a proporem a transferência horizontal como a hipótese mais parcimoniosa para explicar a discrepância entre a filogenia do TE e dos seus hospedeiros. Os autores propuseram que o ancestral das cobras *Colubridae* foi o possível doador das seqüências *Bov-B LINE* para *Ruminantia*, e que a transferência ocorreu à cerca de 40-50 milhões de anos. A relação filogenética de *Bov-B LINE* em várias espécies de

Squamata concorda com a filogenia das espécies, sugerindo que *Bov-B LINE* tem sido transmitido verticalmente desde a origem de *Squamata* na era Mesozóica.

Outros exemplos de transferência horizontal em vertebrados envolvem espécies de três classes diferentes. Destacam-se dentre eles os eventos de transferência entre espécies dos elementos *Tcl-like* e *mermaid*. O elemento *Tcl-like*, da superfamília *mariner-Tcl*, é amplamente distribuído no genoma de animais, chegando a representar 1% do genoma total em peixes e anfíbios. Leaver (2001), estudando estas seqüências móveis no peixe marinho *Pleuronectes platessa*, no salmão *Salmo solar* e na rã *Rana temporaria*, observou que *Tcl-like* apresentam seqüências muito conservadas em relação ao longo tempo de divergência das espécies (400 milhões de anos). O alto grau de similaridade levou o autor a sugerir transferência horizontal para explicar a história evolutiva dessas seqüências. O elemento *mermaid*, da família SINE, foi encontrado no genoma de peixes, anfíbios e primatas (inclusive no homem), mas não em ratos, cobras e aves (Shimoda, 1996). Essa distribuição descontínua, identificada por análises de PCR e *Southern blot*, associada à significativa similaridade de seqüências entre os diferentes vertebrados, levaram os autores a sugerirem o evento de transferência horizontal como a explicação mais parcimoniosa para história evolutiva de *mermaid*.

Transferência horizontal entre reinos

O retrotransposon *Ta* fornece a melhor evidência de transferência horizontal de TEs entre reinos de espécies. Konteczny *et al.* (1991) e Grandbastein (1993) mostraram que as seqüências *Ta* das subfamílias *Ta8-Ta10* de *Arabidopsis thaliana* são mais similares às seqüências dos elementos transponíveis *copia* e *1731* de *Drosophila* do que às outras subfamílias *Ta* (*Ta 1-Ta7*) presentes no genoma dessa planta. Outro exemplo é o do elemento *Bar* de cevada (*Hordeum vulgare*). A subfamília *Bar 29* apresenta maior similaridade de seqüência com *Aprt12* de *Aedes aegypti* do que com as outras subfamílias *Bar* daquela planta.

O elemento *gypsy*, como dito anteriormente, é amplamente distribuído entre os organismos e por isso, acredita-se, que tenha sido principalmente transferido verticalmente ao longo da história evolutiva. Porém, exceções em insetos foram citadas anteriormente. Miller *et al.* (1999), em um amplo estudo sobre a distribuição dessas seqüências em vertebrados observaram a presença de *gypsy* em uma ampla variedade de espécies, com exceção de aves e mamíferos. Comparando as seqüências obtidas com as demais seqüências *gypsy* contidas no GenBank, Miller *et al.* (1999) concluíram que as seqüências de vertebrados eram mais similares às de fungos que às de insetos, sugerindo o evento de transferência horizontal para



explicar tal discordância entre as filogenias dos hospedeiros e a de *gypsy*. Também Poulter & Butler (1998), analisando seqüências *gypsy-like* no peixe *Fugu rubripes*, sugeriram a transferência horizontal entre vertebrados e fungos como o evento mais provável para explicar a descontinuidade entre a filogenia do TE e a das espécies hospedeiras.

Transferência horizontal em humanos

Segundo o Consórcio Internacional do Sequenciamento Humano centenas de genes humanos foram originados por transferência horizontal de bactérias e provavelmente dezenas desses são derivados de elementos transponíveis. Como visto anteriormente, os TEs correspondem a 45% do total do genoma humano, podendo ser quase completamente agrupados em quatro tipos: SINEs (Short Interspersed Nuclear elements); LINEs (Long Interspersed Nuclear elements); retrotransposons com LTR e transposons de DNA. Essas seqüências possuem diferentes estratégias que as habilitam sobreviver ao longo do tempo evolutivo em seus hospedeiros. Por exemplo, os LINEs e os SINEs, que correspondem a 75% dos TEs presentes em humanos, se transpõem quase exclusivamente verticalmente ao longo das gerações, sendo raramente transferidos por transferência horizontal. Os retrotransposons com LTRs, que correspondem a aproximadamente 19% dos TEs humanos, podem ser transmitidos verticalmente e, em menor freqüência, horizontalmente, ao longo do tempo evolutivo. Os transposons de DNA, que correspondem a apenas 6% dos TEs do genoma humano, se transpõem exclusivamente pelo mecanismo *cut and paste*, o qual limita seu número e tempo de vida em seus hospedeiro, apresentam alta freqüência de transferência horizontal. Sendo assim, apenas os TEs de DNA que são transferidos horizontalmente para genomas hospedeiros livres dessas seqüências garantem sua sobrevivência ao longo do tempo evolutivo. Dessa forma, imagina-se que apenas uma pequena parcela dos TEs surgiram na espécie humana por meio de transferência horizontal.

Outro dado interessante obtido pelo Consórcio foi a identificação de 47 genes humanos resultantes de elementos transponíveis (domesticação), sendo que 43 desses são TEs de DNA que correspondem à menor parcela de TEs no genoma humano (International Human Genome Sequencing Consortium, 2001). Dessa forma a presença de TEs de DNA no genoma humano parece estar também relacionada ao fenômeno de transferência horizontal bem como com ao domesticação.



VII. Importância evolutiva dos elementos transponíveis no genoma

Os TEs desempenham um papel evolutivo muito importante no genoma hospedeiro, independente do modo pelo qual foram adquiridos (transferência horizontal, hibridização ou por recombinação). Essas seqüências além de acarretarem efeitos deletérios, devido à inserção em genes e regiões regulatórias, ou neutros, quando se inserem em regiões não-codificantes, também podem produzir mutações e mudanças genéticas que são significantes contribuições para a evolução de seus hospedeiros. Na verdade, um número cada vez maior de evidências sugere que a atividade dos elementos transponíveis pode influenciar o fenótipo dos hospedeiros de muitas maneiras: alguns dos seus efeitos são imediatos, diretos e óbvios; outros, entretanto, são sutis e podem demorar longos períodos de tempo para se manifestar completamente (Kidwell & Lish, 2001).

Como efeito imediato temos o aumento da variabilidade, que serve de material molde para as forças evolutivas. Nesse caso, é o produto da atividade do TE que fornece vantagem adaptativas. Entretanto, deve-se ainda lembrar que os TEs carregam consigo uma preciosa maquinaria enzimática que pode ser, potencialmente, utilizada pelo hospedeiro. Esse recrutamento de função não é detectado facilmente, isto porque, a domesticação geralmente acontece muito tempo depois da invasão do TE no genoma hospedeiro, sendo assim, esses elementos podem ter perdido as características que os distinguiam como seqüências móveis.

Seqüências transponíveis recrutadas podem exercer diferentes funções no genoma, tais como, regulação da expressão gênica, expressão de enzimas e outras proteínas necessárias ao hospedeiro. Um recente exemplo de um TE se tornando parte integral da regulação dos genes hospedeiros é fornecido pela apolipoproteína (a) em humanos (Yang *et al.*, 1998). O ativador desse gene está localizado 20 kb acima do sítio de início da transcrição, localizado inteiramente dentro de um elemento LINE. Outro exemplo é dado pelos elementos *MDG412* e *blastopia* que estão inseridos dentro do primeiro *exon* do gene responsável pela produção da enzima arilalquilamina N-acetil transferase (aaNAT1), em *Drosophila*, e potencializam a expressão dessa.

Exemplos adicionais de recrutamento da função do TE pelo genoma hospedeiro são ilustrados pelos retroelementos *HetA* e *TART* em *D. melanogaster*. Os telômeros são blocos terminais dos cromossomos que têm seu comprimento determinado por um equilíbrio dinâmico entre a perda de nucleotídeos, devida à replicação incompleta ou degradação terminal, e adição, devida à atividade da enzima telomerase. Os retrotransposons *Het-A* e *TART* estão presentes na região terminal dos cromossomos de *D. melanogaster* e parecem



desempenhar um tipo de atividade telomerásica, balanceando a perda de seqüências terminais através da atividade transposicional.

Um exemplo de recrutamento de TEs para a realização de funções enzimáticas é dado pelos genes *rag1* e *rag2*, que produzem as enzimas RAG1 e RAG2 que são requeridas pelo sistema V(D)J do sistema imune dos vertebrados. Tem sido evidenciada uma incrível semelhança entre o modo de recombinação destas seqüências e o modo de transposição dos elementos transponíveis (Lewis & Wu, 1997; Spanopoulou *et al.*, 1996; Agrawal *et al.* 1998; Hiom *et al.*, 1998) o que sugere que o sistema V(D)J poderia ter sua origem a partir de um elemento transponível ancestral.

VIII. Hipóteses alternativas para as incongruências filogenéticas

Como vimos nos exemplos acima, a transferência gênica horizontal vem sendo regularmente invocada para explicar: (i) as inconsistências entre as filogenias dos TEs e de suas espécies hospedeiras; (ii) a distribuição descontínua do elemento e (iii) a alta similaridade de seqüências transponíveis entre espécies distantemente relacionadas. Porém, deve-se lembrar que existem outros eventos, muitas vezes igualmente ou mesmo mais parcimoniosos, para explicar tais incongruências. Esses eventos podem ser a perda estocástica do elemento em algumas linhagens, polimorfismo ancestral, diferentes taxas de evolução de TEs entre hospedeiros, conservação de sítios catalíticos essenciais aos TEs, entre outros (Figura 5 A a D, respectivamente).

Apesar de existirem fortes evidências para a distribuição descontínua de um TE em certas espécies, é difícil provar que esse elemento realmente não esteja presente na espécie. Isto porque, TEs de uma mesma subfamília podem apresentar cerca de 50% de similaridade com outras subfamílias. Assim, oligonucleotídeos ou mesmo sondas específicas para uma subfamília podem não detectar outra. Esse é o caso do elemento *mariner*: sondas da subfamília *mauritiana* não detectam seqüências da subfamília *mellifera* (Capy *et al.*, 1998). Do mesmo modo, sondas internas do elemento *P* canônico que não detectam elementos *P* em espécies dos subgrupos *cordata* e *elliptica* do grupo *saltans* de *Drosophila* (Daniels *et al.*, 1990a). Daniels & Strausbaugh (1986) sugeriram que o elemento *P* pode ter sido perdido estocasticamente nesses dois grupos de espécies (Figura 5A). Entretanto a utilização de oligonucleotídeos degenerados e sondas que hibridizam em elementos muito divergentes e rearranjados propiciam a identificação de TEs divergentes. Utilizando essa estratégia (Castro

& Carareto, 2004) demonstraram a ocorrência de elementos *P* divergente em espécies dos subgrupos *cordata* e *elliptica*.

A evolução de uma família de elementos transponíveis dentro de espécies é diferente da evolução dos genes de cópia única do hospedeiro. Por exemplo, dado elemento pode apresentar diferentes mecanismos moleculares controlando as taxas de transposição e excisão e, portanto, pode evoluir em taxas diferentes daquelas que evoluem os genes dos hospedeiros. Também, os mecanismos de repressão de cada hospedeiro podem ser diferentes. Por exemplo, o grau de disgenesia híbrida em insetos decorrente da falha dos mecanismos de repressão dos elementos que a causam pode ser variado entre as diferentes espécies (Capy *et al.*, 1994b). Também, os mecanismos que controlam o aumento do número de cópias são diferentes entre as espécies hospedeiras. Esse é, por exemplo, o caso de *D. simulans* e *D. melanogaster*, pois a média de inserções de TEs é inferior na primeira em relação à segunda espécie (Vieira *et al.*, 1999).

Também, deve-se lembrar que os TEs variam de sequência dentro da mesma espécie (subfamílias podem ter variações de mais de 30% em suas sequências) e, portanto, todas suas cópias, dentro do mesmo hospedeiro, não são idênticas. Portanto, é possível que diferentes espécies, originadas de um ancestral comum, possuam diferentes sequências. Estas sequências, por sua vez, irão evoluir independentemente, dentro, e entre as diferentes espécies. Desta forma, o polimorfismo ancestral pode explicar as complexas relações filogenéticas observadas nas comparações das filogenias dos TEs e seus hospedeiros (Figura 5B).

Finalmente, outra característica dos TEs que afeta a análise de relações filogenéticas é que suas diferentes partes podem evoluir em diferentes taxas, de acordo com suas funções (Figura 5C). Por exemplo, regiões codificadoras da transposase (essenciais para sua mobilização) estão sob forte pressão seletiva (Figura 5D) e devem evoluir a uma taxa menor quando comparada às outras partes do elemento (Capy *et al.*, 1994a).



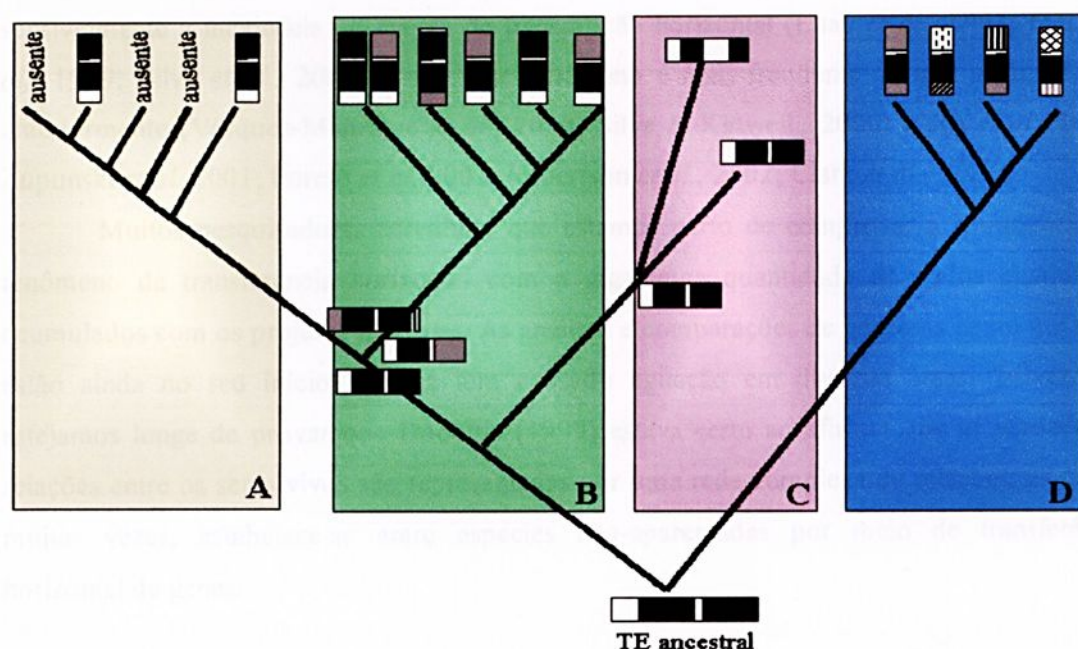


Figura 5. Representação esquemática das hipóteses alternativas para as incongruências filogenéticas: (A) perda estocástica do elemento em algumas linhagens; (B) polimorfismo ancestral; (C) diferentes taxas de evolução de TEs entre hospedeiros e (D) conservação de sítios catalíticos essenciais aos TEs.

IX. Conclusões

Até a primeira metade do século XX, a ciência tinha o genoma como entidade estática e esse dogma contribuiu para estabelecimento dos fundamentos da Genética moderna. Hoje, entretanto, sabemos que o genoma é instável e que essa instabilidade é a principal responsável pela rápida adaptação dos organismos às mudanças ambientais. Entre os responsáveis pela instabilidade genômica estão os elementos transponíveis, seqüências que há décadas atrás foram consideradas como DNA lixo ou parasitas genômicos. Entretanto, esses elementos são reconhecidos como um dos principais fatores envolvidos na evolução dos genomas.

Inúmeros trabalhos, desde a descoberta dos TEs na década de 1940, por Barbara McClintock, corroboram a incrível capacidade de transposição e de promoção de instabilidade genômica desses elementos. Porém, muito permanece obscuro no campo das pesquisas sobre eventos de transferência horizontal dessas seqüências genéticas.

Como seqüências autônomas, os TEs precisam se multiplicar em nível genômico e populacional para se manterem ao longo das gerações. Nesse contexto, a idéia de a transferência horizontal ser uma estratégia que habilita aos elementos transponíveis fugirem da inativação vertical e extinção vem crescendo no meio científico. Evidências mostram que o

modo de transposição e a auto-regulação do número de cópias no genoma hospedeiro afetam sensivelmente a incidência do evento de transmissão horizontal (Luan *et al.*, 1993; Malik *et al.*, 1999; Silva *et al.*, 2004) e que esse fenômeno é mais freqüente do que se acreditava anteriormente (Vásquez-Manrique *et al.*, 2000; Silva & Kidwell, 2000; Koga *et al.*, 2000; Zupunski *et al.* 2001; Loreto *et al.*, 2001; Robertson *et al.*, 2002; Clark *et al.*, 2002).

Muitos pesquisadores acreditam que estamos perto de comprovar a abundância do fenômeno da transferência horizontal com a gigantesca quantidade de dados atualmente acumulados com os projetos genomas. As análises e comparações de genomas como um todo estão ainda no seu início, mas já tem causado agitação em diversas áreas. Talvez não estejamos longe de provar que Doolittle (1999) estava certo ao afirmar que as verdadeiras relações entre os seres vivos são representadas por uma rede complexa de relações, as quais, muitas vezes, estabelece-se entre espécies não-aparentadas por meio de transferência horizontal de genes.

BRITTON, R. J. Mobile elements inserted in the distant past have taken on important functions. *Genet.*, v. 205, p. 177-182, 1997.

BROOK, P.; DOOLITTLE, R. F. Proposed acquisition of an animal protein domain by bacteria. *Proc. Natl. Acad. Sci. USA*, v. 89, p. 8990-8994, 1992.

BRIANET, P. *et al.* The *mariner* transposable element in natural populations of *Drosophila dentissima*. *J. Mol. Evol.*, v. 42, p. 669-673, 1996.

CAPY, P. *et al.* Dynamics and Evolution of Transposable Elements. I^o. France: Landes Bioscience, 1998. 197p.

CAPY, P.; ANXOLABÉHERE, D.; LANGIN, T. The strange phylogenies of transposable elements: are horizontal transfers the only explanation? *Trends Genet.*, v. 10, p. 7-12, 1994.

CAPY, P. *et al.* Horizontal transmission versus ancient origin: mariner in the wicker box. *Genetica*, v. 93, p. 161-170, 1994b.

CAPY, P.; DAVID, J. R.; HARTL, D. L. Evolution of the transposable element *mariner* in the *Drosophila melanogaster* species group. *Genetica*, v. 86, p. 37-46, 1992.

CARLSON, T. A.; CHELM, B. K. Apparent eukaryotic origin of glutamine synthetase II from the bacterium *Bradyrhizobium japonicum*. *Nature*, v. 322, p. 568-570, 1986.

CLARK, J. B.; MADDISON, W. P.; KIDWELL, M. G. Phylogenetic analysis supports horizontal transfer of *P* transposable elements. *Mol. Biol. Evol.*, v. 11, p. 40-50, 1994.

CLARK, J. B.; SILVA, J.; KIDWELL, M. G. Evidence of horizontal transfer of *P* transposable elements. In: Horizontal gene transfer. M. Syvanen, and C.L. Kado, eds. Academic Press, San Diego, p. 161-171, 2002.

CASTRO, J. Revelação da distribuição, diversidade de seqüência, número de cópias e atividade transcricional de elementos transponíveis no grupo *saltans* de *Drosophila*, com ênfase no elemento *P*. *Tese de Doutorado*, Universidade Estadual Paulista, São José do Rio Preto, p.142, 2003.



X. Referências Bibliográficas

- AGRAWAL, A.; EASTMAN, Q.; SCHATZ, D. Transposition mediated by RAG1 and RAG2 and implications for the evolution of immune system. *Nature*, v. 394, p.744-751, 1998.
- ALBEROLA, T. M.; DE FRUTOS, R. Distribution of gypsy sequences in *Drosophila* species of the *obscura* subgroup. *Hereditas*, v. 118, p.131-137, 1993.
- ANXOLABÉHÈRE, D.; KAI, H.; NOUAUD, D. The geographical distribution P-M hybrid dysgenesis in *Drosophila melanogaster*. *Genet. Sel. Evol.*, v. 16, p. 15-26, 1984.
- ARCÀ, B.; SAVAKIS, C. Distribution of transposable element *Minos* in the genus *Drosophila*. *Genetics*, v. 108, p.263-267, 2000.
- BENVENISTE, R. E.; SHERR, C. J.; TODARO, G. J. Evolution of type C viral genes: origin of feline leukemia virus. *Science*, v. 190, p. 886-888, 1975.
- BERG, D. E.; HOWE, M. M. Mobile DNA. American Society for Microbiology, Whashington, D.C., 1989.
- BRITTEN, R. J. Mobile elements inserted in the distant past have taken an important function. *Gene*, v. 205, p. 177-182, 1997.
- BROK, P.; DOOLITTLE, R. F. Proposed acquisition of an animal protein domain by bacteria. *Proc Natl. Acad. Sci. USA*, v. 89, p. 8990-8994, 1992.
- BRUNET, F. *et al.* The *mariner* transposable element in natural populations of *Drosophila teissieri*. *J. Mol. Evol.*, v. 42, p. 669-675, 1996.
- CAPY, P. *et al.* *Dynamics and Evolution of Transposable Elements*. 1st. France: Landes Bioscience, 1998. 197p.
- CAPY, P.; ANXOLABÉHÈRE, D.; LANGIN, T. The strange phylogenies of transposable elements: are horizontal transfers the only explanation? *Trends Genet.*, v. 10, p. 7-12, 1994a.
- CAPY, P. *et al.* Horizontal transmission versus ancient origin: mariner in the witness box. *Genetica*, v. 93, p. 161-170, 1994b.
- CAPY, P.; DAVID, J. R.; HARTL, D. L. Evolution of the transposable element *mariner* in the *Drosophila melanogaster* species group. *Genetica*, v. 86, p. 37-46, 1992.
- CARLSON, T. A.; CHELM, B. K. Apparent eukaryotic origin of glutamine synthetase II from the bacterium *Bradyrhizobium japonicum*. *Nature*, v. 322, p. 568-570, 1986.
- CLARK, J. B.; MADDISON, W. P.; KIDWELL, M. G. Phylogenetic analysis supports horizontal transfer of *P* transposable elements. *Mol Biol Evol*, v. 11, p. 40-50, 1994.
- CLARK, J. B.; SILVA, J.; KIDWELL, M. G. Evidence of horizontal transfer of *P* transposable elements. In: Horizontal gene transfer. M Syvanen, and C.I. Kado, eds. Academic Press, San Diego, p. 161-171, 2002.
- CASTRO, J. Reavaliação da distribuição, diversidade de seqüências, número de cópias e atividade transcricional de elementos transponíveis no grupo *saltans* de *Drosophila*, com ênfase no elemento *P*. *Tese de Doutorado*, Universidade Estadual Paulista, São José do Rio Preto, p.142, 2003.



- CASTRO, J. P.; CARARETO, C. A. *P* elements in *saltans* group of *Drosophila*: a new evaluation of their distribution and number of genomic insertion sites, *Mol. Phylo. Evol.* v. 32, p. 383-387, 2004.
- COFFEY, T. J.; DOWSON, C. G.; DANIELS M. Horizontal spread of an altered penicillin-binding protein 2B gene between *Streptococcus pneumoniae* and *Streptococcus oralis*. *FEMS Microbiol. Lett.*, v. 110, p. 335-339, 1993.
- DANIELS, S. B. *et al.* Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species. *Genetics*, v. 124, p. 339-355, 1990a.
- DANIELS, S. B.; *et al.* Distribution of *hobo* transposable elements in the genus *Drosophila*. *Mol. Biol. Evol.*, v. 7, p. 589-606, 1990b.
- DANIELS, SB; STRAUSBAUGH LD. The distribution of *P* element sequences in *Drosophila*: the *willistoni* and *saltans* species groups. *J. Mol. Evol.*, v.23, p. 138 – 148, 1986.
- DARWIN, C. The origin of species by means of natural selection. John Murray, London. 1859.
- DOOLITTLE, W. F. Phylogenetic classification and the universal tree. *Science*, v. 284, p. 2124-2128, 1999.
- DOOLITTLE, R. H. *et al.* A naturally occurring horizontal gene transfer from a eukaryote to a prokaryote. *J. Mol. Evol.* v. 31, p. 383-388, 1990.
- DOOLITTLE, R. F. *et al.* Origins and evolutionary relationships of retroviruses. *Q. Ver. Biol.*, v. 64, p. 1-29, 1989.
- DRABEK, D. *et al.* Transposition of *Drosophila hydei* Minos transposon in mouse germ-line. *Genomics*, v. 81, p. 108-111, 2003.
- ENGELS, W. R. Hybrid dysgenesis in *Drosophila melanogaster*: rules of inheritance of female sterility. *Genetics*, v. 33, p. 219-236, 1979.
- ENGELS, W. R.; PRESTON, C. R. Hybrid dysgenesis in *Drosophila melanogaster*. The biology of female and male sterility. *Genetics*, v. 92, p. 161-174, 1979.
- FLAVELL, A. J. Long terminal repeat retrotransposons jump between species. *Proc. Natl. Acad. Sci. USA.*, v. 96, p. 12211-12212, 1999.
- FINNEGAN, D. J. Transposable elements in eukaryotes. *Int. Rev. Cytol.*, v. 93, p. 281-326, 1985.
- FRANZ, G.; SAVAKIS, C. Minos, a new transposable element from *Drosophila hydei* is a member de *Tc1*-like family of transposon. *Nucl. Acid Res.*, v. 19, p. 6646, 1991
- FROMAN, B. E.; TAIT, R. C.; GOTTLEIB, L. D. Isolation and characterization of the phosphoglucose isomerase gene from *Echerichia coli*. *Mol. Gen. Genet.*, v. 217, p. 126-131, 1989.
- GRAUR, D.; LI, W. H. Fundamentals of Molecular Evolution, second edition, Sinauer Associates, U.S.A., 2000.
- GRANDBASTEIN, M. A. Retroelements in higher plants. *Trends Genet.*, v. 8, p.103-8, 1992.



- GREEN, M. M. Genetic instability in *Drosophila melanogaster*: de novo induction of putative insertion mutations. *Proc. Natl. Acad. Sci. USA*, v. 74, p. 3490-3493, 1977.
- GREEN, M. M. Genetic instability in *Drosophila melanogaster*: on the identity of the MR and P-M system. *Biol. Zentralbl.* v. 103, p. 1-8, 1984.
- GREEN, M. M. Genetic instability in *Drosophila melanogaster*: the genetics of na MR element that makes complete P insertion mutations. *Proc. Natl. Acad. Sci. USA*, v. 83, p. 4533-4537, 1986.
- GROISMAN, E. A.; SAIER, M. H; OCHMAN, H. Horizontal transfer of a phosphatase gene as evidence for mosaic structure of the *Salmonella* genome. *EMBO J.*, v. 11, p. 1309-1316, 1992.
- HAGEMANN, S.; HARING, E.; PINSKER, W. Repeated horizontal transfer of P transposons between *Scaptomyza pallida* and *Drosophila bifasciata*. *Genetica*, v. 98, 43-51, 1996.
- HALL, R. M. Mobile gene cassettes and integrons: moving antibiotic resistance genes in gram-negative bacteria. *Ciba Foundation Symposium*, 207, 192-202, 1997.
- HARING, E.; HAGEMANN, S.; PINSKER, W. Ancient and recent horizontal invasions of drosophilids by P elements. *J Mol Evol.* v. 51, p. 577-586, 2000.
- HARTL, D. L.; LOZOVSKAYA, E.; LAWRENCE, J. G. Nonautonomous transposable elements in prokaryotes and eukaryotes. *Genetica*, v. 86, p. 47-53, 1992.
- HEREDIA, F. Evolução do retroelemento gypsy em espécies de *Drosophila* e *Zaprionus indianus*. *Tese de Doutorado*, Universidade Federal do Rio Grande do Sul, Porto Alegre, p.188, 2002.
- HIOM, K.; MELEK, M.; GELLERT, M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell*, v. 94, p. 463-470, 1998.
- HOUCK, M. A. *et al.* Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. *Science*, v. 253, p. 1125-1128. 1991.
- JORDAN, I. K.; McDONALD, J. F. Evolution of copia retrotransposon in *Drosophila melanogaster* species subgroup. *Mol. Biol. Evol.*, v. 15, p. 1160-1171, 1998.
- JORDAN, I. K.; MATYUNINA L. V.; McDONALD, J. F. Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. *Pro. Natl. Acad. Sci. USA*, v. 96, p. 12621-12625, 1999.
- KAMINSKI, J. M. *et al.* Design of a nonviral vector for site-selective, efficient integration into the human genome. *FASEB J.*, v. 115, p. 49-63, 2002.
- KAPLAN, N. L.; DARDEN, T.; LANGLEY, C. Evolution and extinction of transposable elements in Mendelian populations. *Genetics*, v. 109, p. 459-480, 1985.
- KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: nature and inheritance of P element regulation. *Genetics*, v. 111, p. 337-350, 1985.
- KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: partial sterility associated with embryonic lethality in P-M system. *Genet. Res.*, v. 44, p. 11-28, 1984.
- KIDWELL, M. G. Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, v. 80, p. 1655-1659, 1983.
- LEVITAN, M. Spontaneous chromosomal alterations in *Drosophila robusta*. *Proc. Natl. Acad. Sci. USA*, 48:930-937, 1962.



- KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: the genetics of cytotype determination in a neutral strain. *Genetics*, v. 98, p. 275-290, 1981.
- KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: the relationship between the P-M and I-R interaction system. *Genet. Res.*, v.33, p. 105-117, 1979.
- KIDWELL, M.G. Transposable elements and the evolution of the genome size in eukaryotes. *Genetica*, v. 115, p. 49-63, 2002.
- KIDWELL, M. G.; LISCH, D.R. Perspectives: transposable elements, parasitic DNA, and genome evolution. *Evolution*, v.55, p. 1-24, 2001.
- KIDWELL, M. G.; LISCH, D.R. Transposable elements and host genome evolution. *Trends Ecol. Evol.*, v. 15, p. 95-99, 2000.
- KIDWELL, M. G.; LISCH, D. R. Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci. USA*, v.94, p. 7704-7711, 1997.
- KIDWELL, M. G.; KIDWELL, J. F.; SVED, J. A. Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics*, v. 86, p. 813-833, 1977.
- KLINAKIS, A. G. *et al.* Genome-wide insertional mutagenesis in human cells by *Drosophila* mobile element *Minos*. *EMBO*, v. 1, p. 416-421, 2000.
- KOGA, A. *et al.* Evidence for recent invasion of the medaka fish genome by the *Tol2* transposable element. *Genetics*, v. 155, p. 273-281, 2000.
- KONTECZNY, A. *et al.* A superfamily of *Arabidopsis thaliana* retrotransposon. *Genetics*, v. 127, p. 801-809, 1991.
- KORDIS, D.; GUBENSEK, F. Molecular evolution of *BOV-B LINEs* in vertebrates. *Gene*, v. 238, p. 171-178, 1999.
- KORDIS, D.; GUBENSEK, F. Unusual horizontal transfer of a long interspersed nuclear element between distant vertebrate classes. *Proc. Natl. Acad. Sci. USA*, v. 95, p. 10704-10709, 1998.
- KORDIS, D.; GUBENSEK, F. *Bov-B* long interspersed repeated DNA (LINE) sequences are present in *Vipera ammodytes* phospholipase A2 genes and in genomes of Viperidae snakes. *Eur. J. Biochem.*, v. 246, p. 772-779, 1997.
- LAWRENCE, J. G.; HARTL, D. L. Inference of horizontal genetic transfer from molecular data: an approach using the bootstrap. *Genetics*, v. 131, p. 753-760, 1992.
- LAWRENCE, J. G.; OCHMAN, H.; HARTL, D. L. The evolution of insertion sequences within enteric bacteria. *Genetics*, v. 131, p. 9-20, 1992.
- LEAVER, M. J. A family of *Tcl1-like* transposon from the genomes of fishes and frogs: evidence for horizontal transmission. *Gene*, v. 271, p. 203-214, 2001.
- LENSTRA, J. A., BOXTEL, V., ZWAAGSTRA, K. A.; SCHWERIN, M. Short interspersed nuclear element (SINE) sequences of Bovidae. *Animal Genet.*, 24:33-39, 1993.
- LEVITAN, M. Spontaneous chromosome aberrations in *Drosophila robusta*. *Proc. Natl. Acad. Sci. USA*, 48:930-937, 1962.



- LEWIN, B. *Genes V*. New York: Oxford University Press, 1994, 1272p.
- LEWIS, S. M.; WU, G. E. The origins of V(D)J recombination. *Cell*, v. 88, p.159-162, 1997.
- LOHE, A.R. *et al.* Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable element. *Mol. Biol. Evol.*, v. 12, p. 62-72, 1995.
- LORETO, E.L.S. *et al.* *Drosophila mediopunctata* P elements: A new example of horizontal transfer. *The Journal of Heredity*, p. 375-381, 2001.
- LORETO, E.L.S.; BASSO DA SILVA, L.; ZAHA, A.; VALENTE, V.L.S. Distribution of transposable elements in neotropical species of *Drosophila*. *Genetics*, v. 101, p. 153-165, 1998.
- LUAN, D. *et al.* Reverse transcription of *R2Bm* RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell*, v. 72, p. 595-605, 1993.
- MALIK, H. S.; BURKE, W. D.; EICKBUSH, T. H. The age and evolution of non-LTR retrotransposable elements. *Mol. Biol. Evol.*, v. 16, p. 793-805, 1999.
- MARIANI, F. *et al.* Characterization of an *IS*-like element from *Mycobacterium tuberculosis*. *J. Gen. Microbiol.*, v. 175, p. 7863-7868, 1993.
- MARUYAMA, K.; HARTL, D. L. Evidence for interspecific transfer of the transposable element *mariner* between *Drosophila* and *Zaprionus*. *J. Mol. Evol.*, v. 33, p. 514-524, 1991.
- MILLER, K. *et al.* Identification of multiple *gypsy* LTR-retrotransposon lineages in vertebrate genome. *J. Mol. Evol.*, v. 49, p. 358-366, 1999.
- MIZROKHI, L. J.; MAZO, A. M. Evidence for transmission of mobile element *jockey* between distant *Drosophila* species. *Proc. Natl. Acad. USA*, v. 87, p. 9216-9220, 1990.
- NISHIKAWA, M.; SUZUKI, K.; YUSHIDA, K. DNA integration into recipient yeast chromosomes by trans-kingdom conjugation between *Echerichia coli* and *Saccharomyces cerevisiae*. *Curr. Genet.*, v. 21, p. 101-108, 1992.
- O'BROCHTA, D. A. *et al.* Gene vector and transposable element behavior in mosquitoes. *J. Exp. Biol.*, v. 206, p. 3823-3834, 2003.
- O'HARE, K.; RUBIN, G.M. Structure of P element transposable and their sites of insertion and excision in *Drosophila melanogaster* genome. *Cell*, v. 34, p. 25-35, 1983.
- OSTERTAG, E.; HAIG, H.; KAZAZIAN, Jr. Twin priming: a proposed mechanism for creation of inversions in *L1* retrotransposon. *Genome Research*, v. 11, p. 2059-2065, 2001.
- PELISSON, A. *et al.* *Gypsy* transposition correlates with the production of retroviral envelope-like protein under the tissue-specific control of the *Drosophila flamenco* gene. *EMBO J.*, v. 13, p. 4401-4411, 1994.
- PINSKER, W. *et al.* The evolutionary life history of P transposons: from horizontal invaders to domesticated neogenes. *Chromosoma*, v. 110, p. 148-158, 2001.
- POULTER, R.; BUTLER, M. A retrotransposon family from pufferfish (fugu) *Fugu rubripes*. *Gene*, v. 215, p. 241-249, 1998.



- ROBERTSON, H. M.; LAMPE, D. J. Recent horizontal transfer of *mariner* transposable element among and between *Diptera* and *Neuroptera*. *Mol. Biol. Evol.*, v. 12, p. 850-862, 1995.
- ROBERTSON, H. M. *et al.* The *mariner* transposon of animals: horizontally jumping genes. In: Horizontal gene transfer. M. Syvanen and C.I. Kado eds. Academic Press, San Diego, CA, 2002.
- ROWE, W. P.; SATO, H. Genetic mapping of the *Fv-1* locus of the mouse. *Science*, v. 180, p. 640-641, 1973.
- SASAKURA, Y.; AWAZU, S.; CHIBA, S.; SATOH, N. Germ-line transgenesis of *Tc1/mariner* superfamily transposon *Minos* in *Ciona intestinalis*. *PNAS*, v. 100(13), p. 7726-7730, 2003.
- SCARPULLA, R. C. Processed pseudogenes for rat cytochrome-c are preferentially derived from one of three alternative mRNAs. *Mol. Cell. Biol.*, v. 4, p. 2279-2288, 1984.
- SCHAEFER, R. E.; FAUSTO-STERLING, A.; KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: morphological and cytological studies of ovarian dysgenesis. *Genetics*, v. 92, p. 1141-1152, 1979.
- SHIMIZU, K. *et al.* Extrachromosomal transposition of the transposable element *Minos* occurs in embryos of the silkworm *Bombyx mori*. *Insect Mol. Biol.*, v. 9(3), p. 277-281, 2000.
- SHIMODA, N. *et al.* *Mermaid*: a family of short interspersed repetitive elements widespread in vertebrates. *Bioch. Biophys. Res. Commun.*, v. 220, p. 226-232, 1996.
- SIKORSKI, R. S. *et al.* Trans-kingdom promiscuity. *Nature*, v. 345, p. 581-82, 1990.
- SILVA, J. C.; KIDWELL, M. G. Horizontal transfer and selection in the evolution of *P* elements. *Mol Biol Evol.*, v. 17, p. 1542-1557, 2000.
- SILVA, J. C.; LORETO, E.; CLARK, J. B. Factors that affect the horizontal transfer of transposable elements. *Curr. Issues Mol. Biol.*, v. 6, p.57-72, 2004.
- SIMMONS G. M. Horizontal transfer of *hobo* transposable element within the *Drosophila melanogaster* species complex: evidence of DNA sequencing. *Mol. Biol. Evol.*, v. 9, p. 1050-1060, 1992.
- SMITH, M. W.; FENG, D. F.; DOOLITTLE, R. F. Evolution by acquisition: the case for horizontal transfer. *Trends Biochem. Sci.*, v. 17, p. 489-493, 1992.
- SMITH, M. W.; DOOLITTLE, R. F. A comparison of evolutionary rates of the two major kinds of superoxide dismutase. *J. Mol. Evol.*, v. 34, p. 175-184, 1992.
- SPANOPOULOU, E. *et al.* The homeodomain region of RAG1 reveals the parallel mechanisms of bacterial and V(D)J recombination. *Cell*, v. 87, p. 263-276, 1996.
- STUART-ROGERS, C.; FLAVELL, A. The evolution of *Ty-1* group retrotransposon in gymnosperms. *Mol. Biol. Evol.*, v. 18, p. 155-163, 2001.
- SYVANEN, M. Cross-species gene transfer: a major factor in evolution? *Trends Genet.*, v. 2, p. 63-66, 1986.
- VANDERWIEL, P. L.; VOYTAS, D. F.; WENDEL, J. F. *Copia-like* retrotransposable element evolution in diploid and polyploid Cotton (*Gossypium* L.). *J.Mol. Evol.*, v. 36, p. 429-447, 1993.

VÁSQUEZ-MANRIQUE, R. P. *et al.* Evolution gypsy endogenous retrovirus in the *Drosophila obscura* species group. *Mol. Biol. Evol.*, v. 17, p. 1185-1193, 2000.

VIEIRA, C. *et al.* Evolution of genome size in *Drosophila*. Is the invader's genome being invaded by transposable elements? *Mol. Biol. Evol.* V. 19, p. 1154-1161, 2002.

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

VOYTAS, D. F. *et al.* *Copia*-like retrotransposons are ubiquitous among plants. *Proc. Natl. Acad. Sci. USA*, v. 89, p. 7124-7128, 1992.

YANG, Z. *et al.* Apolipoprotein (a) gene enhancer resides within a LINE element. *J. Biol. Chem.*, v. 273, p. 3722-3731, 1998

ZAGORAIOU, L. *et al.* In vivo transposition of *Minos*, a *Drosophila* mobile element, in mammalian tissues. *Proc. Natl. Acad. Sci. USA*, p. 11474-11478, 2001.

ZUPUNSKI, V.; GUBENSEK, F.; KORDIS, D. Evolutionary dynamics and evolutionary history in the RTE clade of non-LTR retrotransposons. *Mol. Biol. Evol.*, v.18, p. 1849-1863, 2001.

Capítulo 2



Identificação de duas subfamílias do elemento transponível *micropia* em populações de *Drosophila* do grupo *repleta*.

IDENTIFICAÇÃO DE DUAS SUBFAMÍLIAS DO ELEMENTO TRANSPONÍVEL *MICROPIA* EM POPULAÇÕES DE *DROSOPHILA* DO GRUPO *REPLETA*.

Resumo: A ocorrência, número de inserções no genoma e a expressão de um RNA antissense do elemento transponível *micropia* foram estudadas em 26 espécies pertencentes a três diferentes subgrupos (*obscura*, *mini* e *repleta*) do grupo *repleta* de *Drosophila*. Sequências da PCR, entretanto sob baixa estridência este retratransposon foi detectado em todas as espécies analisadas. O elemento já estava presente no ancestral do grupo *repleta* de *Drosophila*. Análises de Southern blot mostraram uma variação de 0 a 17 sítios de inserção no genoma e a ocorrência de sequências muito específicas. A expressão do RNA antissense de *micropia* foi variável entre os diferentes tecidos (somos e testículos), o que sugere que mais de um mecanismo regula a transposição do elemento *micropia* nessas espécies. A variação na amplificação por PCR e na expressão do RNA antissense, bem como a divergência na sequência de nucleotídeos entre as espécies permite sugerir a existência de pelo menos duas subfamílias *micropia* no genoma das espécies desse grupo.

Palavras-chave: *Drosophila*, grupo *repleta*, retratransposon copia, número de inserções, transferência horizontal

* In press no periódico *Genética* (www.bjweronline.com/issn/0016-6707/content)

Capítulo 2



Identificação de duas subfamílias do elemento transponível *micropia* em populações de *Drosophila* do grupo *repleta*.

Resumo: A ocorrência, número de inserções no genoma e a expressão de um RNA antisense do elemento transponível *micropia* foram estudadas em 26 espécies pertencentes a três diferentes subgrupos (*mercatorum*, *mulleri* e *hydei*) do grupo *repleta* de *Drosophila*. Sequências *micropia* foram detectadas em 11 espécies, sob condições de alta especificidade da PCR, entretanto sob baixa estrigência este retrotransposon foi detectado em todas as espécies analisadas. A ampla distribuição de *micropia* sugere que este elemento já estava presente no ancestral do grupo *repleta* de *Drosophila*. Análises de *Southern blot* mostraram uma variação de 0 a 17 sítios de inserção no genoma e a ocorrência de sequências macho-específicas. A expressão do RNA antisense de *micropia* foi variável entre os diferentes tecidos (soma e testículos), o que sugere que mais de um mecanismo regula a transposição do elemento *micropia* nessas espécies. A variação na amplificação por PCR e na expressão do RNA antisense, bem como a divergência na sequência de nucleotídeos entre as espécies permite sugerir a existência de pelo menos duas subfamílias *micropia* no genoma das espécies desse grupo.

Palavras-chave: *Drosophila*, grupo *repleta*, retrotransposon *copia*, número de inserções, transferência horizontal

* *In press* no periódico *Genetica* (www.kluweronline.com/issn/0016-6707/contents)





Identification of two subfamilies of *micropia* transposable element in species of the *repleta* group of *Drosophila*

Luciane M. de Almeida & Claudia M.A. Carareto*

Departamento de Biologia, Universidade Estadual Paulista – UNESP, 15054-000 São José do Rio Preto, SP, Brazil; *Author for correspondence (Phone: +55-17-2212382; Fax: +55-17-2212390; E-mail: carareto@bio.ibilce.unesp.br)

Received 17 July 2003 Accepted 6 October 2003

Key words: antisense RNA, *Drosophila*, *micropia*, *micropia* subfamilies, number of insertion sites, *repleta* group, retrotransposon

Abstract

The occurrence, number of insertion sites and antisense RNA expression of *micropia* transposable element was studied in 26 species that belong to three subgroups (*mercatorum*, *mulleri* and *hydei*) of *repleta* group of *Drosophila*. Under high specific PCR, *micropia* sequences were detected in 11 species, but under less stringent condition, this retrotransposon was detected in all species. The widespread distribution of *micropia* suggests that this element was already present at the common ancestor of the *repleta* group of *Drosophila*. Southern blot analysis showed a variation from 0 to 17 different insertion sites and the occurrence of male-specific sequences. We found that the expression of the 1.0 kb *micropia* antisense RNA is variable among the species and tissues (soma and testis), which suggests that more than one mechanism regulates transposition in these species. Variation of amplification by PCR and of antisense RNA expression, as well as divergence of nucleotide sequences among the species allow us to suggest that at least two subfamilies of *micropia* transposable element are harbored by the genome of this species group.

Introduction

Micropia is a member of the widespread *Ty-copia* family of retrotransposon that was first identified in *Drosophila hydei* in lampbrush loops present in the Y chromosome during the prophase of meiosis in primary spermatocytes (Huijser et al., 1988). According to Lankenau, Corces and Lankenau (1994) and Lankenau, Huijser and Hennig (1989) *micropia* is flanked by LTRs and contains one open reading frame that encodes putative products that show homology to the nucleocapsid, protease, reverse transcriptase, RNase H and integrase proteins of vertebrate retroviruses. *Micropia* nomenclature refers to the technique employed in its discovery, microcloning experiments and its similarity to copia-like elements. The words *dm* or *dh* are added to the family name *micropia* depending on its origin, for example,

D. melanogaster or *D. hydei* species, respectively. Then, there are four best-characterized *micropia* elements: *micropia-dm11*; *micropia-dm2*; *micropia-dhMiF₂* and *micropia-dhMiF₈* (Lankenau, 1993).

Micropia expression gives rise to a complex set of sense and antisense RNA (Lankenau, Corces & Lankenau, 1994). One of these transcripts is a 5 kb sense product whose size corresponds to the distance between the transcription initiation signals found in 5'LTR and a polyadenylation site located in 3'LTR that is expressed in both testis and somatic tissues of males and females. In addition to the full-length sense transcript, *micropia* encodes two different antisense RNAs 1.0 and 1.6 kb long that have their expression limited to the testis. According to Lankenau, Corces and Lankenau (1994), the pattern of subcellular accumulation of the sense and antisense RNAs is consistent with the hypothesis that the antisense RNA could

act in the control of the germline expression of the full-length transcript or the *micropia*-encoded reverse transcriptase protein.

Besides the four well studied *micropia* sequences (*micropia-dm11*; *micropia-dm2*; *micropia-dhMiF₂* and *micropia-dhMiF₈*), other homologous sequences have been identified in three of four species group of the *Sophophora* subgenus: *melanogaster* group (*D. simulans*, *D. birchii*, *D. yakuba*, *D. ananassae*) by Lankenau (1993); *willistoni* group (*D. willistoni*) by Lankenau (1993); *saltans* group (*D. saltans*) by Lankenau (1993) and (*D. austrosaltans*, *D. prosaltans*, *D. subsaltans*, *D. parasaltans*, *D. sturtevantii*, *D. dacunhai*, *D. milleri*, *D. neocordata*) by Almeida, Castro and Carareto (2001); and in other groups of *Drosophila* subgenus: *immigrans* group (*D. immigrans*); *funnebris* group (*D. funnebris*); *mellanica* group (*D. mellanica*) and in *repleta* group (*D. repleta*, *D. neorepleta*, *D. paranaensis*, *D. mercatorum*, *D. buzzatii*, *D. mulleri*, *D. eohydei*) by Lankenau (1993).

The first step to understand the evolutionary history of a particular transposable element (TE) is to examine its phylogenetic distribution. The distribution of a TE can be resulted from vertical transmission of a TE present in their common ancestor or from horizontal transfer when the high degree of similarity between TE sequences is impossible to reconcile with the long divergence time of their respective host species (Silva & Kidwell, 2000). On the other hand, in order to understand TE dynamics and estimate their possible effects on evolution, data about their number of insertion sites, genome localization and activity in different species are necessary. To the best of our knowledge, literature revision has shown just one study about *micropia* number of insertion sites (Lankenau, 1993), and only *D. melanogaster* and *D. hydei* were studied. There is also only one study (Lankenau, Corces & Lankenau, 1994) that shows the expression of a sense and antisense RNA in seven species (*D. paranaensis*, *D. mercatorum*, *D. limensis*, *D. buzzatii*, *D. eohydei*, *D. hydei* and *D. melanogaster*).

Micropia has been studied in details only in *D. hydei* and *D. melanogaster* species, but there is very little information about polymorphism of insertion sites, occurrence and activity in other *Drosophila* species. In order to get some insight into the evolutionary history of *micropia* retrotransposon, we have analyzed its phylogenetic distribution, number of insertion sites, and diversity of sequence and of antisense RNA expression in 26 species belonging to three subgroups

(*mercatorum*, *mulleri* and *hydei*) of *repleta* species group of *Drosophila*.

Material and methods

Fly stocks

All species and strains used in the present study are listed in Table 1; each list includes the taxonomic nomenclature, location, and stock number or collection date.

DNA amplification and sequencing

PCR reactions were performed in 25 µl volumes using approximately 200 ng of template DNA, 100 µM of each primer, 1.5 mM of MgCl₂, 2.5 µl of DMSO and 1 U of Taq DNA Polymerase (GIBCO-BRL) in 1× Polymerase buffer. After an initial denaturation for 3 min at 95°C, 40 cycles consisting of 1 min denaturation at 95°C, 1 min annealing at 52°C and 2 min extension at 72°C were carried out. An additional extension step of 10 min at 72°C was performed after the last cycle. The primers used (2813: 5'-TTAACTCCTAGAGTTCATCGCTGG-3' and 2814: 5'-CATGTACCTGGTTAACTACTGACC-3') anneal into *micropia* sequence that encodes eight consecutive amino acids highly conserved in *micropia* elements of *D. melanogaster* and *D. hydei* (Lankenau, Corces & Lankenau, 1994). The fragment amplified has 386 bp and corresponds to nucleotides from 2813 to 3198. The amplified fragments were separated by electrophoresis on 1% agarose gel. PCR products were cleaned and directly sequenced (ABI 377 automated sequencer).

Southern blot analysis

Genomic DNA was prepared from 50 adult flies according to Jowett (1986). DNA samples were digested with *Bam* HI and *Xho* I restriction enzymes to determine the number of insertion sites and with *Xho* I to identify male-specific sequences. DNA fragments were separated by electrophoresis on 1% agarose gels and transferred to nylon membranes (Hybond N+/Amersham-Pharmacia). The membranes were hybridized to a random primer-labeled probe at 60°C in 5× SSC; 0.1% SDS; 5% dextran sulfate, and 20-fold dilution of liquid block. The filters were washed twice with 0.2× SSC and 0.5% SDS for

Table 1. Taxonomic list of *Drosophila* species used in this study (according to Vilela, 1983)

Subgroup	Cluster	Species	Location	Stock number	Date
Mercatorum		<i>D. paranaensis</i>	Novo Horizonte, SP, Brazil		1998
		<i>D. mercatorum</i>	Novo Horizonte, SP, Brazil		1998
		<i>D. aldrichi</i>	Tamaulipas, Mexico	15081125.0	
Mulleri		<i>D. mulleri</i>	Tucson, Arizona, USA	150811371.7	
		<i>D. wheeleri</i>	Catalina Island, Los Angeles, USA	150811501.3	
		<i>D. navojoa</i>	Tehuantepec, Mexico	150811374.1	
Mojavensis		<i>D. arizonae</i>	Tucson, Arizona, USA	A1015	
		<i>D. mojavensis</i>	Grand Canyon, Arizona, USA	A870	
		<i>D. koepferae</i>	Tapia-Tucumam, Argentina		
Mulleri		<i>D. serido</i>	Milagres, BA, Brasil	150811431.3	
		<i>D. buzzatti</i>	Novo Horizonte, SP, Brazil		1998
		<i>D. gouveai</i>	Morro do Chapéu, BA, Brazil		
Buzzatti		<i>D. antonietae</i>	Bela Vista, MS, Brazil		
		<i>D. seriema</i>	Serra do Cipó, MG, Brazil		
		<i>D. mainland</i>	Catalina Island, CA, USA	150811315.0	
Longicornis		<i>D. pachuca</i>	Chapingo, Mexico	150811391.0	
		<i>D. propachuca</i>	Hidalgo, Mexico	150811411.1	
		<i>D. hexastigma</i>	Zapotitlan, Puebla, Mexico	150811302.2	
Longicornis		<i>D. spenceri</i>	Gyaymas, Sonora, Mexico	150811441.2	
		<i>D. longicornis</i>	Tucson, Arizona, USA	150811311.8	
		<i>D. eremophila</i>	Zapotitlan, Puebla, Mexico	150811292.1	
Eremophila		<i>D. mettleri</i>	Tucson, Arizona, USA	CAT397	
		<i>D. nigricruria</i>	Novo Horizonte, SP, Brazil		1998
		<i>D. anceps</i>	Infernillo, Michoacán, Mexico	150815031.1	
Hydei		<i>D. ritae</i>	Zapotitlan dos Salinas, Puebla, Mexico		2002
		<i>D. hydei</i>	Encarnación, Paraguay		2000

15 min at 60°C. Hybridization was performed by using Gene Images (Amersham-Pharmacia) according to the manufacturer's instructions. A 739 bp *micropia* sequence amplified from dhMiF₂ plasmid (Huisjer et al., 1988) was used as a probe. This fragment was obtained by PCR using the primers mRNA₂ (5'-AGGGGAAATTCGACCAAAC-3') and mRNA₁ (5'-AATATCCGGGTCCAATTGCT-3') which anneal from nucleotides 2293 to 2313 and from 3012 to 3032, respectively.

RNA isolation and reverse transcription

Total RNA was isolated from 15 testis and 15 heads using TRIZOL Reagent (GIBCO – BRL Technologies) according to the manufacturer's instructions. The *micropia* cDNA library was generated from total RNA using Superscript II RT (GIBCO – BRL Technologies) and Oligo DT primer at 55°C. PCR amplifications were performed in order to amplify the sense as well

the antisense transcripts. To amplify the sense mRNA the above cited mRNA₁/mRNA₂ primers were used (Figure 1(A)). To amplify the antisense mRNA a first round of amplification was done using the mRNA₁ and an Oligo DT primer. Figure 1(B) illustrates how this combination of primers assures amplification of antisense mRNA exclusively since mRNA₁ and Oligo DT primers have the same polarity of amplification (5'–3') and are not able to produce any fragments from the sense RNA transcript.

PCR was done in a volume of 25 µl with 2.5 units Taq DNA polymerase in 1× polymerase buffer, 20 nmol of each primer, 200 µM of each dNTP and 2 mM of MgCl₂. The solutions were heated to 95°C for 5 min and then cycled through 35 reactions of 95°C for 30 s, 58°C for 30 s, 72°C for 2 min, followed by a final extension at 72°C for 10 min. Another round of PCR amplification using the same primers and conditions was necessary to obtain enough products to be visualized in agarose gels. In order to confirm the

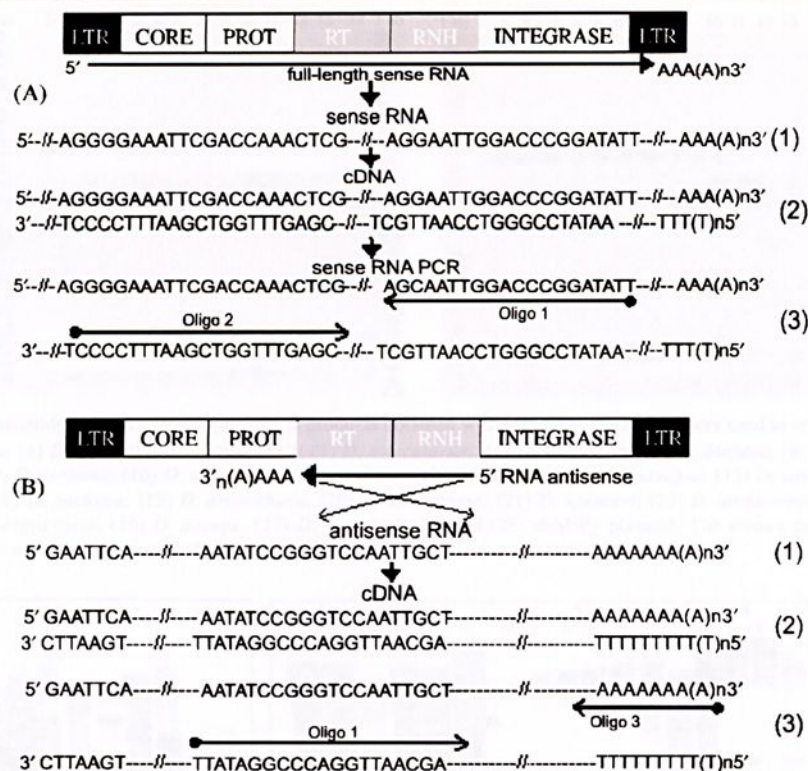


Figure 1. Structure of *micropia* transposable element (LTR, CORE, PRO, RT, RNase H) and diagram of the approach used to amplify the *micropia* mRNAs. (A) *Micropia* sense mRNA sequence (1), cDNA synthesis using Oligo DT primer that anneals in the poli (A) tail (2) and oligonucleotides 1 (mRNA₁: 5'-AATATCCGGGTCCAATTGCT-3') and 2 (mRNA₂: 5'-AGGGGAAATTCGACCAAACCT-3'), which anneals on nucleotides 2293-2313 and 3032-3032. (B) *Micropia* antisense mRNA sequence (1), cDNA synthesis using Oligo DT primer that anneals in the poli (A) tail (2) and oligonucleotides 1 (mRNA₁) and 3 (Oligo DT), which anneals on nucleotides 3012-3032 and poli A tail, respectively. Horizontal arrows indicate the direction of sense and antisense mRNAs transcription.

similarity of PCR products obtained with the *micropia* sequence, the obtained were hybridized to the 734 bp of dhiMF₂.

Phylogenetic analysis

All sequences were compiled into NEXUS files after alignment in CLUSTAL W program (Thompson, Higgins & Gibson, 1994). The phylogenetic relationships among *micropia* sequences were estimated by using the maximum parsimony method, using a branch-and-bound algorithm as implemented in PAUP v.4.0b10 (Swofford, 1997). Branch support was calculated by bootstrap analysis consisting of 500 replicates. The distance matrix used in this work was built according to the Kimura two-parameter method. DNA *micropia* sequences of *D. hydei* (X13304) and *D. melanogaster* (X14037) were obtained from EMBL DNA sequence library.

Results

Distribution among species

DNA amplification by PCR is frequently complicated by problems of low yield and specificity. A common approach to improve such reactions is the addition of small amounts of certain organic chemicals, such as dimethylsulfoxide (Chakrabarti & Schutt, 2001). Under high specific PCR *micropia* element, homologous sequences were detected by PCR in 11 from 26 species tested (Figure 2(A)). The amplified sequence has 386 bp and corresponds to a high conserved sequence of *micropia*. On the other hand, under less stringent PCR (no DMSO was added into the mix), there was weak amplification in those 15 species in which *micropia* had not been previously amplified (Figure 2(B)).

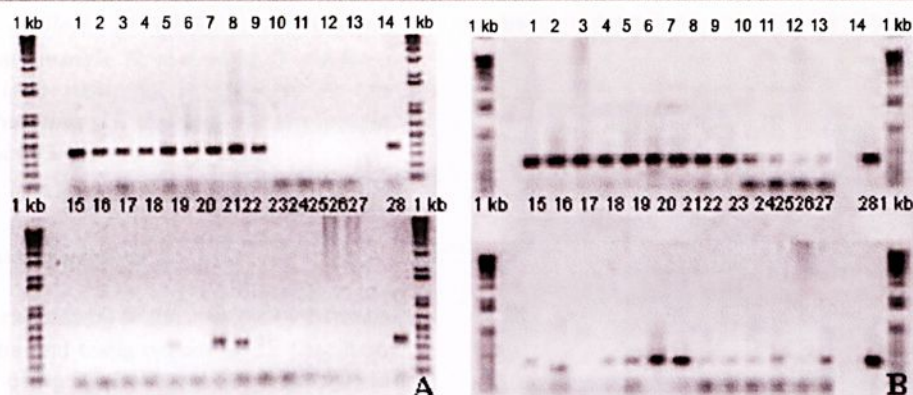


Figure 2. Ethidium bromide-stained agarose gel of PCR products obtained with *micropia*-specific primers used as template genomic DNA from the following species: (1) *D. hydei*; (2) *D. paranaensis*; (3) *D. mercatorum*; (4) *D. koepferae*; (5) *D. buzzatii*; (6) *D. serido*; (7) *D. gouveai*; (8) *D. antonietae*; (9) *D. seriema*; (10) *D. aldrichi*; (11) *D. mulleri*; (12) *D. wheeleri*; (13) *D. navojoa*; (15) *D. arizonae*; (16) *D. mojavenensis*; (17) *D. mainland*; (18) *D. pachuca*; (19) *D. propachuca*; (20) *D. hexastigma*; (21) *D. spenceri*; (22) *D. longicornis*; (23) *D. eremophila*; (24) *D. mettléri*; (25) *D. nigricruria*; (26) *D. anceps*; (27) *D. ritae* and (14) and (28) dhMiF₂ plasmid. The arrows point at the expected 386 bp fragments. (A) Under high stringency and (B) under low stringency.

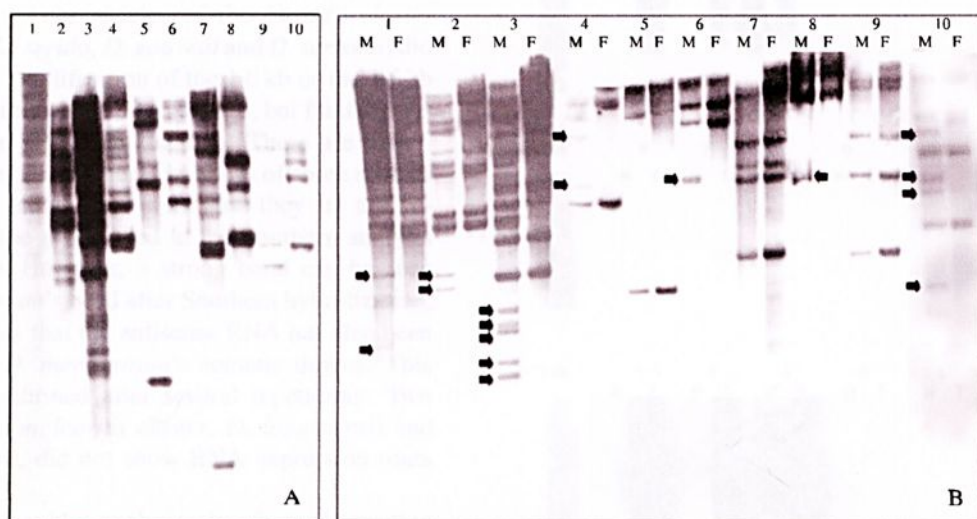


Figure 3. Southern hybridization showing the number of insertion sites (A) and the male-specific fragments of *micropia* element (B) in 10 species of the *repleta* group. (1) *D. hydei*; (2) *D. paranaensis*; (3) *D. mercatorum*; (4) *D. koepferae*; (5) *D. buzzatii*; (6) *D. serido*; (7) *D. gouveai*; (8) *D. antonietae*; (9) *D. seriema* and (10) *D. spenceri*. Arrows indicate male-specific fragments. (A) genomic DNA of 50 individuals digested with *Bam* HI and *Xho* I; (B) genomic DNA of 50 males (M) and 50 females (F) digested with *Xho* I. Both blots were probed with 739 bp fragment amplified from dhMiF₂ plasmid.

Micropia number of insertion sites

Figure 3 shows the number of insertion sites of *micropia* in pools of 50 individuals (Figure 3(A)) and of 50 males and 50 females side by side (Figure 3(B)) in 10 species of the *repleta* group. The strongest signals were detected in *hydei* and *mercatorum* subgroups, with *D. hydei* showing the most intense hybridizing bands. The species surveyed contain variable numbers of *micropia* elements: *D. hydei* has 17 copies; *D. paranaensis* 9; *D. mercatorum* 16; *D. koepferae*

11; *D. buzzatii* 6; *D. serido* 7; *D. gouveai* 11; *D. antonietae* 5; *D. seriema* 7 and *D. spenceri* 9. Although male-specific sequences (Figure 3(B)) have not been detected in three species of the *buzzatii* subgroup (*D. koepferae*, *D. gouveai* and *D. seriema*), most species present these sequences varying from one (*D. hydei*, *D. serido* and *D. antonietae*) to four (*D. mercatorum* and *D. spenceri*). Although the 386 bp fragment has been amplified in *D. longicornis*, it was not detected by the Southern analysis. Also, no *micropia*-hybridizing sequences were detected in

D. aldrichi, *D. mulleri*, *D. whelleri*, *D. navojoa*, *D. arizonae*, *D. mojavenensis*, *D. mainland*, *D. pachuca*, *D. propachuca*, *D. hexastigma*, *D. spenceri*, *D. longicornis*, *D. eremophila*, *D. mettlari*, *D. nigricruria*, *D. nigrospiracula*, *D. anceps*, or *D. ritae*.

Micropia antisense transcripts

We analyzed the expression of the *micropia* retrotransposon in the soma and testis tissues in 12 species of *repleta* group by using RT-PCR and Southern hybridization analysis (Figure 4(A)–(C)). Whereas the sense RNA was detected in all species (data not shown), the expression of antisense RNA was variable among them. Fragments of about 1.0 and 1.6 kb were amplified in testis of *D. hydei*, *D. paranaensis*, *D. mercatorum*, *D. koepferae*, *D. antonietae* and *D. spenceri* (Figure 4(A)). Four species of the *buzzatii* cluster (*D. buzzatii*, *D. serido*, *D. gouveai* and *D. seriema*) did not show any amplification of the 1.0 kb or the 1.6 kb fragments visible on the agarose gels, but faint signals could be seen after hybridization. There are lower-molecular-weight fragments in heads of some species that probably are unspecific because they are not homologous to the probe used in the Southern analysis (Figure 4(B)). However, a strong band can be seen in *D. mercatorum*'s head after Southern hybridization, which indicates that the antisense RNA has also been expressed in *D. mercatorum*'s somatic tissues. This result was confirmed after several repetitions. Two species from *longicornis* cluster, *D. longicornis* and *D. propachuca*, did not show RNA expression (data not shown).

The Southern blot analysis also showed variation in the antisense RNA amplification; it was more intense in species of the *hydei* and *mercatorum* subgroups, and remarkably less intense in the *buzzatii* cluster. To detect the antisense RNA in these species, longer exposition of the blot to the autoradiograph film was necessary (Figure 4(C)). The antisense RNA expression is positively associated with the *micropia* number of insertion sites; for example, species that have higher numbers of *micropia* also have more intense amplification of the antisense RNA transcript. The only exception was *D. spenceri* because, although it had nine copies of *micropia*, this species did not show any signal of antisense RNA hybridization.

The distribution of *micropia* sequences in species of the *repleta* group of *Drosophila*, based on PCR analysis and the expression of the antisense

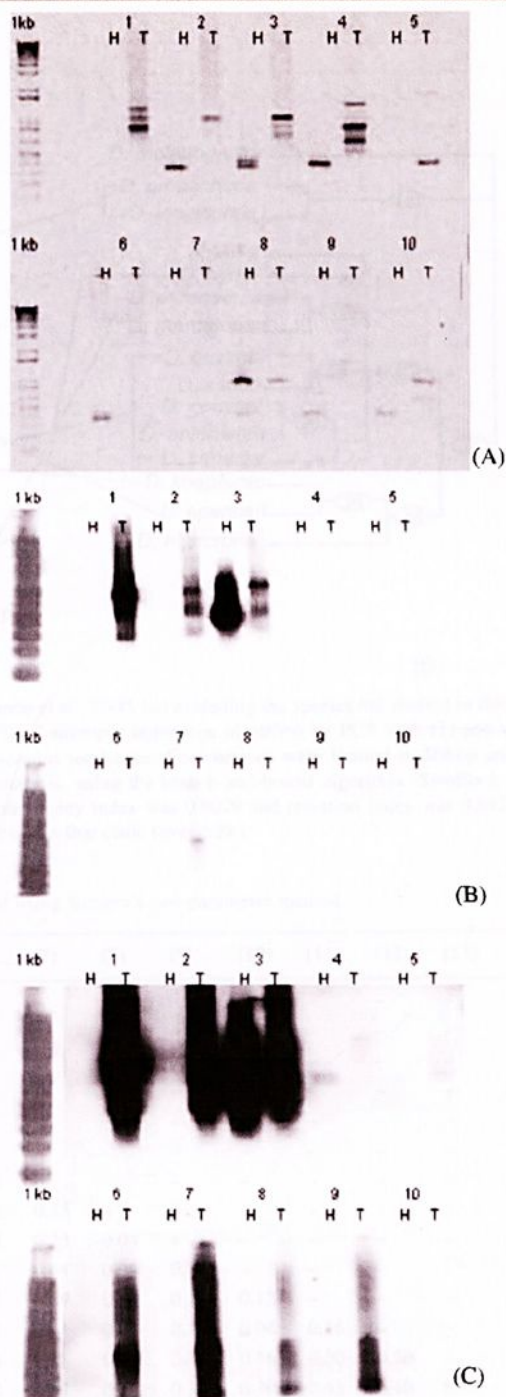


Figure 4. RT-PCR products obtained with *micropia*-specific primers for amplification of antisense RNA (Oligo DT and mRNA₁) using as template genomic cDNA (A) and Southern blot of PCR products using a 739 bp fragment amplified from dhMIF₂ as probe and two (B) and 20 (C) minutes of blot exposition to the autoradiograph film. (1) *D. hydei*; (2) *D. paranaensis*; (3) *D. mercatorum*; (4) *D. koepferae*; (5) *D. buzzatii*; (6) *D. serido*; (7) *D. gouveai*; (8) *D. antonietae*; (9) *D. seriema*; (10) *D. spenceri*.

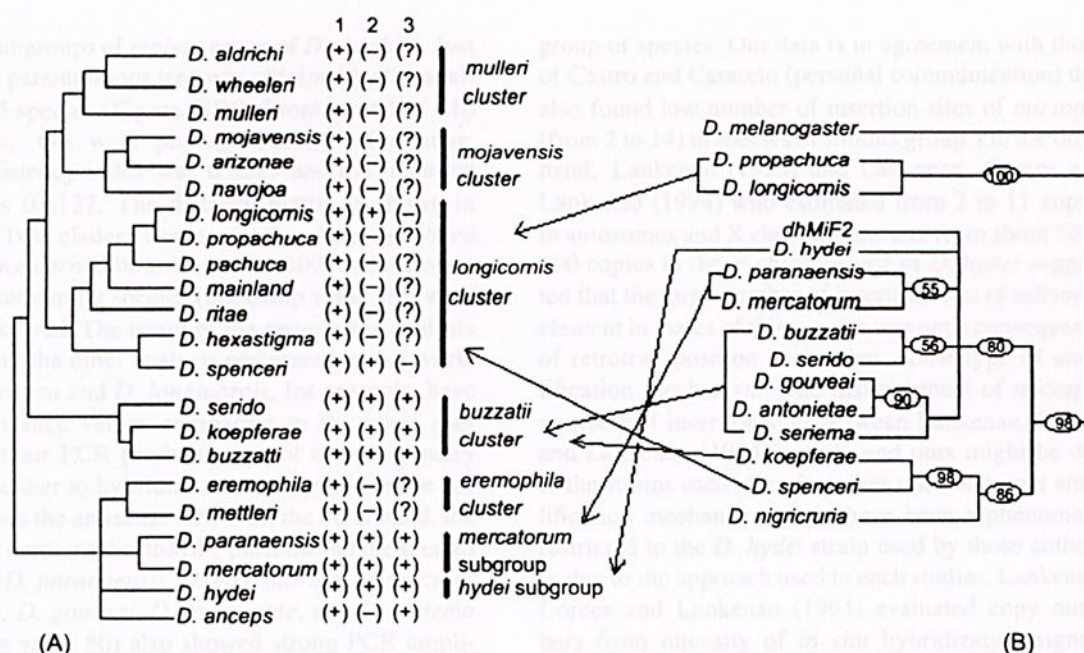


Figure 5. (A) Phylogeny of *repleta* group of *Drosophila* (according to Durando et al., 2000, but excluding the species not studied in this work) with the indication of presence (+), absence (-) or lack of information (?) of *micropia* sequences identified by PCR with (1) and without DMSO (2) and the expression of antisense RNA (3). (B) Phylogeny of *micropia* sequences. Comparisons were limited to 386 bp amplified by primers 2813 and 2814. The cladogram was generated by parsimony analysis, using the branch-and-bound algorithm (Swofford, 1997). A *micropia* sequence from *D. melanogaster* was used as outgroup. The consistency index was 0.8029 and retention index was 0.6127. The numbers in the ovals indicate the percent of 500 bootstrap replications that contain that clade (over 50%).

Table 2. Genetic distances among *micropia* nucleotide sequences calculated using Kimura's two-parameter method

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) <i>D. melanogaster</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(2) <i>D. propachuca</i>	0.34	-	-	-	-	-	-	-	-	-	-	-	-	-
(3) <i>dhMiF2</i>	0.31	0.33	-	-	-	-	-	-	-	-	-	-	-	-
(4) <i>D. hydei</i>	0.35	0.31	0.04	-	-	-	-	-	-	-	-	-	-	-
(5) <i>D. paranaensis</i>	0.33	0.30	0.02	0.01	-	-	-	-	-	-	-	-	-	-
(6) <i>D. mercatorum</i>	0.34	0.33	0.06	0.06	0.00	-	-	-	-	-	-	-	-	-
(7) <i>D. koepferae</i>	0.51	0.48	0.35	0.34	0.31	0.35	-	-	-	-	-	-	-	-
(8) <i>D. buzzatii</i>	0.34	0.30	0.02	0.02	0.01	0.02	0.35	-	-	-	-	-	-	-
(9) <i>D. serido</i>	0.35	0.36	0.06	0.10	0.05	0.08	0.33	0.04	-	-	-	-	-	-
(10) <i>D. gouveai</i>	0.36	0.34	0.05	0.05	0.04	0.05	0.44	0.05	0.08	-	-	-	-	-
(11) <i>D. antonietae</i>	0.43	0.42	0.18	0.18	0.12	0.17	0.39	0.13	0.15	0.12	-	-	-	-
(12) <i>D. seriema</i>	0.41	0.39	0.14	0.14	0.07	0.11	0.41	0.08	0.13	0.04	0.16	-	-	-
(13) <i>D. spenceri</i>	0.54	0.55	0.47	0.43	0.42	0.45	0.38	0.43	0.47	0.56	0.50	0.50	-	-
(14) <i>D. longicornis</i>	0.35	0.12	0.38	0.39	0.39	0.39	0.52	0.38	0.38	0.39	0.43	0.40	0.63	-
(15) <i>D. nigricrura</i>	0.58	0.56	0.47	0.46	0.45	0.45	0.56	0.45	0.48	0.48	0.50	0.49	0.58	0.62

RNA is depicted in Figure 5(A) in which we tried to establish a relationship between *micropia* distribution and characteristics and the phylogeny of its hosts.

Phylogenetic analysis

The aim of the phylogenetic analysis was to determine the relationships among *micropia* sequences in species

of three subgroups of *repleta* group of *Drosophila*. Just one more parsimonious tree was obtained by the analysis of 15 species (Figure 5(B)). From a total of 410 characters, 194 were phylogenetically informative. The consistency index was 0.8029 and the retention index was 0.6127. The distance matrix is shown in Table 2. Two clades, one containing *D. propachuca* and *D. longicornis* (bootstrap value 100) and the other with the remaining species (bootstrap value 98), were highly resolved. The result of the sequencing analysis agrees with the other analysis performed in this work. *D. propachuca* and *D. longicornis*, for example, have higher distance values comparing to the other species and their PCR products are not complementary enough neither to hybridize to the *micropia* probe nor to express the antisense RNA. On the other hand, the clade that contains the dhMIF₂ plasmid and the species *D. hydei*, *D. paranaensis*, *D. mercatorum*, *D. buzzatii*, *D. serido*, *D. gouveai*, *D. antonietae*, and *D. seriema* (bootstrap value 80) also showed strong PCR amplification, intense hybridizing bands and expression of the antisense RNA.

Discussion

Our results show a wide distribution of *micropia* among the 26 species, which suggests that this retrotransposon could be present in the common ancestor of *repleta* group of *Drosophila*. Lankenau (1993) and Lankenau, Corces & Lankenau (1994) reported the occurrence of *micropia* in 24 species belonging to three subgenus (*Sophophora*, *Drosophila* and *Dorsilopha*). These data suggest that *micropia* may be as old as *Drosophila*. Then, *micropia* could have been in the genus *Drosophila* before its diversification and might have been transmitted vertically to its descendent lineages. However, under high stringent PCR only 11 species showed *micropia* sequences and, under less stringent PCR all species seem to harbor sequences of this element. Then, we propose that the 15 species with weak amplification harbor more divergent sequences of *micropia*.

According to Kaplan, Darden and Langley (1985) it is possible that insertion sites of a TE correlate with the time that this TE has been present in its hosts. Accordingly, a species invaded by some TE long time ago probably presents few copies of this sequence in its genome. The low number of insertion sites observed in the *repleta* group (0–17) corroborates the hypothesis that *micropia* is an old component of this

group of species. Our data is in agreement with those of Castro and Carareto (personal communication) that also found low number of insertion sites of *micropia* (from 2 to 14) in species of *saltans* group. On the other hand, Lankenau (1993) and Lankenau, Corces and Lankenau (1994) who estimated from 2 to 11 copies in autosomes and X chromosome and from about 50 to 100 copies in the Y chromosome of *D. hydei* suggested that the large number of insertion sites of *micropia* element in males of this species was not a consequence of retrotransposition but, rather, some type of amplification mechanism. The disagreement of *micropia* numbers of insertion sites between Lankenau, Corces and Lankenau (1994)'s study and ours might be due to the strains used in each studies once *micropia* amplification mechanism might have been a phenomena restricted to the *D. hydei* strain used by those authors or due to the approach used in each studies. Lankenau, Corces and Lankenau (1994) evaluated copy numbers from intensity of *in situ* hybridization signals in chromosomes of larval brains and the frequency of *micropia* containing clones during microcloning experiments. It is also possible that we did not detect *in tandem* sequences since the restriction enzyme used in Southern analysis could have produced fragments of the same length if several copies of *micropia* were located side by side and restriction site polymorphism were not present. Even though, if the male specific bands were *in tandem*, they must have been stronger. Additionally, the ladder pattern characteristic of *in tandem* sequences was not observed in the PCR analysis.

Micropia retrotransposon encodes a 5 kb RNA detectable in adult stages both in somatic and in germline tissues of females as well as of males. In addition to the full-length and heterogeneous high-molecular-weight sense transcript, 1.0 and 1.6 kb long antisense RNAs are also expressed only in testis. The 1.0 kb transcript which is about 30 times more abundant than the full-sized one is initiated from an internal promoter located between the RNase H and the integrase regions (Lankenau, 1993). The 1.6 kb transcript probably arose by rearrangements due to an insertion in the 1.0 kb sequence. These antisense RNAs are produced from *micropia* elements located on the Y chromosomal lampbrush loops during the prophase of male meiosis, while the sense RNA is made at the same time on the autosomes and/or the X chromosomes (Lankenau, 1993).

Lankenau, Corces and Lankenau (1994) also suggested that the male-specific bands of *D. hydei* might

be meaningful with respect to testis specific transcripts since they were strongly detected in this species. Our data, based on the analyses of a higher number of species than that of Lankenau (1993), showed that this relationship seems to be not so straightforward. While *D. spenceri*, a species that possesses four male-specific sequences, seems not to express any antisense transcript, *D. hydei*, with just one male-specific element, encodes the antisense RNA. On the other hand, *D. buzzatii*, with non-detectable male-specific sequence encodes the antisense RNA.

Our results are in agreement with Lankenau (1993) concerning the testis specific expression of this transcript, but with an intriguing exception. *D. mercatorum* also expresses a lower-molecular-weight antisense RNA in somatic tissues. One possible explanation for this unexpected expression of the antisense RNA in heads is that a defective or rearranged *micropia* element of *D. mercatorum* is able to swindle the mechanism that controls the differential expression of antisense RNAs in the soma and testis tissues. There are several experimental evidences for coupled, balanced or unbalanced expression of sense and antisense RNAs that suggest that antisense transcripts may govern the expression of their sense counterparts (Vanheé-Brossollet & Vaquero, 1998). Studies about endogenous antisense RNAs deserves further exploration in order to promote understanding of how, when and where these antisense RNA transcripts regulate gene expression.

Transposable elements can be classified according to their structure into two classes, as proposed by Finnegan (1989). Additionally, several subdivisions within the main classes are established on the basis of their degree of identity and similarity in terms of their DNA or protein sequences. Terms such as megafamily, superfamily, family, and subfamily, are frequently used to define closely related groups of TEs. According to Capy et al. (1998), two elements that are 80% similar belong to the same subfamily, while those that are 40% similar belong to same family but to different subfamilies. Our data suggest that the *repleta* group of *Drosophila* harbors at least two subfamilies of *micropia* elements. The first proposed subfamily is a monophyletic group of sequences comprising sequences of the *hydei* (*D. hydei*) and *mercatorum* (*D. paranaensis* and *D. mercatorum*) subgroups and of four species of the *buzzatii* cluster of *mulleri* subgroup (*D. buzzatii*, *D. gouveai*, *D. antonietae* and *D. seriema*). These *micropia* sequences differ from 0 to 18% among them and more than 30% to the

others. In addition, these elements express the antisense RNA and are strongly detected by Southern analysis. This distribution of *micropia* sequences agrees with the distribution of their host proposed by Tatarenkov and Ayala (2001) who suggested a closer relationship between the *hydei* subgroup and the *buzzatii* cluster. The second subfamily includes *micropia* elements of *D. propachuca* and *D. longicornis*, which differ 12% from each other and more than 30% from the other sequences. Moreover, they do not express the antisense RNA and are not complementary to the *D. hydei micropia* sequence used as a probe. A third clade, which cannot be characterized as a subfamily, comprises sequences of *D. koepferae*, *D. spenceri* and *D. nigricruria* that are from 38 to 56% divergent and also differ in relation to their antisense RNA expression, which is produced only by *D. koepferae*, but in very low amount.

The analysis of the *micropia* element showed a considerable diversity among the various sequences and allowed us to propose the existence of, at least, two subfamilies of this element in the *repleta* species group. Additionally, with only an exception because *D. koepferae micropia* sequence is grouped with *D. spenceri*, there is strong evidence for vertical transmission to be the main process of *micropia* evolutionary dynamics. This suggestion is based on the fact that the *micropia* phylogeny is consistent with the phylogenetic relationships among its host species.

Acknowledgements

The authors thank D.-H. Lankenau for providing us with dhiMF₂plasmid and W. Etges, F.M. Sene, F.R. Torres, C.R. Ceron, and M. Manfrin for providing us with strains of *Drosophila*. This work was supported by FAPESP (Grants 95/7192-2 and 00/11313-0 to C.M.A.C and 97/14646-5 to L.M.A) and CNPq. We also thank Ms Solange Aranha for English review.

References

- Almeida, L.M., J.P. Castro & C.M.A. Carareto, 2001. *Micropia* transposable element presence in *Drosophila saltans* group. *Dros. Inf. Serv.* 84: 114–117.
- Capy, P., C. Bazin, D. Higuier & T. Langin, 1998. Classification of transposable elements, pp. 37–52 in *Dynamics and Evolution of Transposable Elements*, 1st edn. Landes Bioscience, France.
- Chakrabarti, R. & C.E. Schutt, 2001. The enhancement of PCR amplification by low molecular-weight sulfones. *Gene* 274: 293–298.

- Durando, C., R. Baker, W. Etges, B. Heed, M. Wasserman & R. De Salle, 2000. Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. *Mol. Phylogenet. Evol.* 2: 296–307.
- Finnegan, D., 1989. Eukaryotic transposable elements and genome evolution. *Trends Genet.* 5: 103–107.
- Huijser, P., C. Kirchhoff, D.-H. Lankenau & W. Hennig, 1988. Retrotransposon-like sequences are expressed in Y chromosomal lampbrush loops of *Drosophila hydei*. *J. Mol. Biol.* 203: 689–697.
- Jowett, T., 1986. Preparation of nucleic acids, pp. 275–286 in Roberts D.B. (ed), *Drosophila: A Practical Approach*, edited by D.B. Roberts. IRL Press, Oxford.
- Kaplan, N.L., T. Darden & C. Langley, 1985. Evolution and extinction of transposable elements in Mendelian populations. *Genetics* 109: 459–480.
- Lankenau, D.-H., 1993. The retrotransposon family *micropia* in *Drosophila* species, pp. 232–241 in J. McDonald (ed.), *Transposable Elements and Evolution*. Kluwer Publishers, Amsterdam.
- Lankenau, D.-H., P. Huijser & W. Hennig, 1989. Characterization of the long terminal repeats of *micropia* elements dissected from the Y chromosomal lampbrush loops 'Threads' of *Drosophila hydei*. *J. Mol. Biol.* 209: 493–497.
- Lankenau, S., G.V. Corces & D.-H. Lankenau, 1994. The *Drosophila micropia* retrotransposon encodes a testis-specific antisense RNA complementary to reverse transcriptase. *Mol. Cell Biol.* 14: 1764–1775.
- Silva, J.C. & M.G. Kidwell, 2000. Horizontal transfer and selection in the evolution of *P* elements. *Mol. Biol. Evol.* 17(10): 1542–1557.
- Swofford, D., 1997. PAUP: Phylogenetic analysis using parsimony. Version 4.0b10. Smithsonian Institution, Washington, D.C.
- Tatarenkov, A. & J.F. Ayala, 2001. Phylogenetic relationships among species groups of *virilis-repleta* Radiation of *Drosophila*. *Mol. Phylogenet. Evol.* 21: 327–331.
- Vanheé-Brossollet, C. & C. Vaquero, 1998. Do natural antisense transcripts make sense in eukaryotes? *Gene* 211: 1–9.
- Vilela, C.R., 1983. A revision of *Drosophila repleta* species group (Diptera, Drosophilidae) Rve. Bras. Ent. 1–114.
- Thompson, J.D., D.G. Higgins & T.J. Gibson, 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673–4680.

Capítulo 3

Distribuição descontínua e relações filogenéticas entre retrotransposons copia em espécies de *Drosophila* do grupo *repleta*.

**DISTRIBUIÇÃO DESCONTÍNUA E RELAÇÕES
FILOGENÉTICAS ENTRE RETROTRANSPOSONS COPIA
EM ESPÉCIES DE *DROSOPHILA* DO GRUPO *REPLETA*.**

Resumo: O retrotransposon copia foi amplamente estudado em espécies de *Drosophila* do grupo *melanogaster* e foi demonstrado ser um bom marcador para estudos de interações hospedeiro-parasita. Entretanto, pouco é conhecido a respeito da distribuição, do número de inserções e da história evolutiva de copia em espécies do grupo *repleta* de *Drosophila*. As análises realizadas mostraram a existência de apenas uma subfamília copia distribuída descontinuamente entre as 26 espécies do grupo *repleta*. Análises de Southern blot mostram um baixo número de inserções no genoma das espécies desse grupo. A baixa diversidade de nucleotídeos na UTR sugere a existência de forças seletivas que mantêm essa região regulatória ao longo do tempo evolutivo e que a família copia do grupo *repleta* é mais antiga do que a do grupo *melanogaster*. Análises de associações históricas sugerem a ocorrência de eventos de divergência, de perda estocástica e de um evento de transferência horizontal entre *D. buzzatii* e *D. serido* para explicar a história evolutiva de copia nas espécies do cluster *buzzatii*.

Palavras-chave: *Drosophila*, grupo *repleta*, retrotransposon copia, número de inserções, transferência horizontal.

Distribuição descontínua e relações filogenéticas entre retrotransposons *copia* em espécies de *Drosophila* do grupo *repleta*.

Resumo: O retrotransposon *copia* tem sido intensivamente estudado em espécies de *Drosophila* do grupo *melanogaster* e foi demonstrado ser um bom modelo para estudos das interações entre os mecanismos de regulação dos retrotransposons e dos genomas hospedeiros. Entretanto, pouco é conhecido a respeito da distribuição, do número de inserções, do polimorfismo de seqüências e do modo de transmissão desse retrotransposon em espécies do grupo *repleta* de *Drosophila*. As análises realizadas mostraram a existência de apenas uma subfamília *copia* distribuída descontinuamente entre as 26 espécies do grupo *repleta*. Análises de *Southern blot* mostram um baixo número de inserções no genoma das espécies desse grupo. A baixa diversidade de nucleotídeos na ULR sugere a existência de forças seletivas que mantêm essa região regulatória ao longo do tempo evolutivo e que a família *copia* do grupo *repleta* é mais antiga do que a do grupo *melanogaster*. Análises de associações históricas sugerem a ocorrência de eventos de divergência, de perda estocástica e de um evento de transferência horizontal entre *D. buzzatii* e *D. serido* para explicar a história evolutiva de *copia* nas espécies do cluster *buzzatii*.

Palavras-chave: *Drosophila*, grupo *repleta*, retrotransposon *copia*, número de inserções, transferência horizontal.

* Artigo a ser submetido para publicação no periódico *Heredity*.

Keywords: *Drosophila*, *repleta* group, *copia* retrotransposon, number of insertions, horizontal transfer



Discontinuous distribution and phylogenetic relationships between *copia* retrotransposable elements in *Drosophila* species of the *repleta* group.

Luciane M. de Almeida^a and Claudia M.A. Carareto^{a*}

UNESP - Universidade Estadual Paulista, Departamento de Biologia, 15054-000 São José do Rio Preto, SP, Brazil, e-mail: carareto@bio.ibilce.unesp.br, Phone: 55-17-2212382, Fax: +55-17-2212390. *Author for correspondence.

Abstract

The retrotransposon *copia* has been intensively studied in the *melanogaster* group of *Drosophila* and it has been demonstrated to be a good model system for the study of regulatory interactions between retrotransposon and its host genome. However, very little is known about *copia* distribution, number of genomic insertions, sequence polymorphism and mode of transmission in 26 species of the *repleta* group of *Drosophila*. Our analysis showed the existence of just one *copia* subfamily, a discontinuous distribution of *copia* elements among the *repleta* species and a low number of insertions in the genome of the *repleta* group. The low diversity of nucleotides in the ULR suggests that selective constraints maintain this regulatory region over evolutionary time. In addition, divergence analysis reinforces that the *copia* family harbored by species of the *buzzatii* cluster represents a more ancestral group than that harbored by the *melanogaster* species group. Events of divergence, extinction and of horizontal transfer between *D. buzzatii* and *D. serido* based on analysis of historical associations can be suggested to explain the evolution of *copia* in the *buzzatii* cluster.

Keywords: *Drosophila*, *repleta* group, *copia* retrotransposon, number of insertions, horizontal transfer.

Introduction

Many transposable elements in *Drosophila* are distributed discontinuously among related taxa. This discontinuity can reflect their acquisition by means of interspecific transfer (horizontal transfer) or by sexual transmission (vertical transfer) combined with events of ancestral polymorphism, extinction and different rates of selection (Capy *et al.*, 1994). The retrotransposon *copia* is found in many species of *Drosophila* species (Cavarec *et al.*, 1997), but there is very little information about its distribution in the *repleta* group. It has only been identified in only six species of *repleta* group, *D. koepferae*, *D. buzzatii*, *D. martensis*, *D. mojavensis*, *D. mulleri* and *D. mercatorum* (Biémont and Cizeron, 1999), but no information is available about its presence in the other 86 species of the group. *Copia* is 5.4 kb long flanked by 276 bp direct long terminal repeats (LTRs) and contains a single large open reading frame (ORF) capable of encoding a polyprotein of 1049 amino acids (Emori *et al.*, 1985; Mount and Rubin, 1985). The 5'LTR contains promoter sequences and the start site of transcription (Arkhipova *et al.*, 1995). The URL contains a number of sequence motifs, which bind host factors, and function as enhancer for the *copia* element (Seneddon and Flavell, 1989; Cavarec and Heidmann, 1993; Cavarec *et al.*, 1994; 1997; Matyunina *et al.*, 1996; Wilson *et al.*, 1998). Hence, the ULR is suggested to control *copia* transposition, which is a rate-limiting step in the retrotransposition process (Booeke *et al.*, 1985).

Like other TEs, *copia* retrotransposons are believed to maintain themselves despite the cost to their hosts, due to their capacity to replicate faster than their hosts through the transposition process (Hickey, 1982). According to Deininger (1992), mobile sequences that transpose efficiently are most likely to survive and propagate. This situation leads to competition between transposable elements within the genome that is known as interelement selection (McDonald *et al.*, 1997). The result of this competition, plus the interaction with host regulatory mechanism, can result in the maintenance of *copia* sequences (negative purifying selection) or their change (positive diversifying selection). Jordan and McDonald (1998 b) studied the *copia* LTR-URL natural variation among 26 species of the *melanogaster* and two species of the *repleta* group and observed that there is a positive diversifying selection acting between *copia* families (*melanogaster* and *repleta*) and a negative purifying selection acting within the families.

Aiming to broaden our knowledge of the evolutionary history of the *copia* retrotransposon we investigated their phylogenetic distribution, the number of insertions in the genome and the molecular natural variation in the LTR-URL region in 26 species of the *repleta* group of



Drosophila. The sequence analysis allowed us to propose the evolutionary relationships among the *copia* sequences, the occurrence of evolutionary changes during events of vertical transmission and also the occurrence of horizontal transfer events of *copia* elements between species.

Material and methods

Fly stocks

All species and strains used in this study are listed in Table 1; each list includes the taxonomic nomenclature, location, and stock number or collection date.

DNA amplification and sequencing

The regions of *copia* focused on in this study consist of the 5'LTR and the untranslated leader region (ULR). PCR reactions were performed in 25 μ l volumes using approximately 200ng of template DNA, 100 μ M of each primer, 200 μ M of dNTPs, 1.5 mM of $MgCl_2$, 1.25 μ l of DMSO and 1 unit of Taq DNA Polymerase (GIBCO-BRL) in 1x Polymerase buffer. After an initial step of denaturation for 5 min at 94°C, 30 cycles consisting of 45 sec at 94°C, 1 min at 52°C and 1 min at 72°C were carried out, followed by a final extension step of 5 min at 72°C. The primers CoBuz1: 5'-CCCNTATTCCTCCTTCAAAAA-3' and CoBuz2: 5'-CCGCGAAATTAAGAAACGAG-3' were designed from the *D. buzzatii copia* sequence (X96972) and annealed into the LTR-URL *copia* region. The amplified fragment has 615 pb and corresponds to nucleotides from 10 to 625. The amplified fragments were separated by electrophoresis in 1% agarose gel. The PCR products were cloned into a TA cloning vector (Invitrogen). Both strands of three clones, chosen at random, were sequenced for each species.

Southern blot analysis

Genomic DNA was prepared from 50 adult flies according to Jowett (1986). DNA samples were digested with *Bam* HI and *Xho* I restriction enzymes to determine the number of insertion sites. DNA fragments were separated by electrophoresis in 1% agarose gels and transferred to nylon membranes (Hybond N+/Amersham-Pharmacia). The membranes were hybridized to a random primer-labeled probe at 60°C in 5xSSC; 0.1% SDS; 5% dextran sulfate,

and 20-fold dilution of liquid block. The filters were washed twice with 0.2x SSC and 0.5% SDS for 15 min at 60°C. Hybridization was performed using Gene Images (Amersham-Pharmacia) according to the manufacturer's instructions. A 615 bp *copia* sequence amplified from PTZ18 plasmid containing *copia* element of *D. koepferae* (Genbank accession X96971) was used as a probe.

Evolutionary analysis

The multiple alignments of *copia* consensus sequences were performed with the program CLUSTAL W (Thompson *et al.*, 1994). The evolutionary relationships among *copia* sequences were constructed using the maximum parsimony method, using a branch and bound algorithm as implemented in PAUP v.4.0b10 (Swofford, 2000). Branch support was calculated by bootstrap analysis consisting of 500 replicates. The HKY model (Hasegawa *et al.*, 1985) of DNA substitution used in distance analysis as best fitting the data was determined by a likelihood ratio test using MODELTEST 2.0 (Posada and Crandall, 1998). The *copia* sequence of *D. melanogaster* (X02599) was used as outgroup.

Comparison of host and *copia* phylogenies

TreeMap, software that has been recently used in studies of host-parasite, can be an important tool in phylogenetic analysis of transposable element. TreeMap was developed to test for co-evolutionary events working on the assumption of maximizing the number of co-speciation events (Page, 1994). The program maps the TE phylogeny onto the host phylogeny given the distribution of TE among host species and then selects the outcome that maximizes the number of co-speciation events. The maximum number of co-speciation events found is then assessed for significance against a null distribution of co-speciation events generated by randomizing the host phylogeny. Since TreeMap requires fully resolved source phylogenetic tree, for this analysis we used only sequences of *D. koepferae*, *D. buzzatii* and *D. serido* species because there is no resolved host tree for the 26 species of the *buzzatii* cluster that we studied in this work.



Statistical Tests

The levels of nucleotide diversity and numbers of segregating sites were determined for the LTR and ULR regions of the *copia* transposable element using the DnaSP program (Rozas and Rozas, 1997) in order to evaluate whether the sequences evolve randomly (neutral theory) or have been subjected to functional constraints.

Results

Distribution of *copia* among species and number of insertion sites

DNA amplification of *copia* sequences by PCR showed a fragment of the expected size (615 pb) in 11 of the 26 species analyzed (Figure 1), but only six of them showed a strong amplification, all belonging to the *buzzatii* cluster (*D. koepferae*, *D. buzzatii*, *D. serido*, *D. gouveai*, *D. antonietae* and *D. seriema*). In five species (*D. hydei*, *D. paranaensis*, *D. mulleri*, *D. mettleri* and *D. nigrospiracula*) faint amplification was observed, which may be due to the occurrence of very divergent *copia* sequences. In addition to the 615 bp fragments, larger fragments that may be due to amplification of rearranged elements were also observed. No amplification occurred in species of the *mercatorum* subgroup (*D. mercatorum*) and the *mulleri* subgroup (*D. aldrichi*, *D. wheeleri*, *D. navojoa*, *D. arizonae*, *D. mojavenensis*, *D. pachuca*, *D. propachuca*, *D. hexastigma*, *D. eremophila*, *D. spenceri*, *D. longicornis*, *D. nigricruria*, *D. anceps* and *D. ritae*).

Southern blot analysis confirmed the results obtained by PCR and showed that the total number of insertion sites of sequences homologous to *copia* in entire genomes of 11 species ranged from one (*D. navojoa*) to nine (*D. koepferae*). No *copia* hybridizing sequences were detected in *D. hydei*, *D. wheeleri*, *D. mettleri* and *D. nigrospiracula*.

Evolutionary analysis

The aim of the sequencing analysis was to investigate the natural *copia* LTR-URL nucleotide variation and to propose a phylogenetic relationship between these sequences in species of the *buzzatii* cluster of the *repleta* group. Just one more parsimonious tree was obtained by the analysis of seven taxa (Figure 3). From a total of 654 characters, 34 were phylogenetically informative. The consistency index was 0.9492 and the retention index was 0.6977. The distance matrix is shown in Table 2.



Two clades can be observed in the tree, one containing *D. antonietae*, *D. koepferae*, *D. gouveai* and *D. seriema* (bootstrap value 99) and the other with the *D. buzzatii* and *D. serido* (bootstrap value 63). The smallest divergence was found between *D. seriema* and *D. gouveai* (0.73%) and the largest between *D. melanogaster* and *D. seriema* (48.66%). Distance analysis indicates the occurrence of two *copia* families, the *melanogaster* and *repleta* families (Table 2).

Comparison of host and transposable element phylogenies

TreeMap analysis was used to examine the occurrence of possible sorting, duplication, divergence and horizontal transfer events of *copia* element. This analysis produced two trees (Figure 4). Both trees obtained show the ancestral *copia* sequence passing through a divergence process that gives the *copia* sequences of *D. melanogaster* (*melanogaster* family) and of the *repleta* family. Both trees are equally parsimonious and show in the *repleta* family one occurrence of divergence (circle 2), three events of extinction (interrupted lines) and one of horizontal transfer (arrow) to explain the evolutionary changes of this TE in the *repleta* host species. In the first tree, the *copia* sequence was lost in all the *repleta* species, with exception of the *buzzatii* cluster. In the *buzzatii* cluster one more sorting event occurs and the ancestor of *D. buzzatii* loses the *copia* sequence, but later this element was reintroduced into this species by horizontal transfer from *D. serido*. The second tree, like the first, shows the *copia* sequence present only in species of the *buzzatii* cluster, but in this case the ancestor of *D. serido* loses the *copia* element, which was later reintroduced into this species by horizontal transfer from *D. buzzatii*.

Statistical test

The nucleotide diversity and the numbers of segregating sites were calculated in order to determine whether different selective constraints have been imposed under the *copia* LTR and ULR sequences. Our data showed that the polymorphism within the LTR was 0.06718, with 45 segregating sites, and within the ULR was 0.01979, with 18 segregating sites. Hence, the ULR conservation is 3.4 times bigger than the LTR, which suggests that some degree of selective constraint has been imposed on the ULR compared to the LTR.



Discussion

The retrotransposon *copia* has been intensively studied in the *melanogaster* group of species and it has been demonstrated to be a good model system for the study of regulatory interactions between retrotransposon and its host genome (Flavell *et al.*, 1981; Cavarec *et al.*, 1993; Matyunina *et al.*, 1996; Cavarec *et al.*, 1997; Jordan and MacDonald, 1998b). However, little is known about its distribution, number of genomic insertions, sequence polymorphism and mode of transmission in the *repleta* group of *Drosophila*, one of the most numerous neotropical groups of *Drosophila* species.

Our analyses showed a discontinuous distribution of *copia* elements among 26 species of *repleta* group. Discontinuities of TEs distribution can be explained by vertical transmission followed by events of ancestral polymorphism and extinction in particular species or through horizontal transmission events. The last event has been recently suggested in numerous reports (Maruyama and Hartl, 1991; Kidwell, 1993; Loreto *et al.*, 2001; Clark *et al.*, 2002; Robertson *et al.*, 2002; Martinez-Sebastian *et al.*, 2002), which show that this phenomenon is possibly more frequent than was previously believed. A previous study has already proposed the occurrence of horizontal transfer of *copia* retrotransposon between *D. melanogaster* and *D. willistoni* (Jordan *et al.*, 1999). It has been postulated that cross-species transfers may be an effective strategy by which TEs avoid inactivation over evolutionary time (Maruyama and Hartl, 1991). Since *copia* is known to be subject to effective host-mediated repression, selective pressure might favor its horizontal transfer over evolutionary time (Seneddon and Flavell, 1989; Matyunina *et al.*, 1996).

Transposable elements can be classified according to their structure into classes, subclasses, families, and subfamilies. Subfamilies are about 30-40% divergent in terms of their DNA. Jordan and MacDonald (1998 a,b) suggested the occurrence of two *copia* families in the *Drosophila* genus: the *melanogaster* family with three subfamilies and the *repleta* family, based on the analysis of only two species of the *repleta* group (*D. koepferae* and *D. buzzatii*). Looking for other families and subfamilies within the *repleta* group we extended the number of species analyzed to 26, but PCR amplification yielding dense DNA fragments only occurred in the six species of the *buzzatii* cluster. The highest distance value among the *copia* sequence was 7.84% (*D. buzzatii* and *D. seriema*), which reinforces that this group of species harbors just one subfamily of the *copia* family retrotransposon. However, it is important to observe that the faint

fragments amplified in two other species (*D. hydei*, *D. mettleri*) do not cross-hybridize with the *copia* probe and also were not cloned because of its low DNA concentration. However these sequences could represent very divergent *copia* subfamilies or even families.

A previous study realized by Jordan and McDonald (1998 b) in 26 species of the *melanogaster* and only two species of the *repleta* group showed that *copia* LTR-URL has been subject to a negative purifying selection within each family and positive adaptive selection between families. The authors, using only three sequences of the *repleta* family (two from *D. buzzatii* and one *D. koepferae*), showed that the nucleotide diversity of ULR was 2.2 times bigger than the LTR. Using eight species we showed that this ratio is even higher (3.4). This result reinforces the hypothesis of functional constraint of *copia* ULR regulatory regions in the *repleta* family. Comparing our data with that for the *melanogaster* family (Jordan and McDonald, 1998 b), the nucleotide diversity in the *repleta* LTR and ULR is 3.0 and 2.8 times higher than that in the *copia melanogaster* family, respectively. These ratios are lower than previously obtained by the authors (LTR = 4.2 and ULR = 6.1). The differences between the two studies are probably due to the smaller number of *repleta* species studied in that study.

In addition, we also determined the evolutionary relationships between *copia* sequences of the *buzzatii* cluster of species using parsimony method. The *buzzatii* species cluster is one of several complexes constituting the *mulleri* subgroup of the *repleta* group, subgenus *Drosophila*. At present, the *buzzatii* cluster consists of seven species (*D. buzzatii*, *D. koepferae*, *D. borborema*, *D. serido*, *D. gouveai*, *D. antonietae* and *D. seriema*); the last three were recently classified as distinct species based on aedeagus morphology (Tidon-Sklorz and Sene, 1995; Tidon-Sklorz and Sene, 2001). In the tree of the *copia* sequences two clades were formed, one containing *D. antonietae*, *D. koepferae*, *D. gouveai* and *D. seriema* and the other with the *D. buzzatii* and *D. serido*. Interestingly, *D. buzzatii* and *D. serido* are never grouped as sister taxa in the species phylogeny (Durando *et al.*, 2000; Ruiz *et al.*, 2000; Manfrin *et al.*, 2001).

Aiming to broaden our knowledge about the evolutionary history of the *copia* retrotransposon we confronted the host and the TE phylogenies aiming to identify the main events that have occurred during its evolution within the *repleta* group using TreeMap. Since, within the *buzzatii* cluster, the host tree is resolved only for *D. buzzatii*, *D. serido* and *D. koepferae* (Durando *et al.*, 2000), sequences of these three species alone were used in this analysis. The tree obtained suggests that *copia* was present in the ancestor of the *repleta* group



cluster, in which it occurs in all analyzed species. This suggestion is based on the fact that TreeMap assumes that the *copia* element of the outgroup (*D. melanogaster*) and the *copia* element of *repleta* group have a monophyletic origin, what is not necessarily the truth. In addition, the origin of retrotransposon *copia* in the *buzzatii* cluster of species could be explained by an event of horizontal transfer from a species not analyzed in this study, which does not belong to the *repleta* group. The low divergence of nucleotides between the sequences of the five species (1% to 7%) reinforces the hypothesis of recent origin. Independently whether the origin of *copia* has occurred in an ancestral lineage of the *repleta* group or in the ancestral lineage of the cluster *buzzatii*, within this cluster the moment at which the *copia* element was lost is the main difference between both trees. It is impossible to say whether the loss had occurred in the ancestor of *D. buzzatii*, this sequence being reintroduced into this species by horizontal transfer from *D. serido*, or whether it occurred in the ancestor of *D. serido*, and the reintroduction occurred from *D. buzzatii*. Both hypotheses are equally possible since *D. buzzatii* and *D. serido* are sympatric in the "Caatinga" in eastern Brazil, and along the Brazilian coast (Brito *et al.*, 2002). In addition to the proposition of horizontal transfer mediated by vectors, other possible explanation for the transfer of *copia* element between *D. serido* and *D. buzzatii* is the interspecies hybridization. We have previously found the same closer similarity of sequences between two other TEs of *D. serido* and *D. buzzatii* using parsimony, the retrotransposon *micropia* (Almeida and Carareto, 2004) and the transposon *Minos* (Almeida and Carareto, personal communication). Taking these results together, we can suggest that through one or more events of hybridization between both cryptic species the three TEs have jumped, together or separately, from *D. buzzatii* to *D. serido* or in the opposite direction. This hypothesis is corroborated with the observation that there are different levels of isolation mechanism between species of *buzzatii* cluster (Madi-Ravazzi *et al.*, 1997; Machado *et al.*, 1999).

Several phylogenetic studies have been carried out to elucidate the phylogeny of the *buzzatii* cluster (Ruiz *et al.*, 2000; Rodríguez-Trelles *et al.*, 2000; Durando *et al.*, 2000; Manfrin *et al.*, 2001; Kuhn, 2003) but the trees obtained show several polytomies and are generally discordant. Recent studies have shown that retrotransposons can be appropriate tools for elucidating difficult host phylogenies (Graur and Li, 2000). Although TreeMap is excellent software for the analysis of historical associations, it requires fully resolved phylogenetic trees. Because of this, our analysis was restricted to the *D. melanogaster*, *D. koepferae*, *D. buzzatii* and



Because of this, our analysis was restricted to the *D. melanogaster*, *D. koepferae*, *D. buzzatii* and *D. serido* species. When resolved host trees become available it will be possible to expand our analysis including all the species that compose the cluster. Despite being an active element that is prone to jump between species, it is also vertically transferred and reinforces the monophyletic state of the *buzzatii* cluster. This result indicates that *copia* could be a useful molecular marker.

One of the important questions about retrotransposons is whether their evolution has been influenced primarily by vertical or horizontal transmission. The answer to this question is complex due to a lack of enough sampling of the host taxa and because of poorly resolved phylogenies. Recent reports (Terzian *et al.*, 2000; Vazquez-Marinque *et al.*, 2000; Jordan *et al.*, 1999; Flavell *et al.*, 1997) and this study have shown that horizontal transfer has not been limited to class II elements. In addition, this study reinforces that the *copia* family harbored by species of the *buzzatii* cluster represents a more ancestral group than that harbored by the *melanogaster* species group.

Acknowledgments

The authors wish to thank Oriol Cabre for providing us with plasmid PTZ18 and W. Etges, F.M. Sene, L. Madi-Ravazzi, F.R. Torres, C.R. Ceron, H. E. M. C. Bicudo, and M. Manfrin for providing us with strains of *Drosophila*, and F. P. L. Marques for phylogenetic analysis suggestion. This work was supported by FAPESP (Grants nos. 95/7192-2 and 00/ 1371-0 to C.M.A.C and 97/14646-5 to L.M.A) and CNPq. Revision of the English language text was carried out by Peter James Harris.

References

- Almeida LM, Carareto CM (2004). Identification of two subfamilies of *micropia* transposable elements in species of the *repleta* group of *Drosophila*. *Genetica* 121: 155-164.
- Arkhipova IR, Lyubomirskaya NV, Hyin YV (1995). *Drosophila* retrotransposons. R. G. Landes, Austin, TX.
- Biémont C, Cizeron G (1999). Distribution of transposable elements in *Drosophila* species. *Genetica* 105: 43-62.

Flavell AJ, Pearce SR, Heslop-Harrison P, Kumar A (1997). The evolution of *Ty1-copia* retrotransposon in eukaryote genomes. *Genome* 100: 183-195.

Booeke JD, Garfinkel DJ, Styles CA, Fiuk GR (1985). *Ty* elements transpose through an RNA intermediate. *Cell* **40**:491-500.

Brito RA, Manfrin MH, Sene FM (2002). Nested cladistic analysis of Brazilian population of *Drosophila serido*. *Mol. Phylogenet. Evol.* **22**: 131-143.

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

Cavarec L, Heidmam T (1993). The *Drosophila copia* retrotransposon contains binding sites for transcriptional regulation by homeoproteins. *Nucleic Acids Res.* **21**:5041-5049.

Cavarec L, Jensen S, Heidmam T (1994). Identification of a strong transcriptional activator for the *copia* retrotransposon responsible for its differential expression in *Drosophila hydei* and *melanogaster* cell line. *Biochem. Biophys. Res. Commun.* **203**: 392-399.

Cavarec L, Jensen S, Casella JF, Cristescu SA, Heidmam T (1997). Molecular cloning and characterization of a transcription factor for the *copia* retrotransposon with homology to BTB-containing lola neurogenic factor. *Mol. Cell. Biol.* **17**: 482-494.

Clark JB, Silva J, Kidwell MG (2002). Evidence of horizontal transfer of *P* transposable elements. In: Horizontal gene transfer. M Syvanen, and C.I. Kado, eds. Academic Press, San Diego, p. 161-171.

Csinki AK, McDonald JF (1995). Analysis of *copia* sequence variation within and between *Drosophila* species. *Mol. Biol. Evol.* **12**: 83-93.

Deininger PL, Batzer MA, Hutchison CA, Edgel MH (1992). Master genes in mammalian repetitive DNA amplification. *Trends Genet.* **8**: 307-311.

Durando CM, Baker RH, Etges WJ, Heed W, Wasserman M, DeSalle R (2000). Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. *Mol. Phylogenet. Evol.* **16**: 296-307.

Emori Y, Shiba T, Kanaya S, Inouye S, Yuki S, Saigo K (1985). Determination of the nucleotide sequences of *copia* and *copia*-related RNA in *Drosophila* VLP. *Nature* **315**: 773-776.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by molecular clock of mitochondrial DNA. *J. Mol. Evol.* **21**: 160-174.

Hickey DA (1982). Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* **101**: 519-531.

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

Flavell AJ, Pearce SR, Heslop-Harrison P, Kumar A (1997). The evolution of *Ty1-copia* retrotransposon in eukaryote genomes. *Genetica* **100**: 185-195.



Graur D, Li WH (2000) Fundamentals of Molecular Evolution, second edition, Sinauer Associates, U.S.A.

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

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

Jordan IK, Matyunina VL, McDonald JF (1999). Evidence for 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 DB (ed) *Drosophila: a Practical Approach*, edited by D.B. Roberts, IRL Press, Oxford.

Kaplan NL, Darden T, Langley C. (1985) Evolution and extinction of transposable elements in Mendelian populations. *Genetics* **109**: 459-480.

Kidwell MG (1993). Lateral transfer in natural populations of eukaryotes. *Annu. Rev. Genet.* **27**: 235-256.

Kidwell MG, Lisch DR (2000). Transposable elements and host genome evolution. *Trends Ecol. Evol.* **15**:95-99.

Kuhn GCS, Franco FF, Silva Jr. W, Martinez-Rossi NM, Sene FM (2003). On the pBuM189 satellite DNA variability among South American populations of *D. buzzatii*. *Hereditas* **139**: 161-166.

Loreto ELS, Valente VLS, Zaha A, Silva JC, Kidwell MG (2001). *Drosophila mediopunctata* P elements: A new example of horizontal transfer. *The Journal of Heredity* 375-381.

Machado LPB, Madi-Ravazzi L, Tadei WJ (1999). Reproductive and chromosome studies in *Drosophila* species of *buzzatii* cluster. *Genet. Mol. Biol.* **22**: 386.

Madi-Ravazzi L, Bicudo HEM, Manzato JA (1997). Reproductive compatibility and chromosome pairing in the *Drosophila buzzatii* complex. *Cytobios* **89**: 21-30.

Manfrin MH, De Brito ROA, Sene FM (2001). Systematics and evolution of the *Drosophila buzzatii* (Diptera: Drosophilidae) cluster using mtDNA. *Ann. Entomol. Soc. Am.* **94**: 333-346.

Martinez-Sebastian M, Hernández M, Mejias B, Gas M, Pérez A, Pascual L, De Frutos R (2002). Evolutionary patterns of *gypsy* and *bilbo* retrotransposon families in the *Drosophila* species of the *obscura* group. *Mol. Phylogenet. Evol.* **22**: 254-266.

- Maruyama K, Hartl DL (1991). Evidence for interspecific transfer of transposable element *mariner* between *Drosophila* and *Zaprionus*. *J. Mol. Evol.* **33**: 514-524.
- Matyumina LV, Jordan IK, McDonald JF (1996). Naturally occurring variation in *copia* expression is due to both element (cis) and host (trans) regulatory variation. *Proc. Natl. Acad. Sci. USA* **93**: 855-864.
- McDonald JF, Matyumina LV, Wilson SW, Jordan IK, Bowen NJ, Miller WJ (1997). LTR retrotransposon and the evolution of eukaryotic enhancers. *Genetica* **100**: 3-13.
- Mount SM, Rubin GM (1985) Complete nucleotide sequence of *Drosophila* transposable element *copia*: Homology between *copia* and retroviral proteins. *Mol. Cell. Biol.* **5**: 1630-1638.
- Page RDM (1994). Parallel phylogenetic: reconstructing the history of host-parasite assemblages. *Cladistics* **10**: 155-173.
- Pinsker W, Haring E, Hangemann S, Miller WJ (2001) The evolutionary life history of *P* transposons: from horizontal invaders to domesticated neogenes. *Chromosoma* **110**: 148-158.
- Posada D, Cradall KA (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817-818.
- Robertson HM, Soto-Adames FN, Walden KK, Avancini RM, Lampe DJ (2002). The *mariner* transposon of animals: horizontally jumping genes. In: Horizontal gene transfer. M. Syvanen and C.I. Kado eds. Academic Press, San Diego, CA.
- Rodriguez-Trelles F, Alarcón L, Fontdevila A (2000). Molecular evolution and phylogeny of the *buzzatii* complex (*Drosophila repleta* group): a maximum-likelihood approach. *Mol. Biol. Evol.* **17**: 1112-1122.
- Rozas J, Rozas R (1997). DnaSP version 2.0: a novel software package for extensive molecular population genetics analysis. *Comput. Appl. Biosci.* **13**: 307-311.
- Ruiz A, Cansian AM, Kuhn GCS, Alves MA, Sene FM (2000). The *Drosophila serido* speciation puzzle: putting new pieces together. *Genetica* **108**: 217-227.
- Seneddon A, Flavell AJ (1989). The transcriptional control of the *copia* retrotransposon. *Nucleic Acids Res.* **17**: 4025-4035.
- Swofford D (2000). PAUP: Phylogenetic analysis using parsimony (and other methods). Version 4.0.b10. Sinauer, Sunderland, Massachusetts.
- Terzian C, Ferraz C, Demaille J, Bucheton A (2000). Evolution of *gypsy* endogenous retrovirus in *Drosophila melanogaster* subgroup. *Mol. Biol. Evol.* **17**: 908-914.



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

Tidon-Sklorz R, Sene FM (2001). Two new species of the *Drosophila serido* sibling set (Diptera, Drosophilidae). *Iheringia Ser. Zool.* **90**: 141-146.

Tidon-Sklorz R, Sene FM (1995). *Drosophila seriema* nsp: new member of the *Drosophila* of the *Drosophila serido* (Diptera: Drosophilidae) superspecies taxon. *Annals of Entomological Society of America* **88**: 139-142.

Vasquez-Marinque RP, Hernandez M, Martinez-Sebastian MJ, De Frutos R (2000). Evolution of gypsy endogenous retrovirus in *Drosophila obscura* species group. *Mol. Biol. Evol.* **17**: 1185-1193.

Wilson SW, Matyunina LV, McDonald JF (1998). An enhancer region within the copia untranslated leader contains binding sites for *Drosophila* regulatory proteins. *Gene* **209**: 239-246.

	<i>D. longiareolata</i>	Tehuacan, Mexico	19811374.1	1990
	<i>D. serido</i>	Milagres, BA, Brazil	19811431.3	
<i>hazarensis</i>	<i>D. hazarensis</i>	Novo Horizonte, SP, Brazil		1995
	<i>D. gouveai</i>	Morro do Campo, MG, Brazil		
	<i>D. antonietae</i>	Sela Vista, MS, Brazil		1990
	<i>D. seriema</i>	Serra do Cipó, MG, Brazil		1990
	<i>D. longiareolata</i>	Tucson, Arizona, USA	19811371.5	
	<i>D. pachica</i>	Chapingo, Mexico	19811391.0	
<i>longiareolata</i>	<i>D. jessupachica</i>	Hidalgo, Mexico	19811411.1	
	<i>D. hexastigma</i>	Zapotitan, Puebla, Mexico	19811382.2	
	<i>D. spenceri</i>	Guaymas, Sonora, Mexico	19811441.2	
	<i>D. areolaris</i>	Zapotitan, Puebla, Mexico	19811291.1	
<i>areolaris</i>	<i>D. costalis</i>	Tucson, Arizona, USA	CAT392	
	<i>D. nigritaria</i>	Novo Horizonte, SP, Brazil		1995
	<i>D. nigropunctata</i>	Tucson, Arizona, USA	19811503.1	1978
	<i>D. microps</i>	Infernillo, Michoacán, Mexico	19815031.1	
	<i>D. rita</i>	Zapotitan dos Salinas, Puebla, Mexico		2002
<i>hyale</i>	<i>D. hyale</i>	Encarnación, Paraguay		2000

Table 1. Taxonomic list of *Drosophila* species of *repleta* group used in this study (according to Durando, 2000).

Subgroup	Cluster	Species	Location	Stock number	date
mercatorum		<i>D. paranaensis</i>	Novo Horizonte, SP, Brazil		1998
		<i>D. mercatorum</i>	Novo Horizonte , SP, Brazil		1998
mulleri	mulleri	<i>D. aldrichi</i>	Tamaulipas, Mexico	15081125.0	
		<i>D. mulleri</i>	Tucson, Arizona, USA	150811371.7	
		<i>D. wheeleri</i>	Catalina Island, Los Angeles, USA	150811501.3	
	mojavensis	<i>D. navojoa</i>	Tehuantepec -Mexico	150811374.1	
		<i>D. arizonae</i>	Tucson, Arizona, USA	A 1015	
		<i>D. mojavensis</i>	Grand Canyon, Arizona, USA	A870	
	buzzatii	<i>D. koepferae</i>	Tapia-Tucumam, Argentina		1990
		<i>D. serido</i>	Milagres, BA, Brasil	150811431.3	
		<i>D. buzzatii</i>	Novo Horizonte, SP, Brazil		1998
		<i>D. gouveai</i>	Morro do Chapéu, MG, Brazil		
		<i>D. antonietae</i>	Bela Vista, MS, Brazil		1990
mulleri	longicornis	<i>D. seriema</i>	Serra do Cipó, MG, Brazil		1990
		<i>D. longicornis</i>	Tucson, Arizona, USA	150811311.8	
		<i>D. pachuca</i>	Chapingo, Mexico	150811391.0	
		<i>D. propachuca</i>	Hidalgo, Mexico	150811411.1	
		<i>D. hexastigma</i>	Zapotitlan, Puebla, Mexico	150811302.2	
	eremophila	<i>D. spenceri</i>	Gyaymas, Sonora, Mexico	150811441.2	
		<i>D. eremophila</i>	Zapotitlan, Puebla, Mexico	150811292.1	
		<i>D. mettleri</i>	Tucson, Arizona, USA	CAT397	
		<i>D. nigricruria</i>	Novo Horizonte, SP, Brazil		1998
		<i>D. nigroespiracula</i>	Tucson, Arizona, USA	15081150311	1978
		<i>D. anceps</i>	Infernillo, Michoacán, Mexico	150815031.1	
		<i>D. ritae</i>	Zapotitlan dos Salinas, Puebla, Mexico		2002
hydei		<i>D. hydei</i>	Encarnación, Paraguay		2000

Table 2. Genetic distances among *copia* nucleotide sequences calculated using HKY method (Hasegawa *et al.*, 1985).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) <i>D. melanogaster</i>	-	-	-	-	-	-	-
(2) <i>D. koepferae</i>	0.4542	-	-	-	-	-	-
(3) <i>D. gouveai</i>	0.4569	0.0147	-	-	-	-	-
(4) <i>D. seriema</i>	0.4866	0.0220	0.0073	-	-	-	-
(5) <i>D. antonietae</i>	0.4487	0.0114	0.0181	0.0218	-	-	-
(6) <i>D. serido</i>	0.4465	0.0596	0.0577	0.0645	0.0488	-	-
(7) <i>D. buzzatii</i>	0.4415	0.0722	0.0702	0.0784	0.0649	0.0523	-

Figure 1. Ethidium bromide-stained agarose gel of PCR products obtained with *copia*-specific primers (CoBuz1 and CoBuz2) used as template genomic DNA from the following species (A): 1. *D. hydei*; 2. *D. paranaensis*; 3. *D. mercatorum*; 4. *D. koepferae*; 5. *D. buzzatii*; 6. *D. serido*; 7. *D. gouveai*; 8. *D. antonietae*; 9. *D. seriema*; 10. *D. aldrichi*; 11. *D. mulleri*; 12. *D. wheeleri*; 13. *D. novojoa*; 14. *D. arizonae*; 16. *D. mujavensis*; 17. *D. pochuca*; 18. *D. propachuca*; 19. *D. hexastigma*; 20. *D. spenceri*; 21. *D. longicornis*; 22. *D. eremophila*; 23. *D. metleri*; 24. *D. nighracula*; 25. *D. nigroesphracula*; 26. *D. anceps*; 27. *D. ritae* and 15 and 29. positive control. The arrows point to the expected 615 bp fragments.

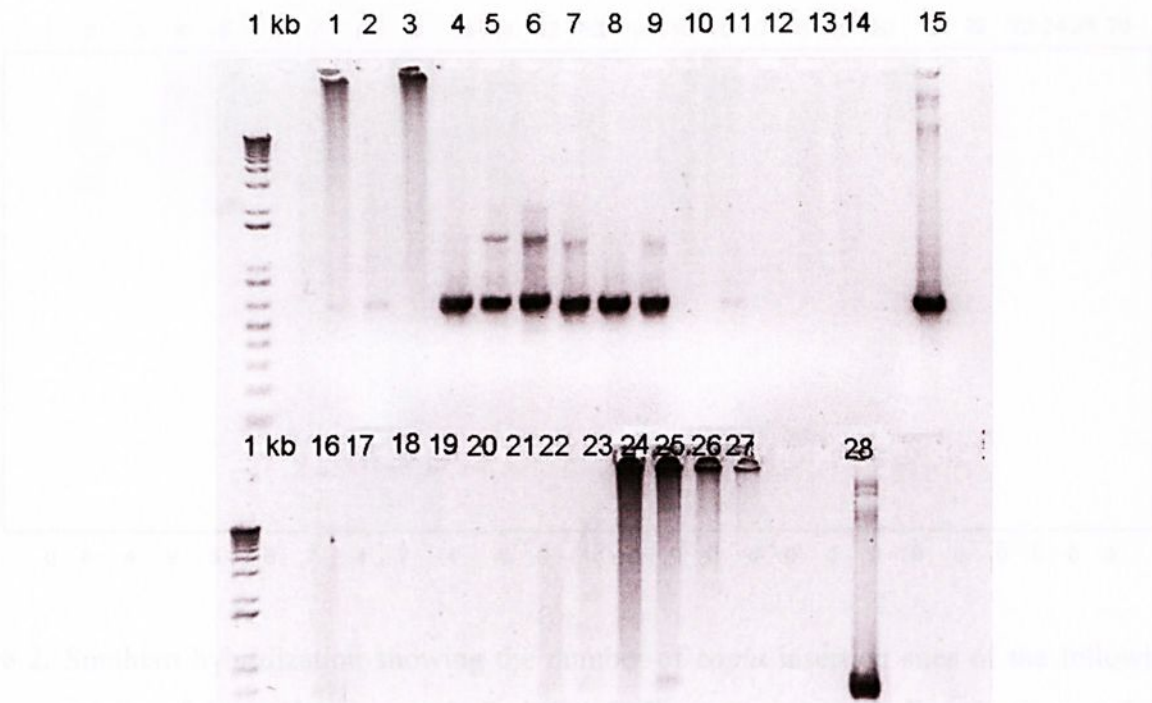


Figure 1. Ethidium bromide-stained agarose gel of PCR products obtained with *copia*-specific primers (CoBuz1 and CoBuz2) used as template genomic DNA from the following species (A): 1. *D. hydei*; 2. *D. paranaensis*; 3. *D. mercatorum*; 4. *D. koepferae*; 5. *D. buzzatii*; 6. *D. serido*; 7. *D. gouveai*; 8. *D. antonietae*; 9. *D. seriema*; 10. *D. aldrichi*; 11. *D. mulleri*; 12. *D. wheeleri*; 13. *D. navojoa*; 14. *D. arizonae*; 16. *D. mojaviensis*; 17. *D. pachuca*; 18. *D. propachuca*; 19. *D. hexastigma*; 20. *D. spenceri*; 21. *D. longicornis*; 22. *D. eremophila*; 23. *D. mettlari*; 24. *D. nigricurria*; 25. *D. nigrospiracula*; 26. *D. anceps*, 27. *D. ritae* and 15 and 29. positive control. The arrows point to the expected 615 bp fragments.

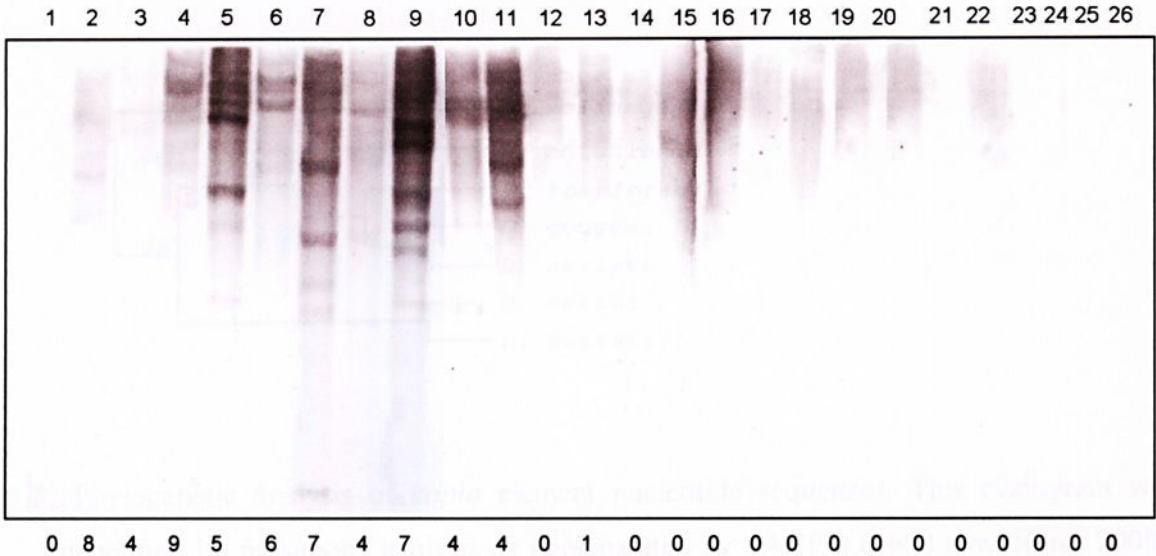


Figure 2. Southern hybridization showing the number of *copia* insertion sites of the following species of the *repleta* group. 1. *D. hydei*; 2. *D. paranaensis*; 3. *D. mercatorum*; 4. *D. koepferae*; 5. *D. buzzatii*; 6. *D. serido*; 7. *D. gouveai*; 8. *D. antonietae*; 9. *D. seriema*; 10. *D. aldrichi*; 11. *D. mulleri*; 12. *D. wheeleri*; 13. *D. navojoa*; 14. *D. arizonae*; 15. *D. mojavenensis*; 16. *D. pachuca*; 17. *D. propachuca*; 18. *D. hexastigma*; 19. *D. spenceri*; 20. *D. longicornis*; 21. *D. eremophila*; 22. *D. mettléri*; 23. *D. nigricruria*; 24. *D. nigrospiracula*; 25. *D. anceps*, 26. *D. ritae*. The number of insertions in the genome is shown for each species. The DNA of 50 individuals were digested with *Bam* HI and *Xho* I and probed with 615 bp fragment amplified from plasmid PTZ18.

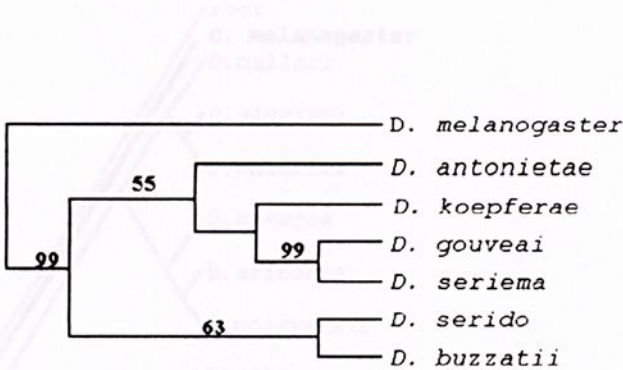


Figure 3. Phylogenetic analysis of *copia* element nucleotide sequences. This cladogram was generated by parsimony analysis as implemented by PAUP 4.0 b10 (Swofford, 2000). The consistency index is 0.9492 and the retention index 0.6977. Bootstrap analysis of 500 replicates was performed on the data.

Figure 4. Historical associations implemented by TreeMap (Page, 1994). Stacked TreeMap reconciliation trees of *Drosophila* phylogeny (blue lines) and *copia* retrotransposon phylogeny (dark lines). Divergence event (circle); horizontal transfer (red arrows); extinction (red interrupted lines).

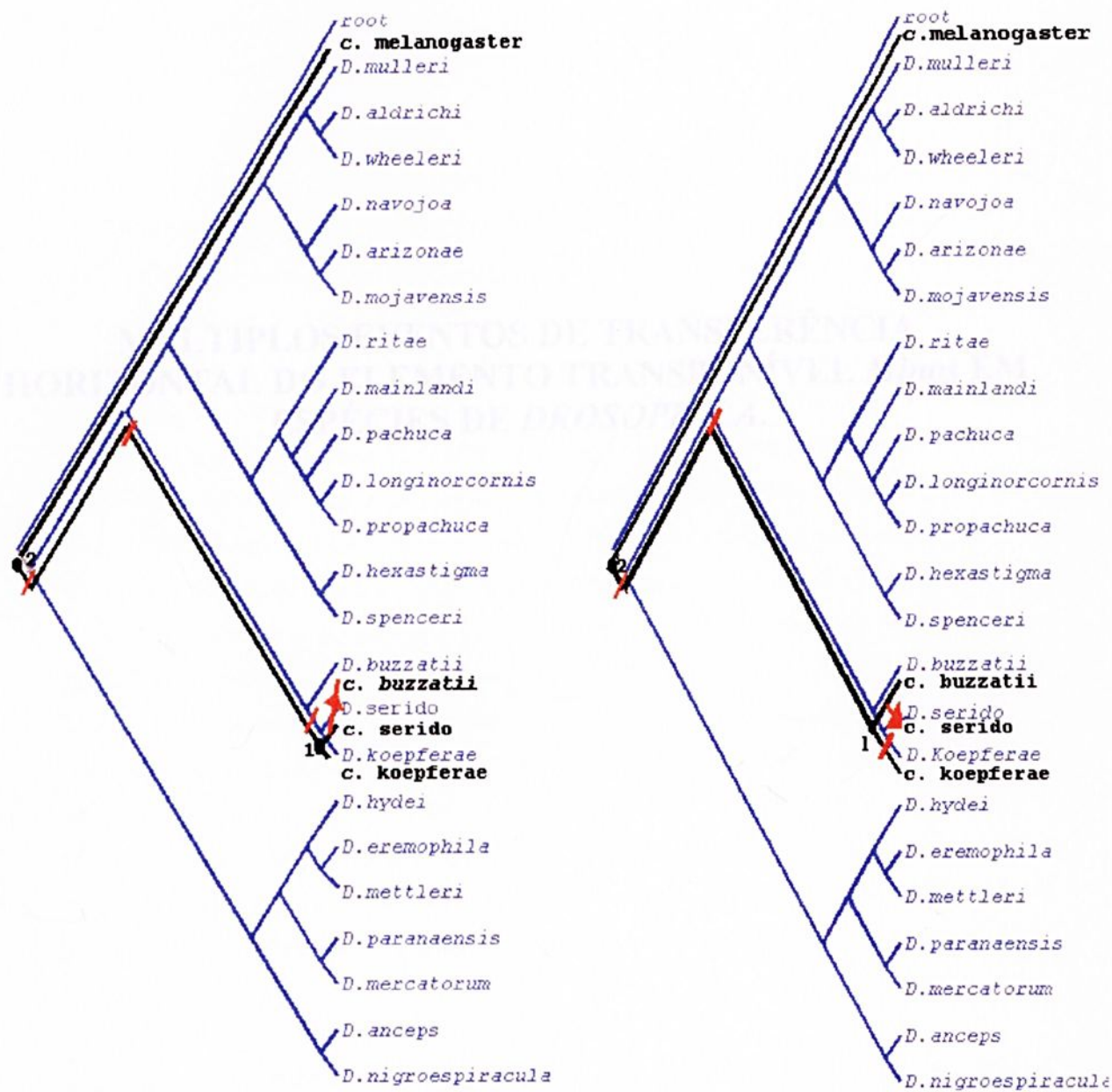


Figure 4. Historical associations implemented by TreeMap (Page, 1994). Stacked TreeMap reconciliation trees of *Drosophila* phylogeny (blue lines) and *copia* retrotransposon phylogeny (dark lines). Divergence event (circle); horizontal transfer (red arrows); extinction (red interrupted lines).

Múltiplos eventos de transferência horizontal do elemento transponível *Minos* em espécies de *Drosophila*.

Resumo: Este estudo compreende a análise do transposon *Minos* em 16 espécies do grupo *repleta* e em sete espécies do grupo *salans* do gênero *Drosophila*. Análises de PCR e *Southern blot* mostraram uma ampla ocorrência do elemento *Minos* e um baixo número de

MÚLTIPLOS EVENTOS DE TRANSFERÊNCIA HORIZONTAL DO ELEMENTO TRANSPONÍVEL *Minos* EM ESPÉCIES DE *DROSOPHILA*.

substituições em 16 espécies do grupo *repleta* e em sete espécies do grupo *salans* sugeriu vários eventos de transferência horizontal, além daquele previamente sugerido por Asa et al. (2006) entre *D. mojavensis* e *D. salans*. Os eventos de transferência horizontal detectados neste estudo foram entre *D. salans* e uma linhagem ancestral das espécies dos clusters *mulleri* e *mojavensis*; entre *D. hydei* e uma linhagem ancestral das espécies dos clusters *mulleri* e *mojavensis*; entre *D. mulleri* e *D. mojavensis*; entre *D. viridis* e *D. buzzatii* e, ainda entre *D. emarginata* e *D. spenceri*. Uma hipótese alternativa que resumiria os três primeiros eventos acima descritos seria a ocorrência de repetidos eventos de transferência horizontal envolvendo a espécie *D. hydei*. Essa proposta está de acordo com os tempos de divergência das espécies hospedeiras (14, 6 e de 4 a 8 milhões de anos, respectivamente para as espécies *D. hydei*, clusters *mulleri* e *mojavensis* e *D. salans*). O presente trabalho sugere que a distribuição do elemento *Minos* entre as espécies dos grupos *repleta* e *salans* é determinada por múltiplos eventos de transferência horizontal balanceado por eventos de inativação vertical e de extinção.

Palavras-chave: elemento *Minos*, número de inserções, transferência horizontal, *Drosophila*, grupo *repleta*, grupo *salans*.



Múltiplos eventos de transferência horizontal do elemento transponível *Minos* em espécies de *Drosophila*.

Luciano M. de Almeida¹ and Claudia M.A. Carreira²

¹Universidade Estadual Paulista (UNESP), Departamento de Biologia, 13034-006 São João do

Resumo: Este estudo compreende a análise do transposon *Minos* em 26 espécies do grupo *repleta* e em sete espécies do grupo *saltans* do gênero *Drosophila*. Análises de PCR e *Southern blot* mostraram uma ampla ocorrência do elemento *Minos* e um baixo número de inserções entre as espécies dos grupos *repleta* e *saltans*. Análises de divergência de nucleotídeos de TEs, de associações históricas entre TEs e hospedeiros e das taxas de substituições em sítios sinônimos dos TEs e do gene *Adh* do hospedeiro permitiram sugerir vários eventos de transferência horizontal, além daquele previamente sugerido por Arcà *et al.* (2000) entre *D. mojavensis* e *D. saltans*. Os eventos de transferência horizontal detectados neste estudo foram entre *D. saltans* e uma linhagem ancestral das espécies dos *clusters mulleri* e *mojavensis*; entre *D. hydei* e uma linhagem ancestral das espécies dos *clusters mulleri* e *mojavensis*; entre *D. mulleri* e *D. mojavensis*; entre *D. serido* e *D. buzzatii* e, ainda entre *D. emarginata* e *D. spenceri*. Uma hipótese alternativa que resumiria os três primeiros eventos acima descritos seria a ocorrência de repetidos eventos de transferência horizontal envolvendo a espécie *D. hydei*. Essa proposta está de acordo com os tempos de divergência das espécies hospedeiras (14, 6 e de 4 a 8 milhões de anos, respectivamente para as espécies *D. hydei*, *clusters mulleri* e *mojavensis* e *D. saltans*). O presente trabalho sugere que a distribuição do elemento *Minos* entre as espécies dos grupos *repleta* e *saltans* é determinada por múltiplos eventos de transferência horizontal balanceado por eventos de inativação vertical e de extinção.

Palavras-chave: elemento *Minos*, número de inserções, transferência horizontal, *Drosophila*, grupo *repleta*, grupo *saltans*.

Keywords: *Minos* element, *Minos* insertion number, horizontal transfer, *Drosophila* *repleta*

*Artigo a ser submetido para publicação no periódico *Molecular Phylogenetics and Evolution*



Multiple events of horizontal transfer of the *Minos* transposable element between *Drosophila* species

Luciane M. de Almeida¹ and Claudia M.A. Carareto*

Universidade Estadual Paulista (UNESP), Departamento de Biologia, 15054-000 São José do Rio Preto, SP, Brazil, e-mail: carareto@bio.ibilce.unesp.br, Phone:55-17-2212382, Fax:+55-17-2212390. *Author for correspondence.

Abstract

In this study the *Minos* element was analyzed in 26 species of the *repleta* group and seven species of the *saltans* group of the genus *Drosophila*. The PCR and Southern blot analysis showed a wide occurrence of the *Minos* transposable element among species of the *repleta* and the *saltans* groups and also a low number of insertions in both genomes. Three different analysis, nucleotide divergence, historical association and comparison between substitution rates (d_N and d_S) of *Minos* sequence and *Adh* host gene, suggest the occurrence of horizontal transfer between *repleta* and *saltans* species group. These data reinforces and extends the Arcà and Savakis (2000) results and suggest five events of horizontal transfer to explain the present *Minos* distribution: between *D. saltans* and the ancestor of the *mulleri* and the *mojavensis* clusters; between *D. hydei* and the ancestor of the *mulleri* and the *mojavensis* clusters; between *D. mulleri* and *D. mojavensis*; between *D. buzzatii* and *D. serido* and between *D. spenceri* and *D. emarginata*. An alternative explanation would be that repeated events of horizontal transfer involving *D. hydei*, which is a cosmopolitan species that diverged from the others *repleta* species as long as 14 Mya, could have spread *Minos* within the *repleta* group and to *D. saltans*. The data presented in this article support a model in which distribution of *Minos* transposon among *Drosophila* species is determined by horizontal transmission balanced by vertical inactivation and extinction.

Keywords: *Minos* element, *Minos* insertion number, horizontal transfer, *Drosophila repleta* group, *Drosophila saltans* group.

1. Introduction

Minos is a mobile element that belongs to the widespread *Tc1/mariner* superfamily. It was first identified as a dispersed repetitive sequence inserted into the rDNA locus of *Drosophila hydei* (Franz and Savakis, 1991). *Minos* is relatively small (1.6 to 1.8 kb) and is characterized by 255 bp perfect inverted terminal repeats (nucleotides 1 to 255 and 1521 to 1775) and the presence of two long non-overlapping open reading frames (ORFs) on the same strand that are separated by a putative 60-nucleotide intron (Franz *et al.*, 1994). *Minos* has been shown to be active in several dipteran (Franz and Savakis, 1991; Kapetanaki *et al.*, 2002; O' Brochta *et al.*, 2003) and lepidopteran (Shimizu *et al.*, 2000) species. In addition to insects, *Minos* is also active in invertebrates (Sasakura *et al.*, 2003), mouse tissues *in vivo* (Zagoraiou *et al.*, 2001; Drabek *et al.*, 2003) and human cells in culture (Klinakis *et al.*, 2000), where it has been used for transposon-mediated mutagenesis.

Despite its biological importance as a tool for genetic manipulation, relatively little is known about *Minos* evolution in the *Drosophila* genus. Several types of events can occur between transposable elements and their hosts during their evolutionary relationships, which can be denominated as historical associations. For example, transposable elements may diverge with their host (co-divergence), diverge independently of their hosts, become extinct or move from one species to another (horizontal transfer). Knowledge about *Minos* distribution among species, its number of insertions within genomes and structural variation among sequences is necessary to understand the evolutionary history of this TE.

There are two manners in which transposable elements can be transferred between species. The first is strictly mating dependent (vertical transmission), and the second occurs between species that are reproductively isolated (horizontal transmission). Vertical transmission predicts that all descendants of an ancestral species that harbors that element should also harbor homologous variants of the element; also the sequence conservation of an element should be congruent with the phylogeny of the host species. On the other hand, the occurrence of horizontal transmission predicts that the distribution of the element may be discontinuous and that the sequence conservation of the element is not necessarily related to the phylogeny of the species in which it resides.

A single previous study by Arcà and Savakis (2000) has shown the distribution of the *Minos* transposable element in 19 species of the subgenus *Sophophora* (including two species of the *saltans* group), eight of the subgenus *Drosophila* (including six species of the *repleta* group) and one species of the *Scaptodrosophila* subgenus. The analysis showed a

discontinuous distribution of the *Minos* transposon in the *Sophophora* subgenus. Distance sequence analysis (Neighbor-joining) realized with only four species (*D. mojavenensis*, *D. hydei*, *D. saltans* and *D. willistoni*) showed that the *Minos* sequence of *D. mojavenensis* is more similar to that of *D. saltans* one than to that of its sibling species *D. hydei*. This finding led Arcá and Savakis (2000) to suggest the occurrence of horizontal transfer of *Minos* between *D. saltans* and *D. mojavenensis*, two distantly related species. Aiming to broaden our knowledge about the evolutionary history of the *Minos* transposable element we extended the number of species to 33 and improved the phylogenetic analysis using the maximum parsimony method and the analysis of historical associations implemented by TreeMap, a software that makes it possible to identify co-divergence, horizontal transfer, duplication and sorting events.

2. Material and Methods

2.1 Fly stocks

All species and strains used in the present study are listed in Table 1; each list includes the taxonomic nomenclature, location, and stock number or collection date.

2.2 DNA amplification and Sequencing

PCR reactions were performed in 25 μ l volumes using approximately 200ng of template DNA, 100 μ M of each primer, 200 μ M of DNTs, 1.5 mM of $MgCl_2$, 1.25 μ l of DMSO and 1 unit of Taq DNA Polymerase (GIBCO-BRL) in 1x Polymerase buffer. After an initial step of denaturation for 10 min at 94°C, 29 cycles consisting of 1 min denaturation at 94°C, 1 min annealing at 54°C and 1 min extension at 72°C were carried out. An additional extension step of 10 min at 72°C was performed after the last cycle. The primers used (M5 5'-TATCGATAATGTCACAATACAGCATG-3' and M3: 5'-ATCAAGCTTGAATTGTGTAACGTCGCC -3') anneal, respectively, at the beginning and at the end of the coding region of the *D. hydei Minos* element (nt 1 to 1081). The fragments amplified are 1081bp long and were separated by electrophoresis in 1% agarose gel stained with ethidium bromide. The PCR products were cloned into a TA cloning vector (Invitrogen) and sequenced. For each amplified fragment two individual clones were chosen at random for sequencing.

2.3 Southern blot analysis

Genomic DNA was prepared from 50 adult flies according to Jowett (1986). DNA samples were digested with *Bam* HI and *Xho* I restriction enzymes to determine the number of insertion sites within the genome. DNA fragments were separated by electrophoresis in 1% agarose gels and transferred to nylon membranes (Hybond N+/Amersham-Pharmacia). The membranes were hybridized to a random primer-labeled probe at 60°C in 5xSSC; 0.1% SDS; 5% dextran sulfate, and 20-fold dilution of liquid block. The filters were washed twice with 0.2x SSC and 0.5% SDS for 15 min at 60°C. Hybridization was performed using Gene Images (Amersham-Pharmacia) according to the manufacturer's instructions. A 1.1 kb *Minos* sequence amplified from pBKSC plasmid containing the *Minos* sequence of *D. mojavensis* (provided by Bruno Árcia) was used as a probe.

2.4 Evolutionary analysis

Alignment of *Minos* sequences were done by program Clustal W (Thompson *et al.*, 1994) to estimate the evolutionary distances. The HKY model (Hasegawa *et al.*, 1985) of DNA substitution used in distance analysis as best fitting the data was determined by a distance ratio test using MODELTEST 2.0 (Posada and Crandall, 1998). DNA *Minos* sequences of *D. mojavensis* (AJ289698), *D. saltans* (AJ289699) and *D. willistoni* (AJ289700) were obtained from the EMBL DNA sequence library.

To the *Minos* element phylogeny, the sequences were aligned again with Clustal W (Thompson *et al.*, 1994) and then split into fragments, to speed up the analysis and to prevent clearly homologous regions from being treated as non-homologous. For the phylogenetic analysis, the gaps were removed from the sub-fragments to obtain unaligned fragments of sequences. The unaligned sequences were analyzed by using the "direct optimization" or "optimization alignment" method developed by Wheeler (1996) and implemented in the computer program POY (Gladstein and Wheeler, 1997). The direct optimization method is a maximum-parsimony algorithm that can process unaligned molecular sequences. The cladogram length is the sum of the costs for all hypothesized nucleotide substitution and insertion/deletion (indels) events. A sensitivity analysis *sensu* (Wheeler, 1995) was performed on the data. The chosen parameter sets are nine possible combinations of the gap: transversion: transition cost ratios (1:1:2; 2:1:1; 2:1:2; 2:2:1; 4:1:1; 4:1:2; 4:2:1; 8:2:1; 8:2:1). The sensitivity plots (nine dark and white fields related to each parameter) show under which of the parameter sets a given group of taxa comes out as monophyletic. In this way, a sensitivity analysis can discern between robust clades (those that under most or all of the

parameter sets) and unstable clades (those that appear only under one or few parameter regimes).

2.5 Historical associations

TreeMap, software that has recently been used in studies of the host-parasite relationship, can be an important tool in phylogenetic analysis of transposable elements. TreeMap was developed to test for co-evolutionary events (co-speciation, horizontal transfer, duplication and sorting events) working on the assumption of maximizing the number of co-speciation events (Page, 1994). The program maps the TE phylogeny onto the host phylogeny given the distribution of TE among host species, and then selects the outcome that maximizes the number of co-speciation (co-divergence) events. The maximum number of co-speciation events found is then assessed for significance against a null distribution of co-speciation events generated by randomizing the host phylogeny. Since TreeMap requires a fully resolved source phylogenetic tree, we excluded from the analysis the *Minos* of sequences *D. antonietae*, *D. gouveai* and *D. seriema* because there is no resolved phylogeny for the cluster *buzzatii* when these species are included.

2.6 The d_N/d_S ratios for host genes and *Minos* elements

In order to confirm the hypothesis of historical association and to investigate whether the low distances of *Minos* between species were due to horizontal transfers or to selection acting on the sequences, we estimated the number of synonymous substitutions per synonymous sites (d_S), the non-synonymous substitution per non-synonymous sites (d_N), the codon bias (CBI; Morton, 1993) and the effective number of codons (N_c ; Wright, 1990) of *Minos* and of a host gene sequence using the DnaSP program (Rozas and Rozas, 1997). The host gene chosen was *Adh* (alcohol dehydrogenase) because given its functional significance selective constraints are expected to be high. Vertical transmission can be assumed if the divergence of *Minos* between species is higher than that observed for *Adh*. On the other hand, if *Minos* divergence is significantly lower than that of *Adh* under similar or higher levels of selective constraints than those operating on the TE horizontal transfer can be inferred. The sequences of *Adh* were obtained from GENBANK: for the *saltans* group only exon 2 of *Adh* was used in the pairwise comparisons (*D. saltans*: AF045115; *D. austrosaltans*: AF045123; *D. emarginata*: AF045126; *D. subsaltans*: AF045117 and *D. sturtevantii*: 045114) and for the *repleta* group only *Adh* 2 was used (*D. hydei*: X58694; *D. buzzatii*: M62743; *D. arizonae*:



M622741; *D. mojavensis*: X12536; *D. mulleri*: X03048; *D. wheeleri*: M62851). For comparisons between groups of species only the exon 2 of both *Adh* was used.

3. Results

3.1 Distribution of Minos among species and number of insertion sites

DNA amplification by PCR showed a fragment of the expected size (1.1 kb) in at least 20 of the 26 species analyzed in the *repleta* group (Figure 1A) and in six of the seven species of the *saltans* group (Figure 1B). In addition to the 1.1 kb fragment, smaller and larger fragments were also observed. These bands probably result from the amplification of deleted or rearranged elements. *D. eremophila*, *D. nigricurria* and *D. nigrospiracula* showed faint lower-molecular-weight fragments and no amplification occurred in *D. hexastigma*, *D. metlleri* and *D. neocordata* genomes.

Sequences homologous to *Minos* element were clearly detected in 21 species of the *repleta* group (Figure 2) by Southern blot. This analysis showed that the total number of genomic insertion sites of *Minos* in entire genomes differs greatly, ranging from one (*D. spenceri*) to 17 (*D. aldrichi*). No *Minos* hybridizing sequences were detected in *D. eremophila*, *D. nigricurria* and *D. nigrospiracula*, which indicates that the faint bands detected by PCR were probably unspecific amplifications or very divergent sequences with a low degree of homology against the probe.

3.2 Evolutionary analysis

A nucleotide differentiation and genetic difference matrix based on Hasegawa *et al.* (1985) method was calculated for 840 nt of *Minos* sequences (Table 2). The overall distance ranged from 2% (between strains of *D. mojavensis*) to 42% (between *D. arizonae* and *D. spenceri*). Within the *saltans* group, the distances ranged from 4% (between strains of *D. saltans* and *D. sturtevantii*) to 13% (between *D. austrosaltans* and *D. emarginata*). The distances within the *repleta* group were 9% (between *D. mercatorum* and *D. paranaensis*) in the *mercatorum* subgroup and from 2% (between strains of *D. mojavensis*) to 42% (between *D. arizonae* and *D. spenceri*) in the *mulleri* subgroup.

The most parsimonious cladogram, resulting from 1:1:1 parameter costs (gap: transversion: transition), is shown in Figure 3. The information from the analysis of all data

under nine parameter sets (gap: transversion: transition cost ratios 1:1:2; 2:1:1; 2:1:2; 2:2:1; 4:1:1; 4:1:2; 4:2:1; 8:2:1; 8:2:1) is summarized in the sensitivity plots (dark and white squares) in the same figure. The tree obtained contains four major clades that show a high degree of resolution. The first clade contains *Minos* sequences of *D. saltans*, *D. seriema*, *D. buzzatii*, *D. serido*, *D. gouveai*, *D. antonietae* and *D. koepferae*. The branch that contains *D. gouveai*, *D. antonietae*, and *D. koepferae* is highly stable and repeats 90% of the analysis. The consistence is reduced to 70% when *D. serido*, *D. buzzatii* and *D. seriema* are included into the clade. Finally, *D. saltans* is grouped into this clade in only 20% of the analysis. The second clade is the largest and groups *Minos* sequences of *D. paranaensis*, *D. mercatorum*, *D. aldrichi*, *D. wheeleri*, *D. navojoa*, *D. arizonae*, *D. mojavenensis*, *D. mulleri*, *D. hydei*, *D. sturtevantii* and *D. subsaltans* obtained in this study and sequences from GenBank (*D. saltans*: AJ289699 and *D. mojavenensis*: AJ289698). The sensitivity analysis shows high consistence of the branch that contains *D. hydei*, *D. aldrichi*, *D. wheeleri*, *D. navojoa*, *D. arizonae*, *D. mojavenensis* and *D. mulleri* (90%). The branches with *D. sturtevantii* and *D. subsaltans* and with *D. paranaensis* and *D. mercatorum* show low consistencies (40%). The third clade includes *Minos* sequences from *D. pachuca*, *D. propachuca*, *D. emarginata* and *D. spenceri*. The sensitivity analysis shows high consistence of the branch containing *D. spenceri* and *D. emarginata* (80%) and of the branch including *D. pachuca* and *D. propachuca* (50%), but the consistence is low when these branches are analyzed together. The last clade, which includes sequences from *D. austrosaltans* and *D. anceps*, shows low branch consistence (20%).

3.3 Historical associations

Data from TE phylogenetic analyses showed that distantly related species harbor *Minos* sequences more closely related to one another than those within the same group of species, which might suggest horizontal transfer. The following pairs illustrate this situation: 1. *D. spenceri* (*repleta* group) and *D. emarginata* (*saltans* group); 2. *D. saltans* (*saltans* group) and the *mulleri* and the *mojavenensis* clusters (*repleta* group); 3. *D. anceps* (*repleta* group) and *D. austrosaltans* (*saltans* group); 4. *D. saltans* (*saltans* group) and *buzzatii* cluster (*repleta* group); and 5. *D. sturtevantii* and *D. subsaltans* and the *mulleri*, the *hydei* and the *mercatorum* subgroups. However, only the first and the second clades present high branch consistence. In order to confirm the hypothesis of horizontal transfer, historical association analysis was performed using TreeMap. Although, the program has constructed several hundreds of trees, all of them showed the same five types of horizontal transfer events;

however, the greatest number of transfer events observed in each tree was three. Figures 4 and 5 show trees that were chosen as the most representative. In 32 trees compared, the most frequent event of horizontal transfer occurred between (a) *D. saltans* and the ancestor of *mulleri* and *mojavensis* clusters (99 %), followed by the events between (b) *D. hydei* and the ancestor of *mulleri* and *mojavensis* clusters (44 %), (c) *D. mulleri* and *D. mojavensis* (31 %), (d) *D. buzzatii* and *D. serido* (44%) and between (e) *D. spenceri* and *D. emarginata* (19 %). These results suggest horizontal transfer to explain some of the inconsistencies observed in Figure 3, as the branch clustering *D. saltans* and species of the clusters *mulleri* and *mojavensis* and the branching of *D. spenceri* and *D. emarginata*, and suggest events previously not observed like the transfer of *Minos* between *D. mojavensis* and *D. mulleri* and between *D. buzzatii* and *D. serido*. Figure 6 shows one tree produced by TreeMap that resume the three first events of horizontal transfer described above (a,b,c) and suggests that *D. hydei* could be the species which dispersed *Minos* between the *saltans* and the *repleta* species groups.

3.4 *Minos* element versus host gene divergence

There are two possibilities for *Minos* sequences to be strongly conserved among different taxa: they are under some kind of strong selection, or they have been transferred relatively recently from one taxon to the other. The strength of selective constraints on *Minos* element sequence was investigated by comparing the estimated divergence among species with that obtained for the *Adh*. Sequences of five species of the *saltans* group and six species of the *repleta* group were used to estimate the numbers of substitution per synonymous (d_S) and non-synonymous (d_N) sites for *Minos* and for *Adh* (Table 3). d_N was smaller than d_S in all pairwise comparisons excepting when comparing *Minos* sequences between the *saltans* and the *repleta* groups. The $d_N : d_S$ ratios of *Minos* were higher when compared with that of *Adh*, which may be due to either weak selection at nonsynonymous sites or very strong selection at synonymous sites. Since differences in the magnitude of d_S among genes have been found to be related to the intensity of natural selection on synonymous codon usage (Shields *et al.*, 1988; Moriyama and Gojobori, 1992), we determined whether the magnitude of d_S between the *saltans* and the *repleta* species groups for *Adh* was correlated with the degree of codon bias (CBI) and the effective number of codons (N_c). N_c varies between 21 (for maximum bias) and 61 (for minimum bias) and a CBI varies between 0 (no bias) and 1 (maximum bias). Table 3 shows low to moderate levels codon bias in the sequences of *Minos* element (N_c of 53.4 and CBI of 0.46, in average), which suggest weak selection acting on these sequences

and higher selection-induced codon bias acting on *Adh* (N_c of 43.8 and CBI of 0.51, in average).

Comparison of d_N and d_S values between *Minos* elements and *Adh* estimated between species of the same group reveals that the average d_N for *Minos* is about 2 times higher than that observed for *Adh* but d_S is similar to (within the *saltans* group) or 3.7 smaller (within the *repleta* group) than that of *Adh*. However, between the species groups, d_N for *Minos* and *Adh* are similar one to the other but d_S is 15 times smaller than that observed for the host gene (Table 3, Figure 7). The smaller values of d_S between two species group might indicate that the *Minos* sequences shared a common ancestor more recently than did the host species.

4. Discussion

Analyses of many transposable elements have shown that their distribution among species does not follow the phylogeny of the host species, and species phylogenies based on their TEs are not always the same as those based on classical characters. Such discrepancies are frequently interpreted as resulting from horizontal transfer of TEs across boundary species (Capy *et al.*, 1994). Moreover, this incongruence can also reflect vertical transfer combined with events of ancestral polymorphism, extinction in particular lineage and different rates of selection. The data presented in this article supports a model in which distribution of *Minos* transposon among *Drosophila* species is determined by horizontal transmission balanced by vertical inactivation and extinction.

PCR and Southern blot analysis showed that the *Minos* transposable element is widely spread in the *repleta* and the *saltans* groups of *Drosophila*: 21 of the 26 species analyzed in the *repleta* group and six of the seven species analyzed in the *saltans* group. Five species of the *repleta* group (*D. hexastigma*, *D. eremophila*, *D. mettleri*, *D. nigricruria*, *D. nigroespiracula*) did not show any evidence of the *Minos* sequence, neither by PCR nor by Southern blot, which could be perfectly explained by stochastic losses or sequence divergence during the evolution of these species. DNA of *D. neocordata* (*saltans* group) did not show any amplification, but, according to Castro and Carareto (personal communication), it shows positive hybridization with the *Minos* probe.

The number of transposable elements within a genome is determined by three factors: the probability that a transposable element produces a new genomic copy; the probability that the element is excised; and the intensity of selection against increased numbers of

transposable elements within the genome (Graur and Li, 2000). Since an increase in copies is harmful to the host species it is believed that host species develop mechanisms to regulate TE copy numbers over time. Our results showed a low number of *Minos* insertions within the genome of the *repleta* group ($\bar{X} \pm S = 9.18 \pm 4.84$) as well as within the *saltans* group ($\bar{X} \pm S = 7.0 \pm 2.57$, Castro and Carareto, personal communication). The wide distribution and the low number of insertions of *Minos* in the genome of two groups of species, one belonging to the *Drosophila* subgenus and the other to the *Sophophora* subgenus, would suggest that this sequence might have been present in the common ancestor of the *Drosophila* and *Sophophora* subgenera. However, it is known that these subgenera may have diverged millions of years ago during the Eocene period (Throckmorton, 1975; Grimaldi, 1987, 1988). Hence, if only vertical transmission of *Minos* has occurred among the host species it would be reasonable to expect highly divergent sequences among these species. On the other hand, if horizontal transfer has also taken place, the *Minos* sequences would be remarkably conserved between certain species. The smallest nucleotide divergence obtained between *Minos* sequences in this study was between two strains of *D. mojavensis* (2%), whilst the largest one was between *D. spenceri* and *D. arizonae* (42%), two species belonging to the *repleta* group. The highest rates of divergence (above 18%) occurred in comparisons including only *D. spenceri* and *D. anceps* species. On the other hand, very low rates were found within the *saltans* group, within the *mulleri* and the *mojavensis* clusters and within the *buzzatii* cluster ($9.07\% \pm 2.2$; $6.06\% \pm 2.3$; $4.07\% \pm 1.1$, respectively). These results suggest that two *Minos* subfamilies harbor the genomes of the *repleta* and the *saltans* species analyzed in this study, one ancient and restricted to *D. spenceri* and *D. anceps*, and the other probably recently introduced and widely spread (*hydei*, *mercatorum* subgroups; *buzzatii*, *mulleri*, *mojavensis* clusters and *saltans* group) as shown by its rate of divergence.

In addition to the divergence rates, phylogenetic analysis also shows discrepancies between the host species and TE phylogenies. Besides the event of horizontal transfer suggested by Arcà and Savakis (2000) between *D. saltans* and *D. mojavensis*, additional events of horizontal transfer are suggested by our data to explain the *Minos* phylogeny within the *saltans* and *repleta* groups. The branching suggests five events of horizontal transfer: between (1) *D. saltans* and an ancestor of the *mulleri* and the *mojavensis* clusters; (2) *D. hydei* and an ancestor of *mulleri* and the *mojavensis* clusters; (3) *D. buzzatii* and *D. serido*, (4) *D. mulleri* and *D. mojavensis* and between (5) *D. spenceri* and *D. emarginata*. The low rates of

divergence of *Minos* between these species and groups reinforce the hypothesis of horizontal transfer, with the exception of the rate obtained between *D. spenceri* and *D. emarginata*.

Our data extend the Arcà and Savakis (2000) results and show that horizontal transfer did not necessarily occur between *D. mojaveensis* and *D. saltans*, but may have occurred between an ancestor of the species included in the *mulleri* and the *mojaveensis* clusters and *D. saltans*, as shown by their low rates of divergence and historical associations. On the other hand, Figure 6 shows an alternative hypothesis suggesting *D. hydei* as the species that spread *Minos* within *repleta* group and *D. saltans*. The last seems to us the most parsimonious hypothesis based on times of divergence of the host species. According to Russo *et al.* (1995), the lineage leading to *D. hydei* diverged from the other *repleta* species at about 14 Mya and the species belonging to the *mulleri* and the *mojaveensis* clusters diverged about 6 Mya. Accordingly, the divergence time for *D. saltans* was estimated at about 4 Mya (Clark and Kidwell, 1997). Hence, *D. hydei* could have transferred *Minos* to an ancestor of the clusters *mojaveensis* and *mulleri* about 6 Mya or even earlier, as well as to *D. saltans* and *D. aldrichi* at any time between 4 Mya and the present. In addition, it is important to observe that the transfers of *Minos* between *D. mojaveensis* and *D. mulleri* and between *D. buzzatii* and *D. serido* might be due to introgression since different levels of intercrosses have been observed between these sibling species (Madi-Ravazzi *et al.*, 1997; Machado *et al.*, 1999; Baffi and Ceron, 2002).

The estimates of d_S and d_N between the *Minos* sequence and host gene (*Adh*) confirm the hypothesis of horizontal transfer suggested by nucleotide divergence and historical associations. The comparison showing that d_S of *Minos* is smaller than that of *Adh* (3.7 times smaller within the *repleta* group and 15 times smaller between the *repleta* and the *saltans* species groups) strongly suggests that the invasion and subsequent spread of the *Minos* element between these species and species groups occurred recently.

The *Drosophila repleta* species group is among the largest of all species groups in the genus *Drosophila*. This group is intensively used in studies of speciation because it displays a range of different degrees of reproductive isolation among its members (Patterson & Stone, 1952). Current knowledge of the phylogeny of this group is based on morphological (Throckmorton, 1982a; Vilela, 1983), cytological (Wasserman, 1982; 1992) and molecular (Russo *et al.*, 1995; Rodríguez-Trelles *et al.*, 2000; Durando *et al.*, 2000; Tartarenkov and Ayala, 2001) data. However, the phylogenetic relationships within the *repleta* species group show several inconsistencies. Transposable elements can be used as molecular markers and might help to elucidate complex phylogenies, like that of the *repleta* group. Taking the

phylogeny of *Minos* and that of the *copia* element (Almeida and Carareto, personal communication), as a whole, we suggest that both transposable elements constitute a monophyletic group of sequences within the cluster *buzzatii* in agreement with the host monophyly (Ruiz *et al.*, 2000; Durando *et al.*, 2000; Manfrin *et al.*, 2001). Despite the relatively good resolution of the TE phylogenies and the agreement with the above cited host phylogenies, further studies, including more TEs and taxa, are necessary in order to be confident that TEs can be good molecular markers of the host species phylogenies in spite of the occasional events of horizontal transfers.

The life cycle of the *Minos* element is similar to that described for members of other class II TE family, such as *P* and *mariner* elements (Engels, 1989; Hartl *et al.*, 1997; Silva and Kidwell, 2000; Pinsker *et al.*, 2001). This cycle starts with the invasion of a naive host, a fast increase in insertion number and a stabilization of the mechanism of transposition repression that reduces the transposition rate. Thus, frequencies of active elements decrease due to mutation, drift and selection and the defective copies tend to extinction. Horizontal transfers have been postulated as a possible strategy developed by TEs to escape from host-mediated silencing mechanisms over evolutionary time (Jordan *et al.*, 1999). Within the past decade, there have been several reports of horizontal transfer in the genus *Drosophila* (Maruyama and Hartl, 1991; Kidwell, 1993; Loreto *et al.*, 2001; Clark *et al.*, 2002; Robertson *et al.*, 2002). Our data expands this list and suggests that this phenomenon is really much more frequent than was believed.

Acknowledgments

The authors thank Bruno Arcà for providing us with pBKSC plasmid, and W. Etges, F.M. Sene, L. M. Ravazzi, H. E. M. C. Bicudo, F.R. Torres, C.R. Ceron, and M. Manfrin for providing us with strains of *Drosophila*. This research was supported by FAPESP (Grants nos. 95/7192-2 and 00/1371-0 to C.M.A.C and 97/14646-5 to L.M.A) and CNPq. Revision of the English language text was carried out by Peter James Harris.

References

- Arcà, B., Savakis C., 2000. Distribution of the transposable element *Minos* in the genus *Drosophila*. *Genetica* 108, 263-267.
- Baffi, M.A., Ceron, C.R., 2002. Molecular analysis of the rDNA ITS1 intergenic spacer in *Drosophila mulleri*, *D. arizonae* and their hybrids. *Biochem. Genet.* 40, 411-421.

- Clark, J.B., Silva, J.C., Kidwell, M.G., 2002. Evidence of horizontal transfer of *P* transposable elements. In: Horizontal Gene Transfer. M. Syvanen and C.I. Kado, eds. Academic Press, San Diego, pp. 161-171.
- Clark, J.B., Kidwell, M.G., 1997. A phylogenetic perspective on *P* transposable element evolution in *Drosophila*. Proc Natl. Acad. Sci. USA. 94(21), 11428-11433.
- Capy, P., Anxolabéhère, D., Thierry L., 1994. The strange phylogenies of transposable elements: are horizontal transfer the only explanation? Trends Genet. 10, 7-12.
- Drabek, D., Zagoraiou, L., de Wit, T., Langeveld, A., Roumpaki, C., Mamalaki, C., Savakis, C. and Grosveld, F., 2003. Transposition of *Drosophila hydei* Minos transposon in mouse germ-line. Genomics 81, 108-111.
- Durando, C.M., Baker, R.H., Etges, W.J., Heed, W., Wasserman, M., DeSalle, R., 2000. Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. Mol. Phylog. Evol. 16, 296-307.
- Engels, W.R. 1989. *P* elements in *D. melanogaster*. Pp. 437-483 in De Berg and M.M. Howe, eds. American Society for Microbiology, Washington, D.C.
- Franz, G., Loukeris, T.G., Dialektaki, G., Thompson, C.R.L., Savakis, C., 1994. Mobile *Minos* elements from *Drosophila hydei* encode two-exon transposase with similarity to paired DNA-binding domain. Proc. Natl. Acad. Sci. USA 91, 4746-4750.
- Franz, G., Savakis, C., 1991. *Minos*, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposon. Nucleic Acids Res. 19, 6646.
- Gladstein, D.S., Wheeler, W.C., 1997. POY: the optimization of alignment characters. Program and documentation. New York, NY. Available at: ftp.amnh.org/pub/molecular. Documentation by D. James and W. Wheeler (POY: "Phylogeny Reconstruction Via Direct Optimization of DNA Data" native code version 2.0 by David Gladstein and Ward Wheeler).
- Graur, D., Li, W.H. 2000. Fundamentals of molecular evolution. Sinauer Associates, Sunderland, U.S.A., second edition, pp. 1-478.
- Grimaldi, D. A., 1987. Amber fossil Drosophilidae (Diptera) with particular reference to the Hispaniolan taxa. Am. Mus. Novitates 2880, 1-23.
- Grimaldi, D. A., 1988. Relicts in the Drosophilidae (Diptera). In: Zoogeography of Caribbean insects, edit by J.K. Liebherr. Ithaca, Cornell University Press, pp. 183-213.
- Hartl, D.L., Lozovkaya, E.R., Lohe, R. 1997. What is restricts the activity of *mariner*-like transposable element. Trends Genet. 13, 197-201.
- Hasegawa, M., Kishino, H.; Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 21, 160-174.



- Jordan, I.K., Matyunina V.L., McDonald, J.F., 1999. Evidence for recent horizontal transfer of long terminal repeat retrotransposon. *Proc. Natl. Acad. Sci. USA* 96, 12621-12625.
- Jowett, T., 1986. Preparation of nucleic acids. In: Roberts DB (ed) *Drosophila: a Practical Approach*, edited by D.B. Roberts, IRL Press, Oxford, pp. 275-286.
- Kapetanaki, M.G., Loukeris, T.G., Livadaras, I., Savakis, C., 2002. High frequencies of *Minos* transposon mobilization are obtained in insects by using in vitro synthesized mRNA as a source of transposase. *Nucleic Acids Res.* 30, 3333-3340.
- Kidwell, M.G., 1993. Lateral transfer in natural populations of eukaryotes. *Annu. Rev. Genet.* 27, 235-256.
- Klinakis, A.G., Zagoraiou, L., Vassilatis, D.K., Savakis, C., 2000. Genome-wide insertional mutagenesis in human cells by *Drosophila* mobile element *Minos*. *EMBO* 1, 416-421.
- Loreto, E.L.S., Valente V.L., Zaha A., Silva, J.C., Kidwell, M.G., 2001 *Drosophila mediopunctata* P elements: A new example of horizontal transfer. *J. Heredity* 375-381.
- Machado L.P.B, Madi-Ravazzi L, Tadei W.J. 1999. Reproductive and chromosome studies in *Drosophila* species of *buzzatii* cluster. *Genet. Mol. Biol.* 22, 386.
- Madi-Ravazzi L, Bicudo H.E.M, Manzato J.A 1997. Reproductive compatibility and chromosome pairing in the *Drosophila buzzatii* complex. *Cytobios* 89, 21-30.
- Magalhães, L. E. 1962. Notes on the taxonomy, morphology and distribution of *saltans* group of *Drosophila*, with description of four new species. *Univ. of Texas Publ.* 6205, 135-154.
- Manfrin, M. H., De Brito, R.O.A., Sene, F. M., 2001. Systematics and evolution of the *Drosophila buzzatii* (Diptera: Drosophilidae) cluster using mtDNA. *Ann. Entomol. Soc. Am.* 94, 333-346.
- Maruyama, K., Hartl, D.L., 1991. Evidence for interspecific transfer of the transposable element *mariner* between *Drosophila* and *Zaprionus*. *J. Mol. Evol.* 33, 514-524.
- Mortan, B.R., 1993. Chloroplast DNA codon use: evidence for selection at the *psb a* locus based on tRNA availability. *J. Mol. Evol.* 37, 273-280.
- Moryama, E.N., Gojobori, T. 1992. Rates of synonymous substitution and base composition of nuclear genes in *Drosophila*. *Genetics* 130, 855-864.
- O'Brochta, D.A., Sethuraman, N., Wilson, R., Hice, R.H., Pinkerton, A.C., Levesque, C.S., Bideshi, D.K., Jasinskiene, N., Coater, C.J., James, A.A., Lehane, M.J., and Alinson, P.W., 2003. Gene vector and transposable element behavior in mosquitoes. *J. Exp. Biol.* 206, 3823-3834.
- Page R.D.M. 1994. Parallel phylogenetic: reconstructing the history of host-parasite assemblages. *Cladistics* 10, 155-173.



- Patterson, J.T., Stone, W.S., 1952. Evolution in the genus *Drosophila*. Macmillan Co., New York.
- Pinsker W, Haring E, Hangemann S, Miller W.J. 2001. The evolutionary life history of *P* transposons: from horizontal invaders to domesticated neogenes. *Chromosoma* 110, 148-158.
- Posada D, Cradall K.A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817-818.
- Robertson, H.M., Soto-Adames, F.N., Walden, K.K., Avancini, R.M., Lampe, D.J., 2002. The *mariner* transposon in animals: horizontally jumping genes. In: Horizontal Gene Transfer. M. Syvanen, and C.I. Kado, eds. Academic Press, San Diego, CA.
- Rozas J, Rozas R 1997. DnaSP version 2.0: a novel software package for extensive molecular population genetics analysis. *Comut. Appl. Biosci.* 13, 307-311.
- Russo, C.A., Takezaki, N., Nei, M., 1995. Molecular phylogeny and divergence times of drosophilid species. *Mol. Biol. Evol.* 12, 391-404.
- Rodríguez-Trelles, F., Alarcón, L., Fontdevila, A. 2000. Molecular evolution and phylogeny of the *buzzatii* complex (*Drosophila repleta* group): a maximum-likelihood approach. *Mol. Biol. Evol.* 17, 1112-1122.
- Ruiz, A. Cansian, A.M., Kuhn, G.C.S, Alves, M., Sene, F.M., 2000. The *Drosophila serido* speciation puzzle: putting new pieces together. *Genetica* 108, 217-227.
- Russo, C.A.; Takezaki, N., Nei, M. 1995. Molecular phylogeny and divergence times of drosophilid species. *Mol. Biol. Evol.* 12, 391-404.
- Sasakura, Y., Awazu, S., Chiba, S., Satoh, N. 2003. Germ-line transgenesis of *Tc1/mariner* superfamily transposon *Minos* in *Ciona intestinalis*. *Proc. Natl. Acad. Sci. USA.* 100, 7726-7730.
- Shields, D.C., Sharp, P.M. Hiddins D.G and Wrigt. 1988. Silent sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. *Mol. Biol. Evol.* 5, 704-716.
- Shimizu, K., Kamba, M., Sonobe, H., Kanda, T., Klinakis, A.G., Savakis, C., Tamura, T., 2000. Extrachromosomal transposition of the transposable element *Minos* occurs in embryos of the silkworm *Bombyx mori*. *Insect Mol. Biol.* 9, 277-281.
- Tatarenkov, A., Ayala, F., 2001. Phylogenetic relationships among species groups of the *virilis-repleta* radiation of *Drosophila*. *Mol. Phylogen. Evol.* 21, 327-331.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680.
- Throckmorton, L. 1975. The phylogeny, ecology and geography of *Drosophila*. In: R.C. King, ed. *Handbook of Genetics*. vol. 3. Plenum, New York, pp. 421-469.

Throckmorton, L.H., 1982. Pathways of evolution in the genus *Drosophila* and the founding of *repleta* group. In: Ecological Genetics and Evolution: The cactus-yeast- *Drosophila* Model System (J. S. F. Barker and W. T. Starmer, Eds) p. 33-48. Academic press, New York.

Vilela, C. R., 1983. A revision of *Drosophila repleta* species group (Diptera: Drosophilidae). Rev. Bras. Ent. 1-114.

Wasserman, M., 1982. Cytological evolution in *Drosophila repleta* species group. In Ecological Genetics and Evolution: The cactus-yeast- *Drosophila* Model System (J. S. F. Barker and W. T. Starmer, Eds), p. 49-64. Academic press, New York.

Wasserman, M., 1992. Cytological evolution in *Drosophila repleta* species group. In: *Drosophila* inversion polymorphism (C.B Krimbas and J. R. Powell, Eds), p. 455-552 CRC Press, Boca Raton, F.L

Wheeler, W.C., 1995. Sequence alignment, parameter sensitivity and the phylogenetic analysis of molecular data. Syst. Biol. 44, 321-331.

Wheeler, W.C., 1996. Optimization alignment: The end of multiple sequence alignment in phylogenetics? Cladistics 12, 1-9.

Wright, F., 1990. The effective number of codons used in gene. Gene 87, 23-29.

Zagoraiou, L., Drabek, D., Alexaki, S., Guy, J.A., Klinakis, A.G., Langeveld, A., Skavdis G., Mamalak, C., Grasveld, F., Savakis, C., 2001. In vivo transposition of *Minos*, a *Drosophila* mobile element, in mammalian tissues. Proc. Natl. Acad. Sci. USA 11474-11478.



Table 1. Taxonomic list of *Drosophila* species used in this study (according to Durando, 2000 and Magalhães, 1962).

Group	Subgroup	Cluster	Species	Location	Stock number	date
saltans	<i>cordata</i>		<i>D. neocordata</i>	Minas Gerais, Brazil		1959
	<i>elliptica</i>		<i>D. emarginata</i>	Vera Cruz, México		1962
	<i>parasaltans</i>		<i>D. parasaltans</i>	Tapuruquara, AM, Brazil		1962
			<i>D. subsaltans</i>	Belém, PA, Brazil		1959
	<i>sturtevantii</i>		<i>D. sturtevantii</i>	Matlapa, San Luis Potosí, Mexico		1998
			<i>D. austrosaltans</i>	Pirassununga, SP, Brazil		1959
	<i>saltans</i>		<i>D. saltans</i>	Chilpancingo, Mexico		1962
repleta	<i>mercatorum</i>		<i>D. paranaensis</i>	Novo Horizonte, SP, Brazil		1998
			<i>D. mercatorum</i>	Novo Horizonte , SP, Brazil		1998
	<i>mulleri</i>		<i>D. aldrichi</i>	Tamaulipas, Mexico	15081125.0	
			<i>D. mulleri</i>	Tucson, Arizona, USA	150811371.7	
			<i>D. wheeleri</i>	Catalina Island, Los Angeles, USA	150811501.3	
	<i>mulleri</i>	<i>mojavensis</i>	<i>D. navojoa</i>	Tehuantepec –Mexico	150811374.1	
			<i>D. arizonae</i>	Tucson, Arizona, USA	A 1015	
			<i>D. mojavensis</i>	Grand Canyon, Arizona, USA	A870	
	<i>buzzatii</i>		<i>D. koepferae</i>	Tapia-Tucumam, Argentina		
			<i>D. serido</i>	Milagres, BA, Brasil	150811431.3	
			<i>D. buzzatii</i>	Novo Horizonte, SP, Brazil		1998
			<i>D. gouveai</i>	Morro do Chapéu, MG, Brazil		
			<i>D. antonietae</i>	Bela Vista, MS, Brazil		1990
			<i>D. seriema</i>	Serra do Cipó, MG, Brazil		1990
	<i>longicornis</i>		<i>D. longicornis</i>	Tucson, Arizona, USA	150811311.8	
			<i>D. pachuca</i>	Chapingo, Mexico	150811391.0	
			<i>D. propachuca</i>	Hidalgo, Mexico	150811411.1	
			<i>D. hexastigma</i>	Zapotitlan, Puebla, Mexico	150811302.2	
			<i>D. spenceri</i>	Gyaymas, Sonora, Mexico	150811441.2	
	<i>eremophila</i>		<i>D. eremophila</i>	Zapotitlan, Puebla, Mexico	150811292.1	
			<i>D. mettleri</i>	Tucson, Arizona, USA	CAT397	
			<i>D. nigricruria</i>	Novo Horizonte, SP, Brazil		1998
			<i>D. nigroespiracula</i>	Tucson, Arizona, USA	15081150311	1978
<i>hydei</i>			<i>D. anceps</i>	Infernillo, Michoacán, Mexico	150815031.1	
			<i>D. ritae</i>	Zapotitlan dos Salinas, Puebla, Mexico		2002
			<i>D. hydei</i>	Encarnación, Paraguay		2000

Table 2. Genetic distances among *Minos* nucleotide sequences of different *Drosophila* species of the *replete* and *saltans* group calculated using HKY (Hasegawa *et al.*, 1985).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1. <i>D. willistoni</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2. <i>D. saltans</i>	0.06	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3. <i>D. saltans (GB)</i>	0.10	0.08	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4. <i>D. subsaltans</i>	0.11	0.09	0.09	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5. <i>D. austrosaltans</i>	0.07	0.08	0.09	0.11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6. <i>D. sturtevantii</i>	0.06	0.04	0.08	0.09	0.07	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7. <i>D. emarginata</i>	0.11	0.10	0.10	0.12	0.13	0.09	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8. <i>D. hydei</i>	0.11	0.10	0.02	0.09	0.12	0.09	0.13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9. <i>D. hydei (GB)</i>	0.09	0.03	0.03	0.10	0.11	0.08	0.10	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10. <i>D. mercatorum</i>	0.12	0.12	0.10	0.15	0.13	0.12	0.14	0.13	0.12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11. <i>D. paranaensis</i>	0.10	0.10	0.10	0.13	0.11	0.09	0.12	0.12	0.11	0.09	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12. <i>D. buzzatii</i>	0.10	0.09	0.10	0.11	0.11	0.09	0.12	0.11	0.09	0.14	0.12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
13. <i>D. koepferae</i>	0.14	0.14	0.11	0.15	0.15	0.13	0.17	0.14	0.13	0.18	0.15	0.07	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
14. <i>D. serido</i>	0.10	0.09	0.09	0.11	0.11	0.08	0.12	0.11	0.10	0.14	0.12	0.07	0.07	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15. <i>D. gouveai</i>	0.11	0.11	0.11	0.13	0.13	0.10	0.13	0.13	0.10	0.16	0.13	0.04	0.08	0.04	-	-	-	-	-	-	-	-	-	-	-	-	-
16. <i>D. antonietae</i>	0.09	0.09	0.10	0.11	0.11	0.08	0.11	0.11	0.09	0.14	0.11	0.03	0.05	0.02	0.03	-	-	-	-	-	-	-	-	-	-	-	-
17. <i>D. seriema</i>	0.14	0.14	0.24	0.15	0.13	0.14	0.18	0.17	0.17	0.16	0.18	0.08	0.09	0.07	0.08	0.09	-	-	-	-	-	-	-	-	-	-	-
18. <i>D. mulleri</i>	0.16	0.14	0.06	0.16	0.18	0.14	0.16	0.09	0.08	0.17	0.16	0.15	0.17	0.15	0.16	0.14	0.16	-	-	-	-	-	-	-	-	-	-
19. <i>D. wheeleri</i>	0.12	0.10	0.03	0.13	0.13	0.10	0.12	0.06	0.05	0.13	0.13	0.12	0.15	0.12	0.13	0.12	0.18	0.10	-	-	-	-	-	-	-	-	-
20. <i>D. aldrichi</i>	0.13	0.12	0.03	0.14	0.14	0.12	0.13	0.05	0.05	0.14	0.14	0.13	0.15	0.13	0.14	0.12	0.18	0.11	0.04	-	-	-	-	-	-	-	-
21. <i>D. navojou</i>	0.14	0.12	0.02	0.13	0.13	0.11	0.12	0.06	0.05	0.15	0.14	0.13	0.13	0.13	0.14	0.12	0.17	0.11	0.06	0.06	-	-	-	-	-	-	-
22. <i>D. arizonae</i>	0.13	0.11	0.03	0.13	0.15	0.11	0.13	0.06	0.05	0.16	0.15	0.12	0.17	0.12	0.15	0.12	0.13	0.11	0.05	0.04	0.05	-	-	-	-	-	-
23. <i>D. mojavenensis</i>	0.12	0.10	0.02	0.12	0.13	0.09	0.12	0.05	0.04	0.13	0.12	0.12	0.15	0.11	0.13	0.11	0.14	0.09	0.03	0.04	0.04	0.03	-	-	-	-	-
24. <i>D. mojavenensis (GB)</i>	0.11	0.01	0.01	0.11	0.12	0.09	0.10	0.04	0.04	0.13	0.11	0.11	0.15	0.11	0.11	0.10	0.18	0.09	0.03	0.04	0.04	0.03	0.02	-	-	-	-
25. <i>D. pachuca</i>	0.13	0.12	0.11	0.14	0.12	0.12	0.15	0.13	0.13	0.13	0.12	0.14	0.16	0.13	0.15	0.14	0.16	0.17	0.14	0.15	0.14	0.15	0.13	0.14	-	-	-
26. <i>D. propachuca</i>	0.16	0.13	0.13	0.16	0.15	0.14	0.16	0.14	0.12	0.16	0.15	0.16	0.17	0.15	0.16	0.15	0.16	0.17	0.16	0.17	0.18	0.17	0.15	0.13	0.12	-	-
27. <i>D. anceps</i>	0.21	0.20	0.16	0.22	0.18	0.19	0.23	0.21	0.20	0.23	0.21	0.21	0.21	0.21	0.39	0.20	0.21	0.23	0.21	0.22	0.22	0.20	0.21	0.21	0.21	0.21	-
28. <i>D. spenceri</i>	0.34	0.32	0.26	0.32	0.32	0.34	0.37	0.36	0.34	0.36	0.35	0.34	0.37	0.34	0.27	0.35	0.21	0.36	0.36	0.38	0.35	0.42	0.36	0.34	0.35	0.40	0.41

Table 3. Effective number of codons (Nc), average codons bias (CBI), synonymous substitutions per synonymous sites (d_s) and the ratio between non-synonymous substitutions per non-synonymous and synonymous substitutions per synonymous sites ($d_N:d_s$) in the *Minos* sequence and *Adh* host gene.

Gene, Species group (N° of Sequences)	Nc Mean (range)	CBI Mean (range)	d_N (mean $\pm$ S)	d_s (mean $\pm$ S)	$d_N:d_s$ (mean $\pm$ S)
<i>Minos</i> transposase (exon1)					
<i>saltans</i> (5)	54.18 (47.79-58.40)	0.45 (0.37-0.51)	0.066 $\pm$ 0.001	0.094 $\pm$ 0.038	0.70
<i>repleta</i> (6)	52.87 (50.30-55.19)	0.48 (0.43-0.50)	0.057 $\pm$ 0.040	0.076 $\pm$ 0.038	0.75
Overall Mean	53.53	0.47			
<i>Minos</i> transposase (exon1)					
<i>saltans</i> (5) vs <i>repleta</i> (6)	53.21 (41.09-61.00)	0.45(0.42-0.50)	0.080 $\pm$ 0.036	0.065 $\pm$ 0.040	1.23
<i>Adh</i>					
<i>saltans</i> (5)	44.63 (41.64-47.45)	0.52 (0.49-0.53)	0.032 $\pm$ 0.009	0.130 $\pm$ 0.047	0.24
<i>repleta</i> (6)	42.50 (38.87-45.88)	0.48 (0.43-0.58)	0.022 $\pm$ 0.012	0.241 $\pm$ 0.108	0.09
Overall Mean	43.53	0.50			
<i>Adh</i> (exon2)					
<i>saltans</i> (5) vs <i>repleta</i> (6)	44.44 (39.09-47.45)	0.53(0.49-0.60)	0.064 $\pm$ 0.001	1.052 $\pm$ 0.789	0.13

Figure 1. Ethidium bromide-stained agarose gel of PCR products obtained with *Minos*-specific primers (M3 and M5) used as template genomic DNA from the following species (A): 1. *D. hydei*; 2. *D. parandensis*; 3. *D. mercatorum*; 4. *D. kneyferi*; 5. *D. buzzatii*; 6. *D. serrice*; 7. *D. guenei*; 8. *D. antonietae*; 9. *D. sericea*; 10. *D. aldrichi*; 11. *D. mulieri*; 12. *D. wheeleri*; 13. *D. navajoa*; 14. *D. arizonae*; 16. *D. mafavensis*; 17. *D. pachuca*; 18. *D. propachuca*; 19. *D. hexastigma*; 20. *D. spenceri*; 21. *D. longicornis*; 22. *D. eremophila*; 23. *D. metlieri*; 24. *D. nigricaria*; 25. *D. nigrospiracula*; 26. *D. anceps*; 27. *D. ritoe* and 13 and 30 positive control. (B): 1. *D. azocordata*; 2. *D. emarginata*; 3. *D. parotallanti*; 4. *D. subsaltans*; 5. *D. sturievanti*; 6. *D. antrosaltans*; 7. *D. saltans* and 8. Positive control. The arrows point to the expected 1.1kb fragments.

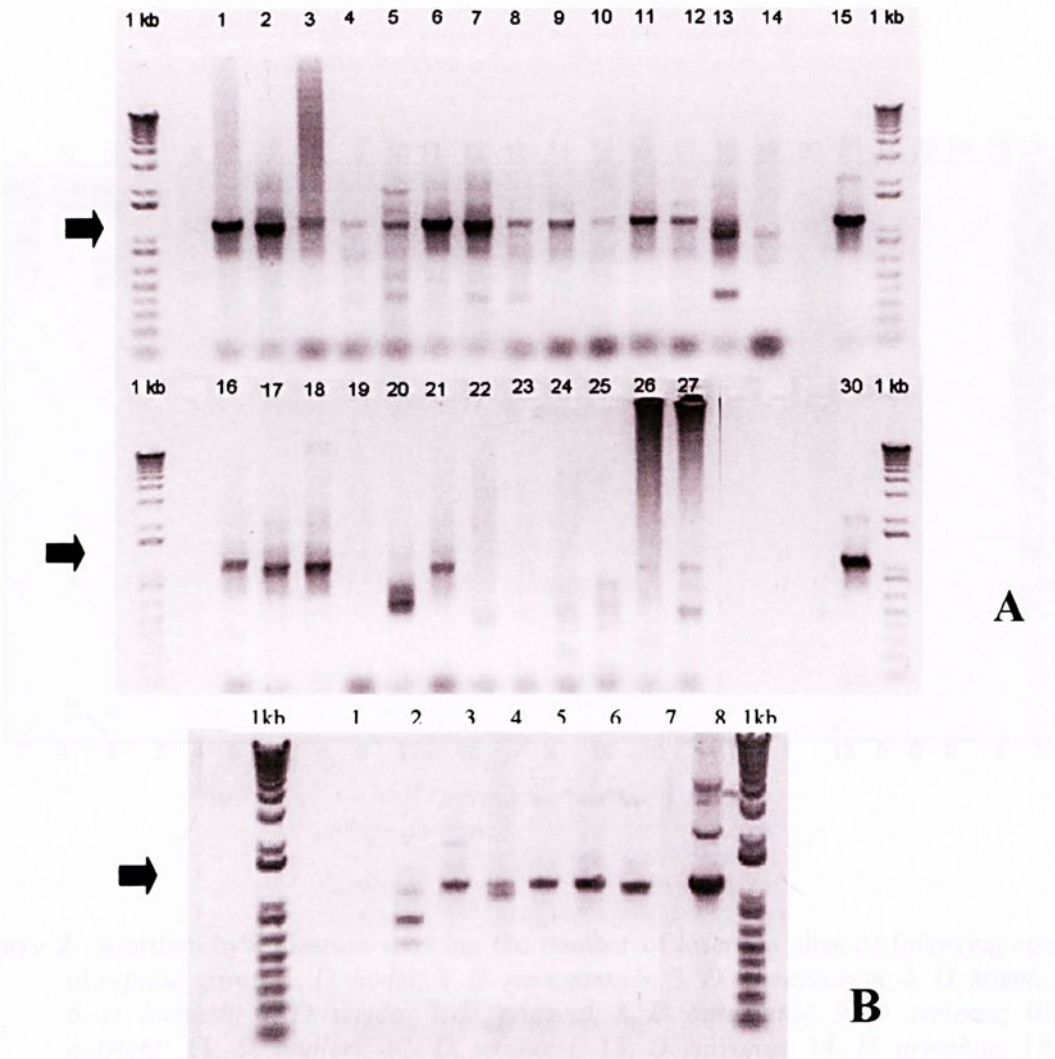


Figure 1. Ethidium bromide-stained agarose gel of PCR products obtained with *Minos* – specific primers (M3 and M5) used as template genomic DNA from the following species **(A)**: 1. *D. hydei*; 2. *D. paranaensis*; 3. *D. mercatorum*; 4. *D. koepferae*; 5. *D. buzzatii*; 6. *D. serido*; 7. *D. gouveai*; 8. *D. antonietae*; 9. *D. seriema*; 10. *D. aldrichi*; 11. *D. mulleri*; 12. *D. wheeleri*; 13. *D. navojoa*; 14. *D. arizonae*; 16. *D. mojavenensis*; 17. *D. pachuca*; 18. *D. propachuca*; 19. *D. hexastigma*; 20. *D. spenceri*; 21. *D. longicornis*; 22 *D. eremophila*; 23. *D. mettleri*; 24. *D. nigricurria*; 25. *D. nigroespiracula*, 26. *D. anceps*, 27. *D. ritae* and 15 and 30 positive control. **(B)**: 1. *D. neocordata*; 2. *D. emarginata*; 3. *D. parasaltans*; 4. *D. subsaltans*; 5. *D. sturtevantii*; 6. *D. austrosaltans*; 7. *D. saltans* and 8. Positive control. The arrows point to the expected 1.1kb fragments.

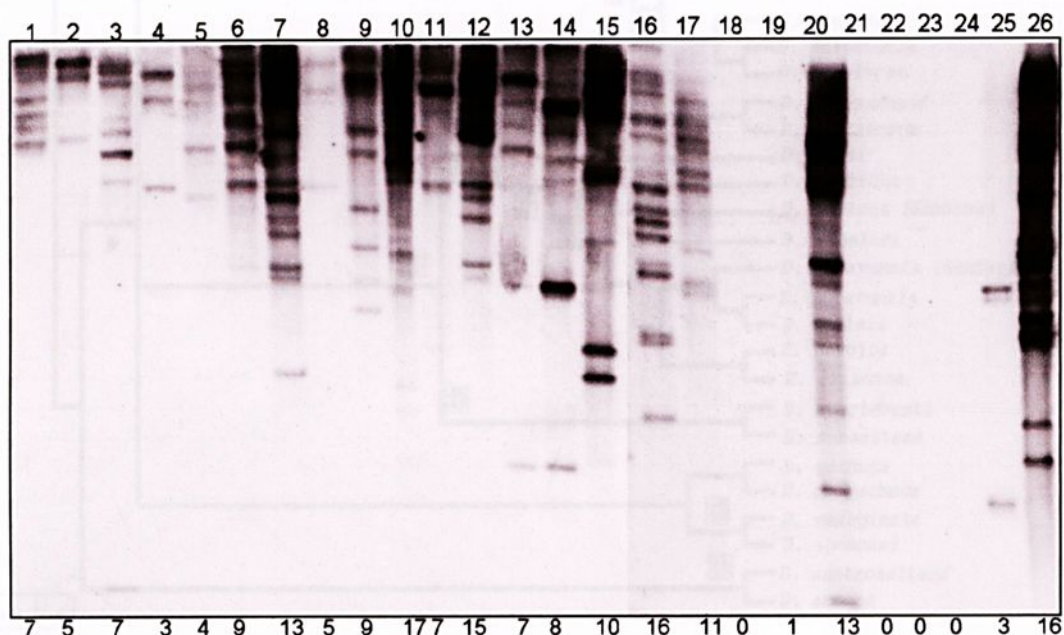


Figure 2. Southern hybridization showing the number of insertion sites of following species of *repleta* group: 1. *D. hydei*; 2. *D. paranaensis*; 3. *D. mercatorum*; 4. *D. koepferae*; 5. *D. buzzatii*; 6. *D. serido*; 7. *D. gouveai*; 8. *D. antonietae*; 9. *D. seriema*; 10. *D. aldrichi*; 11. *D. mulleri*; 12. *D. wheeleri*; 13. *D. navojoa*; 14. *D. arizonae*; 15. *D. mojaviensis*; 16. *D. pachuca*; 17. *D. propachuca*; 18. *D. hexastigma*; 19. *D. spenceri*; 20. *D. longicornis*; 21. *D. eremophila*; 22. *D. mettléri*; 23. *D. nigricruria*; 24. *D. nigrospiracula*; 25. *D. anceps*; 26. *D. ritae*. The DNA of 50 individuals were digested with *Bam* HI and *Xho* I and probed with 1.1 kb fragments amplified from pBKSC plasmid.

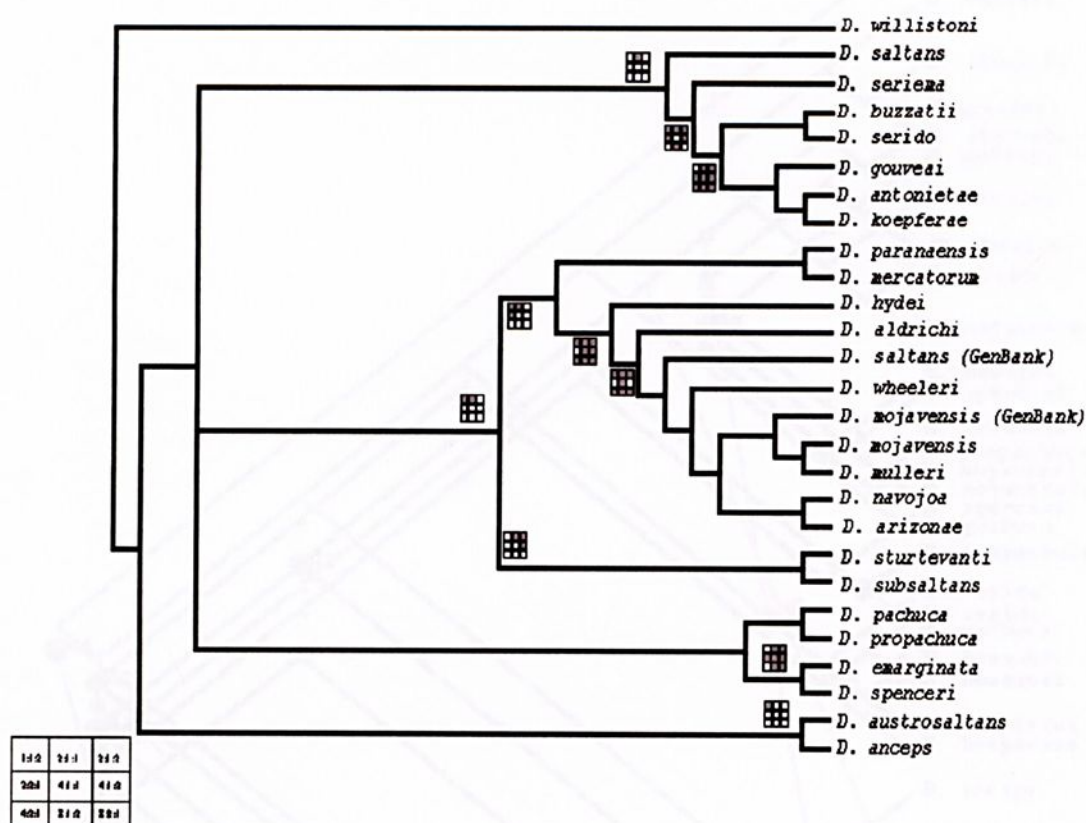


Figure 3. The most parsimonious cladogram of *Minos* sequences (1.1kb amplified by primers M5 and M3) resulted from 1:1:1 parameters costs implemented by POY (Gladstein and Wheeler, 1997). The information from analysis of all data under nine parameter sets (gap: transversion: transition cost ratios 1:1:2; 2:1:1; 2:1:2; 2:2:1; 4:1:1; 4:1:2; 4:2:1; 8:2:1; 8:2:2) is summarized in the sensitivity plots (dark and white squares). The dark fields in the sensitivity plots indicate those parameter sets in which the respective group revealed as monophyletic.

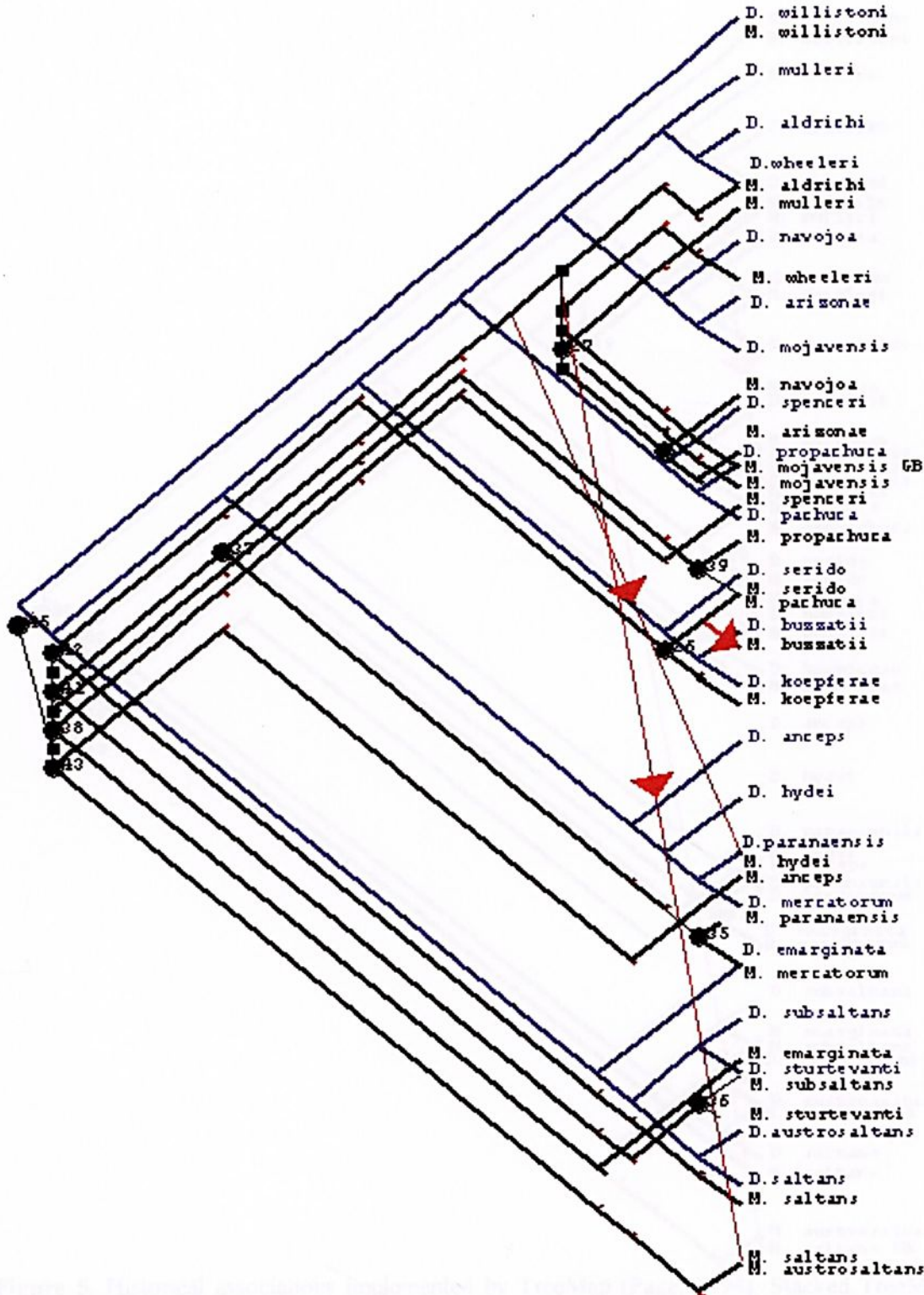


Figure 4. Historical associations implemented by TreeMap (Page, 1994).

Stacked TreeMap reconciliation trees of *Drosophila* phylogeny (blue lines) and *Minos* retrotransposon phylogeny (dark lines). Figure shows 12 events of divergence (circles); 34 of extinctions (interrupted lines); 7 of duplications (squares) and 3 of horizontal transfers (arrows: from an ancestor of the *mulleri* and the *mojavensis* clusters to *D. saltans* and to *D. hydei*, and from *D. serido* to *D. buzzatii*).

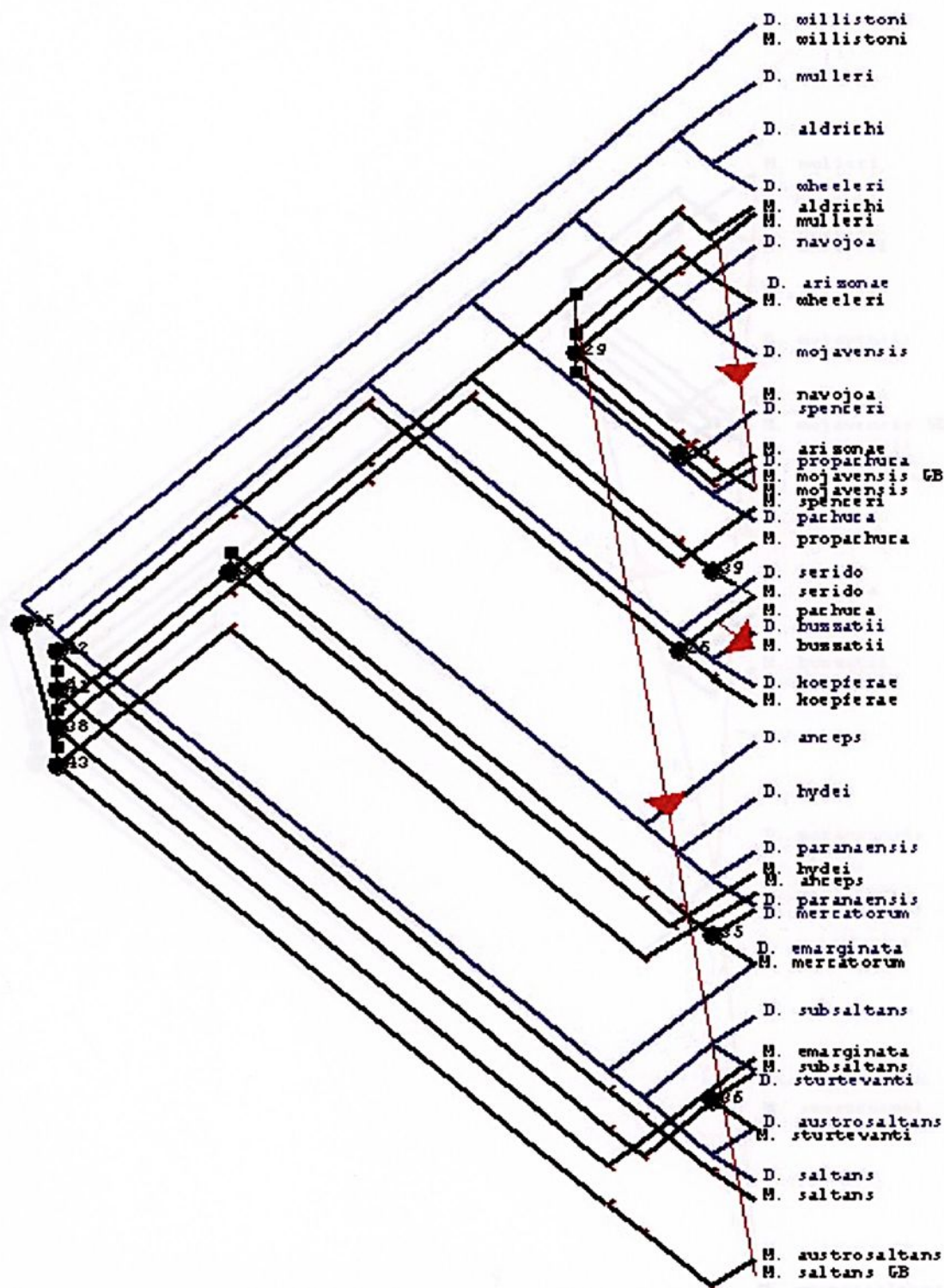


Figure 5. Historical associations implemented by TreeMap (Page, 1994). Stacked TreeMap reconciliation trees of *D.rosophila* phylogeny (blue lines) and *Minos* retrotransposon phylogeny (dark lines). Figure shows 12 events of divergence (circles); 34 of extinctions (interrupted lines); 7 of duplications (squares) and 3 of horizontal transfers (arrows from an ancestor of the *mulleri* and the *mojavenensis* clusters to *D. saltans*; from *D. mulleri* to *D. mojavenensis* and from *D. serido* to *D. buzzatii*).

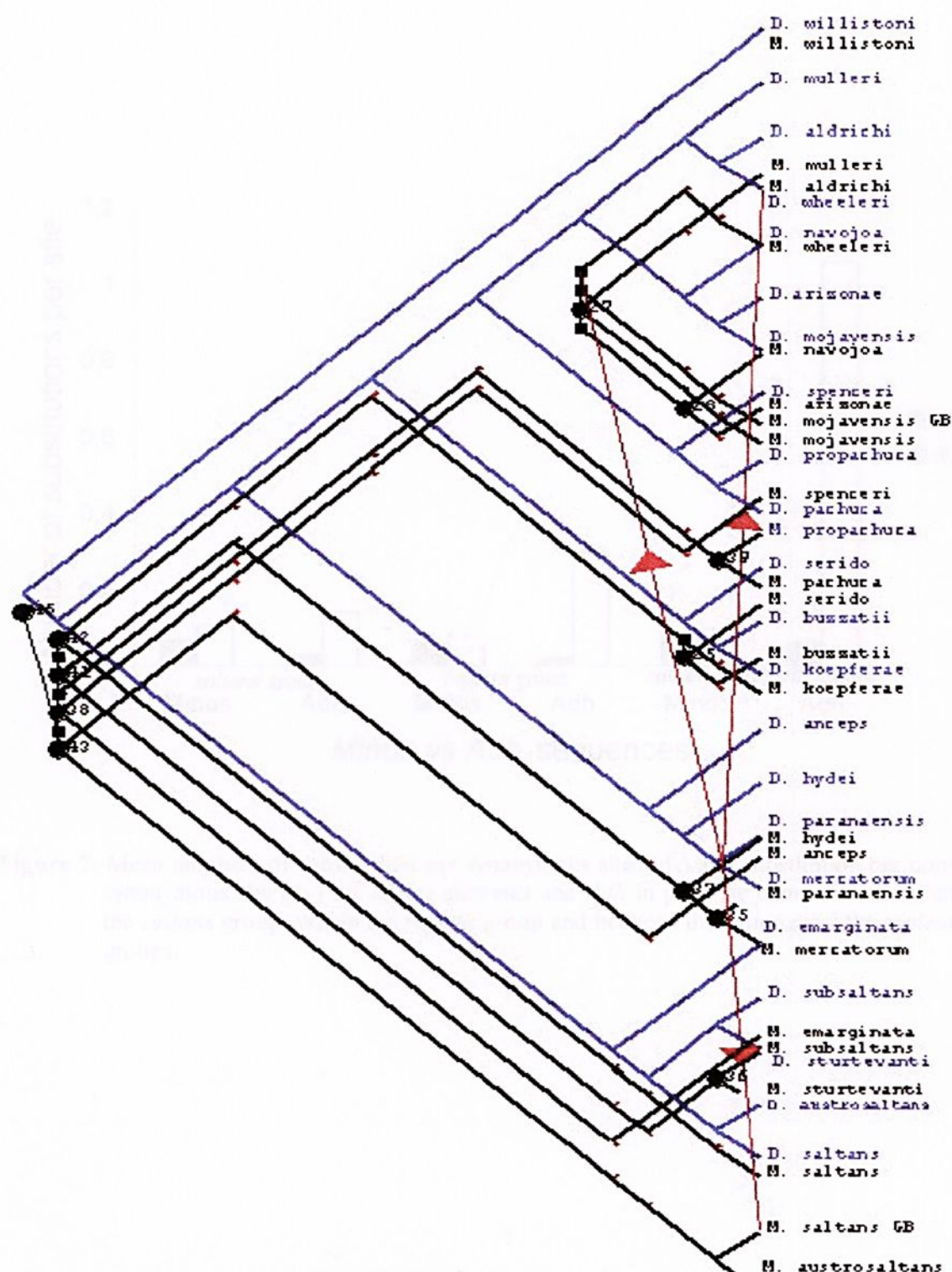


Figure 6. Historical associations implemented by TreeMap (Page, 1994). Stacked TreeMap reconciliation trees of *Drosophila* phylogeny (blue lines) and *Minos* retrotransposon phylogeny (dark lines). Figure shows 12 events of divergence (circles); 28 of extinctions (interrupted lines); 7 of duplications (squares) and 3 of horizontal transfers (arrows: from *D. hydei* to an ancestor of the *mulleri* and the *mojaveensis* clusters; from *D. hydei* to *D. aldrichi* and from *D. hydei* to *D. saltans*).

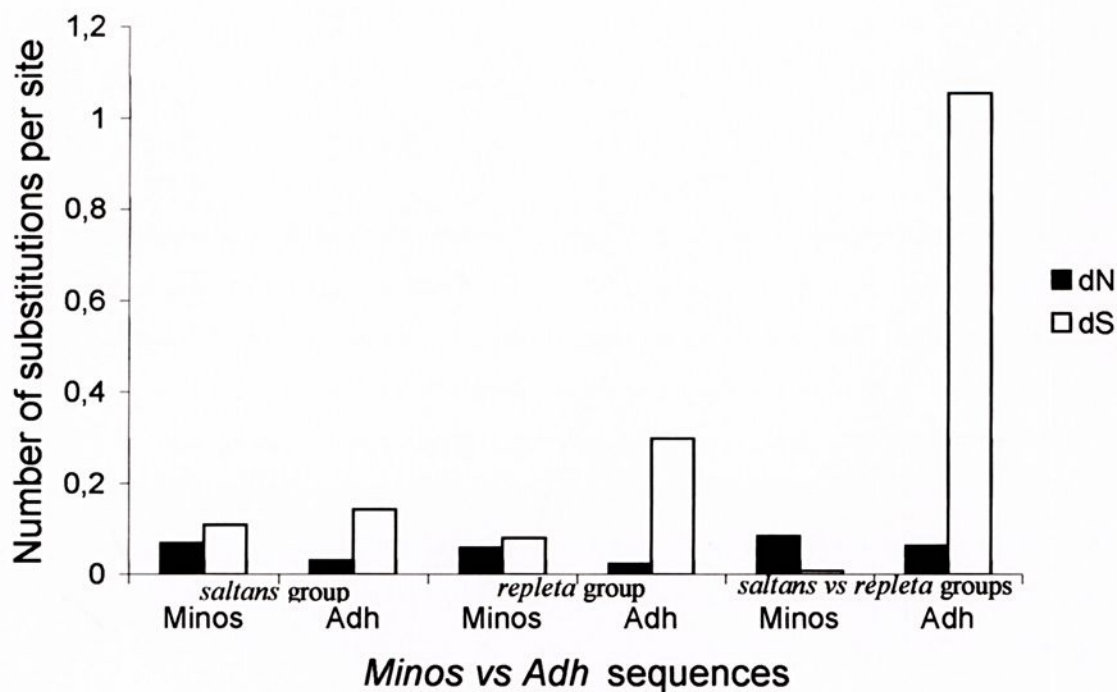


Figure 7. Mean numbers of substitution per synonymous sites (d_S) and substitution per non-synonymous site (d_N) for *Minos* elements and *Adh* in pairwise comparisons within the *saltans* group; within the *repleta* group and between the *saltans* and the *repleta* groups.

IV. DISCUSSÃO

O progresso do conhecimento sobre a estrutura, a função e a dinâmica dos elementos transponíveis, bem como, sobre as possíveis consequências evolutivas de sua presença nos genomas hospedeiros vem ocorrendo em ritmo impressionante e acelerado. Há pouco mais de duas décadas, alguns pesquisadores acreditavam que essas seqüências móveis não passavam de DNA lixo, sem qualquer importância evolutiva, enquanto outros supunham que tinham apenas um efeito deletério, pois alteravam, com sua movimentação, os genomas altamente adaptados e estáveis. Hoje, apesar de existirem muitas questões ainda sem resposta, um panorama geral já pode ser estabelecido, no qual os elementos transponíveis assumem importante papel na evolução das espécies. Sabe-se que a presença desses elementos pode afetar o genoma hospedeiro, aumentando seu tamanho e a variabilidade genética, reestruturando a arquitetura genômica, regulando a expressão gênica e expressando novos genes. Além desses efeitos aparentemente benéficos, a inserção de elementos transponíveis em regiões codificadoras é, na grande maioria das vezes, deletéria e eliminada pela ação da seleção natural ao longo do tempo evolutivo. Uma linha de pesquisa atual relaciona inserções de TEs com doenças humanas, tais como, câncer de mama, hemofilia, neurofibromatose, carcinoma ovário, entre outras (Domingo & Slazinik 1999). Os elementos transponíveis podem facilitar a ocorrência da transposição horizontal de genes.

IV. DISCUSSÃO

Apesar de grande quantidade de informação disponível a respeito dos elementos transponíveis, há muito ainda a ser pesquisado, principalmente a respeito da ocorrência de diferentes famílias em muitas espécies e da frequência de eventos de transferência horizontal. Apesar de existirem várias evidências da ocorrência de tal fenômeno, ainda não foi demonstrado como os TEs podem se mover entre organismos eucariotes distintos. Há, ainda, outras questões igualmente interessantes, tais como: qual é a origem dessas seqüências móveis? Quanto da evolução do genoma hospedeiro é devida a interações com os TEs? Qual a

freqüência da ocorrência do evento de domesticação de TEs? Qual é a importância da transferência horizontal na sobrevivência dos elementos transponíveis? Quais fatores influenciam a ocorrência de tal fenômeno? Por que diferentes espécies hospedeiras possuem diferentes quantidades dessas seqüências? Enfim, são muitas e outras questões que tornam os elementos transponíveis material de pesquisa cada vez mais interessante.

O primeiro capítulo apresentado nesta Tese é uma revisão sobre os principais aspectos de transferência horizontal de elementos transponíveis já descritos. De bactérias ao homem, no qual se pode perceber que a freqüência do fenômeno está diretamente relacionada com a classe de TE predominante nos organismos. Silva *et al.* (2004), em fase do estado de

trabalho. O progresso do conhecimento sobre a estrutura, a função e a dinâmica dos elementos transponíveis, bem como, sobre as possíveis conseqüências evolutivas de sua presença nos genomas hospedeiros vem ocorrendo em ritmo impressionante e acelerado. Há pouco mais de duas décadas, alguns pesquisadores acreditavam que essas seqüências móveis não passavam de DNA lixo, sem qualquer importância evolutiva, enquanto outros supunham que tinham apenas um efeito deletério, pois alteravam, com sua movimentação, os genomas altamente adaptados e estáveis. Hoje, apesar de existirem muitas questões ainda sem resposta, um panorama geral já pode ser estabelecido, no qual os elementos transponíveis assumem importante papel na evolução das espécies. Sabe-se que a presença desses elementos pode afetar o genoma hospedeiro aumentando seu tamanho e a variabilidade genética, reestruturando a arquitetura genômica, regulando a expressão gênica e expressando novos genes. Além desses efeitos aparentemente benéficos, a inserção de elementos transponíveis em regiões codificadoras é, na grande maioria das vezes, deletéria e eliminada pela ação da seleção natural ao longo do tempo evolutivo. Uma linha de pesquisa atual relaciona inserções de TEs com doenças humanas, tais como, câncer de mama, hemofilia, neurofibromatose, carcinoma ovariano, entre outras (Deininger & Blatzer, 1999). Os elementos transponíveis podem, ainda, atuar como vetores naturais da transposição horizontal de genes.

Apesar da grande quantidade de informação disponível a respeito dos elementos transponíveis, há muito ainda a ser pesquisado, principalmente a respeito da ocorrência de diferentes famílias em muitas espécies e da freqüência de eventos de transferência horizontal. Apesar de existirem várias evidências da ocorrência de tal fenômeno, ainda não foi demonstrado como os TEs podem se mover entre organismos eucariotos distintos. Há, ainda, outras questões igualmente interessantes, tais como: qual é a origem dessas seqüências móveis? Quanto da evolução do genoma hospedeiro é devida a interações com os TEs? Qual a

freqüência da ocorrência do evento de domesticação de TEs? Qual é a importância da transferência horizontal na sobrevivência dos elementos transponíveis? Quais fatores influenciam a ocorrência de tal fenômeno? Por que diferentes espécies hospedeiras possuem diferentes quantidades dessas seqüências? Enfim, são essas e outras questões que tornam os elementos transponíveis material de pesquisa cada vez mais interessante.

O primeiro capítulo apresentado nesta Tese é uma revisão sobre os principais eventos de transferência horizontal de elementos transponíveis já descritos, de bactérias ao homem, no qual se pode perceber que a freqüência do fenômeno está diretamente relacionada com a classe de TE predominante nos organismos. Silva *et al.* (2004), em face do modo de transposição e da capacidade de auto-regulação do número de cópias de cada elemento, sugeriram uma escala de ocorrência de transferência horizontal: TEs da Classe II são os que apresentam maior freqüência desse fenômeno, seguidos pelos elementos da Classe I, com LTRs, os quais, por sua vez, apresentam maior ocorrência de transferência horizontal que os elementos da Classe I, sem LTRs. Neste trabalho, avaliamos a ocorrência de transferência horizontal dos retrotransposons *micropia* e *copia* (Classe I com LTRs) e do transposon *Minos* (Classe II) em espécies do gênero *Drosophila*. Os resultados obtidos são concordantes com a proposição de Silva *et al.* (2004), pois a distribuição do transposon *Minos*, entre as espécies dos grupos *saltans* e *repleta*, sugere que esse elemento tenha sido transferido horizontalmente, ao menos três vezes durante sua história evolutiva, entre espécies desses dois grupos, enquanto que a distribuição do retrotransposon *copia* sugere apenas um provável evento de transferência horizontal. Por outro lado, a distribuição do retroelemento *micropia* sugere que este tenha sido transmitido verticalmente ao longo de sua história evolutiva.

Kaplan *et al.* (1985) propuseram que no curso de sua coexistência com os respectivos genomas hospedeiros, os TEs estão sujeitos a curtas explosões de atividade (períodos de invasão) seguidas por longos períodos de inativação. Dessa forma, os elementos que não se expressam há muito tempo estariam condenados à extinção. Sem pressões seletivas mantendo a funcionalidade da seqüência, um elemento inativo irá se degradar devido ao acúmulo de mutações, um processo que finalmente leva à inativação vertical (Miller *et al.*, 1995). Assim, a idéia de que transferência horizontal seja uma estratégia evolutiva que habilita aos elementos transponíveis fugirem da inativação vertical e da extinção vem crescendo no meio científico. Provavelmente, um dos grandes desafios da era genômica seja dimensionar o fenômeno de transferência horizontal entre os seres vivos.



Nesse contexto, um dos objetivos propostos para este estudo foi buscar novas evidências da ocorrência do fenômeno de transferência horizontal de elementos transponíveis. De uma forma particularizada, investigou-se a ocorrência da transmissão horizontal dos elementos *micropia*, *copia* e *Minos* em espécies do grupo *repleta* de *Drosophila*. Além disso, informações que ampliam o entendimento do dinamismo dessas seqüências dentro dos genomas, tais como o número de cópias, o polimorfismo de seqüências e a distribuição ao longo das espécies de um grupo, também foram averiguadas.

Micropia foi o primeiro elemento a ser analisado, estando os resultados de seu estudo apresentados no Capítulo 2. A ampla distribuição e o baixo número de inserções desse elemento no genoma (3 a 17 cópias) sugere que o retrotransposon *micropia* esteja presente no grupo *repleta* antes da diversificação das espécies, tendo sido transmitido verticalmente ao longo das gerações durante a história evolutiva desse grupo de espécies. O número de inserções previamente avaliado na espécie *D. hydei*, a primeira espécie do grupo a ser analisada quanto à presença desse elemento, variou de 2 a 11 nos autossomos e de 50 a 100 no cromossomo Y (Lankenau, 1993). O autor sugeriu que o elevado número de seqüências *micropia* no cromossomo Y era resultante de um processo de amplificação gênica e não da transposição. Duas hipóteses podem justificar a discrepância entre nossos resultados e os de Lankenau (1993). A primeira é que o evento de amplificação possa ser restrito à linhagem de *D. hydei* utilizada pelo autor. A segunda é que a técnica utilizada no nosso estudo para detecção do número de inserções (*Southern blot*) não é ideal para detecção de cópias *in tandem* (resultantes de processo de amplificação), uma vez que os fragmentos gerados pela digestão enzimática, se total, terão o mesmo comprimento e formarão uma única banda no gel. Porém, deve-se lembrar que vários fragmentos de mesmo comprimento geram bandas com maior intensidade de coloração que as demais. Um resultado alternativo seria a detecção de um padrão de bandas em escada, resultante de digestão parcial da seqüência *in tandem*. Neste trabalho, as bandas referentes ao cromossomo Y (presentes apenas em machos) não apresentaram maior intensidade de coloração, nem mesmo um padrão de bandas em escada, indicando que a hipótese da amplificação gênica de *micropia*, restrita à linhagem utilizada por Lankenau (1993), seja a causa mais provável da diferença entre os resultados.

Em um estudo posterior, Lankenau *et al.* (1994) mostraram que as cópias transcionalmente ativas de *micropia* de algumas espécies do grupo *repleta* (*D. hydei*, *D. eohydei*, *D. paranaensis*, *D. mercatorum*, *D. limensis* e *D. buzzatii*) produzem dois tipos de RNA, um RNA *sense* de 5,0 kb e dois RNAs *antisense*, um com de 1,0 kb e outro com 1,6 kb.



A expressão dos RNAs *antisense* é tecido-específica, sendo produzidos apenas nos testículos e em quantidade cerca de 30 vezes superior a do RNA *sense*, sendo o primeiro repressor da expressão do segundo. Os resultados das análises de *Southern blot* deste trabalho mostraram que, além daquelas espécies descritas por Lankenau *et al.* (1994), as espécies *D. koepferae*, *D. serido*, *D. gouveai*, *D. antonietae* e *D. seriema* também produzem RNA *antisense* restrito aos testículos. Contudo, uma única exceção foi detectada: *D. mercatorum* apresentou expressão do RNA *antisense* também na cabeça. O estabelecimento de estratégias de repressão dos mecanismos de controle da expressão diferencial nos tecidos somáticos dos elementos *micropia* presentes em *D. mercatorum* podem ser a causa dessa diferença entre os resultados. Além da expressão diferencial do RNA *antisense* nos tecidos somáticos e germinativos, Lankenau *et al.* (1994) também sugeriram a ocorrência de uma relação entre inserções *micropia* no cromossomo Y (bandas específicas de macho) e expressão do RNA *antisense*. Tal associação não foi verificada em nosso estudo. *D. spenceri*, por exemplo, que possui quatro bandas específicas de macho, não expressa o RNA *antisense*, enquanto que *D. hydei*, que apresenta apenas uma banda específica de macho, e *D. buzzatii*, que não possui bandas macho-específicas, expressam RNA *antisense* nos testículos.

Os TEs são classificados de acordo com as características estruturais e moleculares das suas seqüências, podendo ser agrupados em classes, subclasses, superfamílias, famílias e subfamílias. A classificação em subfamílias é baseada na identidade de nucleotídeos, sendo que elementos de uma mesma subfamília podem apresentar cerca de 30 a 40% de diversidade em suas seqüências de nucleotídeos. Foram identificadas pelo menos duas subfamílias *micropia* neste trabalho. A primeira subfamília, formada por elementos *micropia* que apresentam divergência de nucleotídeos de 1 a 18%, foi identificada em uma espécie do subgrupo *hydei* (*D. hydei*), duas do subgrupo *mercatorum* (*D. mercatorum* e *D. paranaensis*) e nas cinco espécies do *cluster buzzatii* (*D. buzzatii*, *D. serido*, *D. gouveai*, *D. antonietae* e *D. seriema*). A segunda subfamília é formada por elementos com divergência de 12% e ocorrem em duas espécies do subgrupo *mulleri*, *cluster longicornis* (*D. longicornis* e *D. propachuca*). Entre essas duas subfamílias, a taxa de divergência variou de 30% a 58%. Embora as taxas de divergência dos elementos das duas subfamílias estejam compreendidas dentro dos limites propostos por Capy *et al.* (1998), sugerimos que esses elementos pertençam a diferentes subfamílias, uma vez que a capacidade de expressão do RNA *antisense* é diferente para os dois grupos de seqüências. O elemento *micropia* nos subgrupos *hydei*, *buzzatii* e *mercatorum* produz o RNA *antisense*, porém essa produção é inexistente no *cluster longicornis*.



O segundo artigo apresentado nesta Tese amplia os conhecimentos a respeito da distribuição, polimorfismo de sequência e taxas de evolução das sequências *copia* inicialmente identificadas em apenas duas espécies do grupo *repleta* (*D. buzzatii* e *D. koepferae*) por Jordan & McDonald (1998a). As análises de PCR e de *Southern blot* mostraram que o elemento *copia*, apesar de amplamente distribuído entre os drosofilídeos, apresenta uma distribuição heterogênea dentro do grupo *repleta*, estando presente em 11 das 26 espécies analisadas. Por outro lado, no cluster *buzzatti* a distribuição é contínua, uma vez que seis das sete espécies do cluster (*D. borborema* não esteve disponível para o estudo) apresentam o elemento *copia*.

Além da distribuição, neste trabalho também foi avaliado o número de inserções do elemento *copia* nas espécies do grupo *repleta*, o que até o momento era desconhecido. Esse número variou de 1 a 9 cópias, valores significativamente inferiores aos detectados em espécies do grupo *melanogaster* (20 a 60 cópias, de acordo com Csinki & McDonald, 1995). Além do baixo número de inserções, foram detectadas sequências *copia* com menor grau de conservação de sítios importantes para a sua atividade transcricional em espécies do grupo *repleta* que em espécies do grupo *melanogaster* (Jordan *et al.*, 1999). Esta informação indica que as sequências *copia* presentes nas espécies do grupo *repleta* são mais antigas do que aquelas presentes nas espécies do grupo *melanogaster*.

As análises das sequências *copia* permitiram a construção de uma possível filogenia desse TE no grupo *repleta*. A análise de parcimônia formou dois clados, um contendo as sequências das espécies *D. serido* e *D. buzzatii* e outro contendo as sequências das espécies *D. antonietae*, *D. koepferae*, *D. gouveai* e *D. seriema*. As análises de sequências permitiram a identificação de uma única subfamília nesse grupo de espécies, além da detecção da existência de forças seletivas atuando na conservação de sequências da ULR essenciais à atividade do elemento *copia*.

A comparação entre a filogenia das espécies hospedeiras e a das sequências do TE permitiu a identificação de possíveis eventos de divergência, duplicação, perda estocástica e transferência horizontal. A análise de associações históricas entre o elemento *copia* e as espécies hospedeiras produziu duas árvores igualmente parcimoniosas; ambas sugerem a ocorrência de três perdas estocásticas, um evento de divergência e um evento de transferência horizontal para explicar a atual distribuição desse retrotransposon nas espécies do grupo *repleta*. Um ponto comum entre as duas filogenias apresentadas para o grupo é que as sequências *copia* foram perdidas em todas as espécies do grupo *repleta*, com exceção, do

cluster buzzatii. Essa hipótese fundamenta-se no fato que o programa TreeMap assume que o elemento *copia* do grupo externo (*D. melanogaster*) e o do grupo *repleta* formam um grupo monofilético, mas essa não é necessariamente a única possibilidade. Outra hipótese seria que o *cluster buzzatii* recebeu a sequência *copia* por transferência horizontal de uma espécie não utilizada durante esse estudo. Dentro desse *cluster*, duas hipóteses poderiam explicar a evolução desse elemento. Na primeira, a sequência *copia* foi perdida no ancestral da espécie *D. buzzatii*, sendo posteriormente re-introduzida por transferência horizontal a partir de *D. serido*. Na segunda, a sequência *copia* foi perdida no ancestral de *D. serido*, sendo posteriormente re-introduzida horizontalmente de *D. buzzatii*. Apesar do evento de transferência horizontal ser menos freqüente na Classe I de elementos transponíveis, um possível evento de transferência do elemento *copia* entre *D. melanogaster* e *D. willistoni* havia sido sugerido por Jordan *et al.* (1999).

O terceiro elemento analisado nesta Tese foi o transposon de DNA *Minos*, que apresentou ampla distribuição e baixo número de inserções no genoma das espécies do grupo *saltans* (subgênero *Sophophora*) e *repleta* (subgênero *Drosophila*). As análises de sequências permitiram estabelecer as relações filogenéticas entre as sequências *Minos* desses dois grupos de espécies. Várias discrepâncias entre a filogenia dos hospedeiros e a filogenia do TE levaram à utilização do programa TreeMap, que realiza associações históricas. Centenas de possíveis árvores foram construídas, sendo selecionada três árvores, que representaram os principais eventos evolutivos ocorridos durante a evolução do elemento. Vários eventos de transferência poderiam ter ocorrido entre as espécies do grupo *repleta*: entre *D. mulleri* e *D. mojaviensis* e entre *D. serido* e *D. buzzatii*, entre a espécie *D. emarginata* (grupo *saltans*) e *D. spenceri* (grupo *repleta*) e, ainda, entre *D. saltans* (grupo *saltans*) e uma linhagem ancestral dos clusters *mulleri* e *mojavensis*, ou entre a linhagem ancestral de *D. hydei* e uma linhagem ancestral das espécies dos clusters *mulleri* e *mojavensis*. Uma hipótese alternativa seria que uma sequência *Minos* presente na linhagem ancestral de *D. hydei* tenha sido horizontalmente transferida para as espécies dos clusters *mulleri* e *mojavensis* (grupo *repleta*) e também para a espécie *D. saltans*. Essa última proposta está de acordo com os tempos de diversificação das espécies. De acordo com Russo *et al.* (1995), a linhagem que originou *D. hydei* divergiu à cerca de 14 milhões de anos, enquanto que as espécies pertencentes aos clusters *mulleri* e *mojavensis* divergiram a cerca de 6 milhões de anos. Adicionalmente, o tempo estimado para diversificação da espécie *D. saltans* é de 4 milhões de anos (Clark & Kidwell, 1997). Dessa forma, *D. hydei* poderia ter transferido o elemento *Minos* para o ancestral dos clusters *mulleri*



e *mojavensis* a cerca de 6 milhões de anos e para *D. aldrichi* e *D. saltans* em qualquer época entre 4 milhões de anos até o presente.

1993) Comparações da taxa de substituições sinônimas em sítios sinônimos (d_s) e de substituições não sinônimas em sítios não sinônimos (d_n), entre o elemento *Minos* e o gene *Adh*, foram realizadas com o objetivo de comparar a ocorrência de transferência horizontal, sugeridas pelas análises de parcimônia e de associações históricas. Por essa análise, transmissão vertical é assumida se a divergência do TE é maior do que aquela observada para o gene hospedeiro. Por outro lado, se a divergência do Te é significativamente menor do que a do gene hospedeiro o fenômeno de transferência horizontal pode ser invocado para explicar tal discrepância. As estimativas do valor médio de d_s das seqüências *Minos* no grupo *repleta* foram 3.7 vezes menor do que as detectadas para o gene da *Adh*. Adicionalmente, as estimativas dos valores médios de d_s das seqüências *Minos* entre os grupos *saltans* e *repleta* foram 15 vezes menores do que aqueles observados para o gene da *Adh* do hospedeiro. Dessa forma, esses resultados confirmam a hipótese de transferência horizontal dentro do grupo *repleta* e também entre espécies do grupo *repleta* e *saltans* sugeridas anteriormente pelas análises de associações históricas e pela análise filogenética.

O grupo *repleta* é um dos maiores grupos da radiação *Drosophila* e, por apresentar espécies com diferentes níveis de isolamento reprodutivo, vem sendo muito utilizado em estudos de especiação. Uma das classificações existentes para esse grupo divide-o em 5 subgrupos (Vilela, 1983): *hydei*, *mercatorum*, *mulleri*, *fasciola* e *repleta*. Muitos estudos morfológicos, citogenéticos, enzimáticos e moleculares vêm sendo empregados na tentativa de esclarecer as relações filogenéticas do grupo e, com isso, um panorama geral já foi criado, porém, muito ainda resta a ser esclarecido.

2003) O *cluster buzzatii* (complexo *buzzatii*, subgrupo *mulleri*, grupo *repleta*) atualmente é composto por sete espécies (*D. buzzatii*, *D. koepferae*, *D. borborema*, *D. serido*, *D. seriema*, *D. gouveai* e *D. antonietae*). De todas características morfológicas estudadas, apenas a morfologia do edeago (órgão copulador masculino) pode ser utilizada como caráter diagnóstico para a identificação das espécies do *cluster* (Silva & Sene, 1991; Tidon-Sklorz & Sene, 2001). As espécies *D. seriema*, *D. gouveai* e *D. antonietae* foram elevadas à categoria de espécie apenas recentemente por (Tidon-Sklorz & Sene, 1995; Tidon-Sklorz & Sene, 2001) e correspondiam, respectivamente, aos tipos morfológicos C, B e D da superespécie *D. serido* (Silva & Sene, 1991). As relações filogenéticas entre as espécies desse *cluster* vêm sendo investigadas há anos. Inicialmente, essas relações foram estabelecidas com base em

inversões cromossômicas (Ruiz & Wasserman, 1993; Tosi & Sene, 1989; Ruiz *et al.*, 2000). Posteriormente, foram realizados estudos com aloenzimas (Heed *et al.*, 1990; Lapenta *et al.*, 1998) e estudos moleculares utilizando genes nucleares (Rodrigues-Trelles *et al.*, 2000; Durando *et al.*, 2000) genes mitocondriais (Manfrin *et al.*, 2001) e DNA satélite (Kuhn, 2003) como marcadores genéticos na tentativa de esclarecer a filogenia do *cluster*. Porém, até o momento, a falta de congruência entre as árvores geradas por esses estudos impossibilita a elaboração de uma árvore de consenso que represente com segurança as relações filogenéticas entre as espécies do *cluster*.

Sabe-se que os elementos transponíveis são importantes ferramentas da Genética moderna. Recentemente, têm sido propostos como eficientes marcadores moleculares, levando em conta que possuem inúmeras cópias inativas que ocupam posições fixas no genoma e podem ser usados como os demais marcadores moleculares. Considerando essa possibilidade, embora esse não tenha sido um objetivo deste trabalho, as filogenias obtidas de cada elemento podem ser utilizadas para sugerir as relações de parentesco entre suas espécies hospedeiras. Por exemplo, as filogenias apresentadas para os três elementos indicaram a origem monofilética dessas seqüências no *cluster buzzatii*, concordante com a monofilia proposta para as espécies hospedeiras (Durando *et al.*, 2000; Ruiz *et al.*, 2000; Manfrin *et al.*, 2001; Kuhn, 2003). Esses resultados indicam que os elementos transponíveis podem contribuir para esclarecer as complexas relações filogenéticas do grupo *repleta*, apesar da ocorrência ocasional de eventos de transferência horizontal.

Interessantemente, as filogenias apresentadas para os três elementos mostram maior grau de parentesco entre as seqüências *D. serido* e *D. buzzatii*. Essa relação não é observada nas filogenias das espécies do *cluster buzzatii* (Ruiz *et al.*, 2000; Manfrin *et al.*, 2001; Kuhn, 2003). Nessas, *D. buzzatii* é sempre a espécie mais divergente e *D. serido* se agrupa com as demais espécies do *cluster* em diferentes associações, dependendo do marcador utilizado. Duas hipóteses podem explicar essa discordância. A primeira é que essas espécies realmente tenham uma maior proximidade filogenética, o que não foi detectado previamente em outros estudos e, que poderia ser comprovada pela ampliação do número de marcadores genéticos em estudos futuros. A segunda é que esses elementos tenham sido transferidos, simultaneamente ou não, por hibridização introgressiva entre essas duas espécies, que são atualmente simpátricas na Caatinga e ao longo do litoral brasileiro, ou o foram mais amplamente no passado, durante as retrações e expansões das populações de *D. serido* que supõe-se terem ocorrido no Quaternário (Brito *et al.*, 2002). A transferência de elementos

transponíveis via hibridização é reforçada por estudos que mostram variados níveis de fertilidade dos híbridos de diferentes espécies do grupo (Madi-Ravazzi *et al.*, 1997; Machado *et al.*, 1999).

Este trabalho permitiu ampliar o conhecimento sobre a distribuição e o número de inserções genômicas de três elementos transponíveis no grupo *repleta*. Foram apresentadas quatro novas evidências da ocorrência de transferência horizontal entre eucariotos e um possível modo de transmissão desses elementos entre as espécies de *Drosophila*. Até o momento, todas as dúvidas a respeito da existência do fenômeno transferência horizontal e seus possíveis efeitos na evolução dos organismos eucariotos são respondidas apenas com hipóteses, mas uma vez que um número cada vez maior de evidências corroborem esse fenômeno, estaremos mais próximos de avaliar o impacto da transferência horizontal na evolução das espécies e o papel dos TEs nessa transferência.

V. CONCLUSÕES



O presente trabalho permitiu estabelecer as seguintes conclusões:

1. O elemento transponível *micropia* é um antigo componente do genoma das espécies do grupo *repleta*, uma vez que são amplamente distribuídos e possuem baixos números de inserções. Os elementos *copia* e *Afinos*, embora apresentem baixo número de inserções, parecem ter sido introduzidos recentemente no genoma das espécies do grupo *repleta* por eventos de transferência horizontal.
2. Análises de divergência de sequências mostraram que as espécies do grupo *repleta* abrigam duas subfamílias *micropia*, uma subfamília *copia* e duas subfamílias *Afinos* em seus genomas.
3. Análises do polimorfismo de sequências da região da LTR-ULR do elemento *copia* mostraram menor variação nucleotídica na ULR região do que nas demais regiões da LTR, sugerindo que forças seletivas atuaram ao longo da história evolutiva conservando esta região que é essencial à atividade do elemento e que as sequências *copia* presentes nos genomas das espécies do grupo *repleta* são mais antigas do que aquelas encontradas nos genomas das espécies do grupo *melanogaster*.
4. Os elementos transponíveis *micropia*, *copia* e *Afinos* corroboram a hipótese da origem monofilética de *cluster buzzatii*.
5. Análises de divergência de nucleotídeos de TEs, de associações históricas entre TEs e hospedeiros e das taxas de substituições em sítios sinônimos dos TEs e do gene *Adh* permitiram sugerir vários eventos de transferência horizontal dos elementos *copia* e *Afinos* entre as espécies do grupo *repleta* e *saltans*. A ocorrência de um evento de transferência horizontal dentro do *cluster buzzatii* (dos elementos *copia* e *Afinos*, entre *D. serido* e *D. buzzatii*), de eventos dentro do subgrupo *mülleri* (do elemento *Afinos* entre *D. serido* e *D. mulleri* e uma linhagem ancestral das espécies dos *clusters mulleri* e *melanogaster*), bem como a de eventos entre espécies que pertencem a subgêneros diferentes (do elemento *Afinos* entre *D. spenceri*, grupo *repleta*, subgênero *Drosophila* e *D. emarginata*, grupo *saltans*, subgênero *Sophophora*; e entre *D. saltans*, grupo *saltans*, subgênero *Sophophora* e uma linhagem ancestral das espécies dos *clusters mulleri* e *melanogaster*, grupo *repleta*, subgênero *Drosophila*) sugerem que o fenômeno de transferência horizontal é relativamente frequente nesses grupos de espécies.

V. CONCLUSÕES

O presente trabalho permitiu estabelecer as seguintes conclusões:

1. O elemento transponível *micropia* é um antigo componente do genoma das espécies do grupo *repleta*, uma vez que são amplamente distribuídos e possuem baixos números de inserções. Os elementos *copia* e *Minos*, embora apresentem baixo número de inserções, parecem ter sido introduzidos recentemente no genoma das espécies do grupo *repleta* por eventos de transferência horizontal
2. Análises de divergência de seqüências mostraram que as espécies do grupo *repleta* abrigam duas subfamílias *micropia*, uma subfamília *copia* e duas subfamílias *Minos* em seus genomas;
3. Análises do polimorfismo de seqüências da região da LTR-ULR do elemento *copia* mostraram menor variação nucleotídica na ULR região do que nas demais regiões da LTR, sugerindo que forças seletivas atuaram ao longo da história evolutiva conservando essa região que é essencial à atividade do elemento e que as seqüências *copia* presentes nos genomas das espécies do grupo *repleta* são mais antigas do que aquelas encontradas nos genomas das espécies do grupo *melenogaster*;
4. Os elementos transponíveis *micropia*, *copia* e *Minos* corroboram a hipótese da origem monofilética do *cluster buzzatti*;
5. Análises de divergência de nucleotídeos de TEs, de associações históricas entre TEs e hospedeiros e das taxas de substituições em sítios sinônimos dos TEs e do gene *Adh* permitiram sugerir vários eventos de transferência horizontal dos elementos *copia* e *Minos* entre as espécies do grupo *repleta* e *saltans*. A ocorrência de um evento de transferência horizontal dentro do *cluster buzzatii* (dos elementos *copia* e *Minos*, entre *D. serido* e *D. buzzatii*), de eventos dentro do subgrupo *mulleri* (do elemento *Minos* entre *D. mojavenensis* e *D. mulleri* e entre *D. hydei* e uma linhagem ancestral das espécies dos clusters *mulleri* e *mojavensis*), bem como e de eventos entre espécies que pertencem a subgêneros diferentes (do elemento *Minos* entre *D. spenceri*, grupo *repleta*, subgênero *Drosophila* e *D. emarginata*, grupo *saltans*, subgênero *Sophophora*; e entre *D. saltans*, grupo *saltans*, subgênero *Sophophora* e uma linhagem ancestral das espécies dos clusters *mulleri* e *mojavensis*, grupo *repleta*, subgênero *Drosophila*) sugerem que o fenômeno de transferência horizontal é relativamente freqüente nesses grupos de espécies.



- ADAMS, M. D., *et al.* The genome sequence of *Drosophila melanogaster*. *Science*, v. 287, p. 2185-2195, 2000.
- AGRAWAL, A.; EASTMAN, Q.; SCHATZ, D. Transposition mediated by RAG1 and RAG2 and implications for the evolution of immune system. *Nature*, v. 394, p. 744-751, 1998.
- ALBEROLA, T. M.; BORI, L.; DE FRUTOS, R. Structural analysis of *Drosophila subobscura* gypsy elements (gypsy Ds). *Genetica*, v. 100, p.39-48, 1997.
- ALMEIDA, L. M.; CASTRO, J. P.; CARARETO, C. M. A. *Microplia* transposable element presence in *Drosophila saltans* group. *Dev. Inf. Serw.*, v. 84, p. 114-117, 2001.
- ARCA, B.; SAVAKIS, C. Distribution of transposable element *African* in the genus *Drosophila*. *Genetics*, v. 108, p.263-267, 2000.
- ARNAULT, C.; DUFOURNEL, J. Genome and stresses' reactions against Aggregations' behavior of transposable elements. *Genetica*, v. 93, p. 149-160, 1994.
- BADGE, R.; BROOKFIELD, J. F. The role of host factors in population dynamics of selfish transposable elements. *J. Theor. Biol.*, v. 187, p. 261-271, 1997.
- BEALL, E. L.; RIO, D. C. *Drosophila* TRIP/Ku P70 corresponds to the mutagen-sensitive *mus309* gene and is involved in *P* element excision *in vivo*. *Genes Dev.*, v. 10, p. 921-933, 1996.
- BEST, S.; TISSIER, P. L. Positional cloning of mouse retrovirus restriction gene *Fel*. *Nature*, v. 382, p. 826-829, 1996.

VI. REFERÊNCIAS

- BIÉMONT, C. Transposable element copy number in natural populations of *Drosophila melanogaster* and *D. simulans*. *Genetica*, v. 100, p. 161-166, 1997.
- BIÉMONT, C.; CIZERON, G. Distribution of transposable elements in *Drosophila* species. *Genetica*, v. 103, p. 43-62, 1999.
- BINGHAM, P.; KIDWELL, M. G.; RUBIN, G. M. The molecular basis of P-M hybrid dysgenesis: the role of *P* element, a *P* strain transposons family. *Cell*, v. 29, p. 995-1004, 1982.
- BUCHETON, A., *et al.* The molecular basis of I-R hybrid dysgenesis: identification, cloning and properties of *I* factor. *Cell*, v. 38, p. 153-163, 1984.
- CLARK, J. B.; KIDWELL, M. G. A phylogenetic perspective on *p* transposable element evolution in *Drosophila*. *Proc Natl. Acad. Sci. USA.*, v. 94, p. 11428-11433, 1997.

- ADAMS, M. D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science*, v. 287, p. 2185-2195, 2000.
- AGRAWAL, A.; EASTMAN, Q.; SCHATZ, D. Transposition mediated by RAG1 and RAG2 and implications for the evolution of immune system. *Nature*, v. 394, p. 744-751, 1998.
- ALBEROLA, T. M.; BORI, L.; DE FRUTOS, R. Structural analysis of *Drosophila subobscura* gypsy elements (gypsy Ds). *Genetica*, v. 100, p. 39-48, 1997.
- ALMEIDA, L. M.; CASTRO, J. P.; CARARETO, C. M. A. *Micropia* transposable element presence in *Drosophila saltans* group. *Dros. Inf. Serv.*, v. 84, p. 114-117, 2001.
- ARCÀ, B.; SAVAKIS, C. Distribution of transposable element *Minos* in the genus *Drosophila*. *Genetics*, v. 108, p. 263-267, 2000.
- ARNAULT, C.; DUFOURNEL, I. Genome and stresses: reactions against aggressions, behavior of transposable elements. *Genetica*, v. 93, p. 149-160, 1994.
- BADGE, R.; BROOKFIELD, J. F. The role of host factors in population dynamics of selfish transposable elements. *J. Theor. Biol.*, v. 187, p. 261-271, 1997.
- BEALL, E. L.; RIO, D. C. *Drosophila* IRBP/Ku P70 corresponds to the mutagen-sensitive *mus309* gene and is involved in *P* element excision *in vivo*. *Genes Dev.*, v. 10, p. 921-933, 1996.
- BEST, S.; TISSIER, P. L. Positional cloning of mouse retrovirus restriction gene *Fv1*. *Nature*, v. 382, p. 826-829, 1996.
- BIÉMONT, C., *et al.* Maintenance of transposable element copy number in natural populations of *Drosophila melanogaster* and *D. simulans*. *Genetica*, v. 100, p. 161-166, 1997.
- BIÉMONT, C.; CIZERON, G. Distribution of transposable elements in *Drosophila* species. *Genetica*, v. 105, p. 43-62, 1999.
- BINGHAM, P.; KIDWELL, M. G.; RUBIN, G. M. The molecular basis of P-M hybrid dysgenesis: the role of *P* element, a *P* strain transposons family. *Cell*, v. 29, p. 995-1004, 1982.
- BUCHETON, A., *et al.* The molecular basis of I-R hybrid dysgenesis: identification, cloning and properties of *I* factor. *Cell*, v. 38, p. 153-163, 1984.
- CLARK, J. B.; KIDWELL, M. G. A phylogenetic perspective on *p* transposable element evolution in *Drosophila*. *Proc Natl. Acad. Sci. USA.*, v. 94, p. 11428-11433, 1997.



- CAPY, P., *et al.* *Dynamics and Evolution of Transposable Elements*. 1st. France: Landes Bioscience, 1998. 197p.
- CELNIKER, S. E.; RUBIN, G. M. The *Drosophila melanogaster* genome. *Annual Reviews of Genomics and Human Genetics*, v. 4, p. 89-117, 2003.
- CHARLESWORTH, B. The population genetics of transposable elements. *In: T. Ohta and K. Aoki (eds), Population Genetics and Molecular Evolution*, Springer, Berlin, 1985.
- CHARLESWORTH, B.; LANGLEY, C. H. The evolution of self-regulated transposition of transposable elements. *Ann. Ver. Genet.*, v. 23, 251-287, 1986.
- CSINKI, A. K.; MCDONALD, J. F. Analysis of *copia* sequence variation within and between *Drosophila* species. *Mol. Biol. Evol.*, v. 12, p. 83-93, 1995.
- DANIELS, S. B. *et al.* Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species. *Genetics*, v. 124, p. 339-355, 1990.
- DEININGER, P. L.; BATZER, M. *Alu* repeats and human disease. *Mol. Gen. Metab.*, v. 67, p. 183-193, 1999.
- DOOLITTLE, W. F.; SAPIENZA, C. Selfish genes, the phenotypic paradigm and genomic evolution. *Nature*, v. 284, p. 601-603, 1980.
- DRABEK, D. *et al.* Transposition of *Drosophila hydei* Minos transposon in mouse germ-line. *Genomics*, v. 81, p. 108-111, 2003.
- DURANDO, C., *et al.* Phylogenetic analysis of the *repleta* species group of the genus *Drosophila* using multiple sources of characters. *Mol. Phylogenet. Evol.*, v. 2, p. 296-307, 2000.
- ENGELS, W. R. Hybrid dysgenesis in *Drosophila melanogaster*: rules of inheritance of female sterility. *Genetics*, v. 33, p. 219-236, 1979.
- ENGELS, W. R.; PRESTON, C. R. Hybrid dysgenesis in *Drosophila melanogaster*. The biology of female and male sterility. *Genetics*, v. 92, p. 161-174, 1979.
- EMORI, Y., *et al.* Determination of the nucleotide sequences of *copia* and *copia*-related RNA in *Drosophila* VLP. *Nature*, v. 315, p. 773-776, 1985.
- FLAVELL, A. J. Long terminal repeat retrotransposons jump between species. *Proc. Natl. Acad. Sci. USA*, v. 96, p. 12211-12212, 1999.

- FINNEGAN, D. J. Transposable elements in eukaryotes. *Int. Rev. Cytol.*, v. 93, p. 281-326, 1985.
- FRANZ, G.; SAVAKIS, C. *Minos*, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposon. *Nucl. Acids Res.*, v. 19, p. 6646, 1991.
- GRANDBASTIEN, M. A., *et al.* The expression of *tobacco Tnt1* retrotransposon linked to plant defense response. *Genetica*, v. 100, p. 241-52, 1997.
- GRANDBASTIEN, M. A. Activation of plant retrotransposons under stress conditions. *Trends Plant Sci.*, v. 3, p. 181-187, 1998.
- GRAUR, D.; LI, W. H. *Fundamentals of Molecular Evolution*, second edition, Sinauer Associates, U.S.A., 2000.
- GREEN, M. M. Genetic instability in *Drosophila melanogaster*: de novo induction of putative insertion mutations. *Proc. Natl. Acad. Sci. USA*, v. 74, p. 3490-3493, 1977.
- GREEN, M. M. Genetic instability in *Drosophila melanogaster*: on the identity of the MR and P-M system. *Biol. Zentralbl.* v. 103, p. 1-8, 1984.
- GREEN, M. M. Genetic instability in *Drosophila melanogaster*: the genetics of na MR element that makes complete *P* insertion mutations. *Proc. Natl. Acad. Sci. USA*, v. 83, p. 4533-4537, 1986.
- GREIDER, C. W. Telomere length regulation. *Annu. Rev. Biochem.*, v. 65, p. 337-365, 1996.
- HARTL, D. L.; LOZOVSKAYA, E.; LAWRENCE, J. G. Nonautonomous transposable elements in prokaryotes and eukaryotes. *Genetica*, v. 86, p. 47-53, 1992.
- HEED, W. B., *et al.* Genetic differentiation among island populations and species of cactophilic *Drosophila* in West Indies. In *Ecological and Evolutionary Genetics of Drosophila* (J.S.F. Barker, W.T. Starmer, R.J. MacIntyre, Eds), p. 447-490. Plenum Press, New York, 1990.
- HIOM, K.; MELEK, M.; GELLERT, M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell*, v. 94, p. 463-470, 1998.
- HURST, G. D.; WERREN, J. H. The role of selfish genetic elements eukaryotic evolution. *Nature*, v. 2, p. 597-606, 2001.
- JORDAN, I. K.; McDONALD, J. F. Evolution of *copia* retrotransposon in *Drosophila melanogaster* species subgroup. *Mol. Biol. Evol.*, v. 15, p. 1160-1171, 1998a.

JORDAN, I. K.; McDONALD, J. F. Interelement selection in the regulatory region of the *copia* retrotransposon. *J. Mol. Evol.*, v. 47, p. 670-676, 1998b.

JORDAN, I. K.; MATYUNINA, V. L.; MCDONALD, J. F. Evidence for recent horizontal transfer of long terminal repeat retrotransposon. *Proc. Natl. Acad. Sci. USA* v. 96, p. 12621-12625, 1999.

KAPLAN, N. L.; DARDEN, T.; LANGLEY, C. Evolution and extinction of transposable elements in Mendelian populations. *Genetics*, v. 109, p. 459-480, 1985.

KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: nature and inheritance of P element regulation. *Genetics*, v. 111, p. 337-350, 1985.

KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: partial sterility associated with embryonic lethality in P-M system. *Genet. Res.*, v. 44, p. 11-28, 1984.

KIDWELL, M. G. Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA*, v. 80, p. 1655-1659, 1983.

KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: the genetics of cytotype determination in a neutral strain. *Genetics*, v. 98, p. 275-290, 1981.

KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: the relationship between the P-M and I-R interaction system. *Genet. Res.*, v.33, p. 105-117, 1979.

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

KIDWELL, M. G.; KIDWELL, J. F.; SVED, J. A. Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics*, v. 86, p. 813-833, 1977.

KIDWELL, M. G.; LISCH, D. R. Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. USA.*, v. 94, p. 7704-7711, 1997.

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

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

- KLINAKIS, A. G. *et al.* Genome-wide insertional mutagenesis in human cells by *Drosophila* mobile element *Minos*. *EMBO*, v. 1, p. 416-421, 2000.
- KUHN, G. C. S. Dinâmica evolutiva de DNAs satellites em populações e espécies de *Drosophila* do cluster *buzzatii*. Ribeirão Preto, 2003. Tese de Doutorado. Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto.
- LABRADOR, M.; FONTDEVILA, A. High transposition rates of *Oswaldo*, a new *Drosophila buzzatii* retrotransposon. *Mol. Gen. Genet.*, v. 245, p. 661-674, 1994.
- LABRADOR, M., *et al.* Interspecific hybridization increases transposition rates of *Oswaldo*. *Mol. Biol. Evol.*, v. 16, p. 931-937, 1999.
- LANKENAU, D-H; HUIJER, P.; HENNIG, W. Characterization of the long terminal repeats of *micropia* elements dissected from the Y chromosomal lamprush loops "Threads" of *Drosophila hydei*. *J. Mol. Biol.*, v. 209, p. 493-497, 1989.
- LANKENAU, D-H. The retrotransposon family *micropia* in *Drosophila* species, pp. 232-241 in J. McDonald (ed.), *Transposable elements and evolution*. Kluwer Publishers, Amsterdam, 1993.
- LANKENAU, S; CORCES, G. V.; LANKENAU, D-H. The *Drosophila micropia* retrotransposon encodes a testis-specific antisense RNA complementary to reverse transcriptase. *Mol. Cell Biol.*, v. 14, p. 1764-1775, 1994.
- LAPENTA, A. S.; BICUDO, H. E. M.; CERON, C. R.; CORDEIRO, J. Esterase patterns and phylogenetic relationships of species and strains included in *Drosophila buzzatii* cluster. *Cytobios*, v. 96, p. 95-107, 1998.
- LEVITAN, M. Spontaneous chromosome aberrations in *Drosophila robusta*. *Proc. Natl. Acad. Sci. USA.*, v. 48, p. 930-937, 1962.
- LOHE, A. R.; MORIYAMA, E. N.; LIDHOLM, D.; HARTL, D. Horizontal transmission, vertical inactivation, and stochastic loss of *mariner-like* transposable elements. *Mol. Biol. Evol.*, v. 12, p. 62-72, 1995.
- MACHADO, L. P. B.; MADI-RAVAZZI, L.; TADEI, W. J. Reproductive and chromosome studies in *Drosophila* species of *buzzatii* cluster. *Genet. Mol. Biol.*, v. 22, p. 386, 1999.
- MADI-RAVAZZI, L.; BICUDO, H. E. M.; MANZATO, J. A. Reproductive compatibility and chromosome pairing in the *Drosophila buzzatii* complex. *Cytobios*, v. 89, p. 21-30, 1997.
- MANFRIN, M. H.; DE BRITO, R. O. A.; SENE, F. M. Systematics and evolution of the *Drosophila buzzatii* (Diptera: Drosophilidae) cluster using mtDNA. *Ann. Entomol. Soc. Am.*, v. 94, p. 333-346, 2001.



MARTIN, G.; WIERNASZ, D.; SCHEDL, P. Evolution of *Drosophila* repetitive-dispersed DNA. *J. Mol. Evol.*, v. 19, p. 203-213, 1983.

MILLER, W. J. *et al.* Structure and expression of clustered *P* element homologous in *Drosophila subobscura* and *Drosophila guanche*. *Gene*, v. 156, p. 167-174, 1995.

MONTCHAMP-MOREAU, C. S., *et al.* D. Distribution and conservation of sequences homologous to the 1731 retrotransposon in *Drosophila*. *Mol. Biol. Evol.*, v.10, p. 791-803, 1993.

MORAN, J. V.; DEBERARDINIS, R. J.; KAZAZIAN Jr, H. H. Exon shuffling by L1 retrotransposition. *Science*, v. 283, p. 1530-1535, 1999.

MOUNT, S.; RUBIN, G. M. Complete nucleotide sequence of the *Drosophila* transposable element *copia*: homology between *copia* and retroviral proteins. *Mol. Cell. Biol.*, v. 5, p. 1630-1638, 1985.

NAKAMURA, T. M.; CECH, T. R. Revising time: origin of telomerases. *Cell*, v. 92, p. 587-590, 1998.

NIKITIN, A. G.; SHMOOKLER REIS, R. J. Role of transposable elements in age-related genomic instability. *Genet. Res.*, v. 69, p. 183-195, 1997.

NOUAUD, D.; ANXOLABÉHÈRE, A. *P* element domestication: a stationary truncated *P* element may encode a 66-Kda repressor-like protein in the *Drosophila montium* species subgroup. *Mol. Biol. Evol.*, v. 14, p. 1132-1144, 1997.

NOUAUD, D.; QUESNEVILLE, H.; ANXOLABÉHÈRE, D. Recurrent exon shuffling between distant *P*-element families. *Mol. Biol. Evol.*, v. 20, p. 190-199, 2003.

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

O'HARE, K.; RUBIN, G.M. Structure of *P* element transposable and their sites of insertion and excision in *Drosophila melanogaster* genome. *Cell*, v. 34, p. 25-35, 1983.

PARDUE, M. L., *et al.* *Drosophila* telomeres: new views on chromosome evolution. *Trends Genet.*, v. 12, p. 48-52, 1996.

PELISSON, A., *et al.* Gypsy transposition correlates with the production of retroviral envelope-like protein under the tissue-specific control of the *Drosophila* flamenco gene. *EMBO J.*, v. 13, p. 4401-4411, 1994.



PINSKER, W., *et al.* The evolutionary life history of *P* transposons: from horizontal invaders to domesticated neogenes. *Chromosoma*, v. 110, p. 148-158, 2001.

PEREIRA, M. A.; VILELA, C. R.; SENE, F. M. Notes on breeding and feeding sites in some species of *repleta* group of genus *Drosophila* (Diptera, Drosophilidae). *Cien. e Cult.*, v. 35, p. 1313-1319, 1983.

RICHARDSON, R. H., RICHARDSON, M. E.; SMOUSE, P. E. Evolution of electrophoretic mobility in *Drosophila mulleri* complex. In: *Isoenzymes IV, Genetics and Evolution* (C.L. Markert, Ed) p. 533-545. *Academic Press*, New York, 1975.

RICHARDSON, R. H., SMOUSE, P. E. Patterns of molecular variation. I. Interspecific comparison of electromorphs in *Drosophila mulleri* complex. *Biochem. Genet.*, v. 14, p. 447-465, 1976.

RICHARDSON, R. H.; SMOUSE, P. E.; RICHARDSON, M. E. Patterns of molecular variation. II. Associations of electrophoretic mobility and larval substrate within species of *Drosophila mulleri* complex. *Genetics*, v. 85, p. 141-154, 1977.

RODRÍGUEZ-TRELLES, F.; ALARCÓN, L.; FONTDEVILA, A. Molecular evolution and phylogeny of the *buzzatii* complex (*Drosophila repleta* group): a maximum-likelihood approach. *Mol. Biol. Evol.*, v. 17, p. 1112-1122, 2000.

RUIZ, A.; FONTDEVILA, A. Ecología y evolución del grupo *mulleri* de *Drosophila* en Venezuela y Colombia. *Acta Cient. Venezolana*, v.32, p. 338-345, 1981.

RUIZ, A.; WASSERMAN, M. Evolutionary cytogenetics of the *Drosophila buzzatii* species complex. *Heredity*, v. 70, p. 597-603, 1993.

RUIZ, A. *et al.* The *Drosophila serido* speciation puzzle: putting new pieces together. *Genetica*, v. 108, p. 217-227, 2000.

RUSSO, C. A.; TAKEZAKI, N.; NEI, M. Molecular phylogeny and divergence times of drosophilid species. *Mol. Biol. Evol.*, v. 12, p. 391-404, 1995.

SENE, F. M.; PEREIRA, M. A. Q.; VILELA, C. R. Evolutionary aspects of cactus breeding *Drosophila* species in South America. In: J.S.F. Barker and W.T. Starmer (eds). *Ecological Genetics and Evolution: the cactus-yeast Drosophila model system*. Academic Press, Sidney, Australia, p. 97-106, 1982.

SILVA, J. C.; LORETO, E.; CLARK, J. B. Factors that affect the horizontal transfer of transposable elements. *Curr. Issues Mol. Biol.*, v. 6, p. 57-72, 2004.

SILVA, A. F. G.; SENE, F. M. Morphological geographic variability in *Drosophila serido* (Diptera, Drosophilidae). *Revta. Bras. Ent.*, v. 35, p. 455-468, 1991.

- SCHAEFER, R. E.; FAUSTO-STERLING, A.; KIDWELL, M. G. Hybrid dysgenesis in *Drosophila melanogaster*: morphological and cytological studies of ovarian dysgenesis. *Genetics*, v. 92, p. 1141-1152, 1979.
- SPICER, G. S. Phylogenetic utility of mitochondrial cytochrome oxidase gene: Molecular evolution of *Drosophila buzzatii* species complex. *J. Mol. Evol.*, v. 41, p. 749-759, 1995.
- SPICER, G. S.; PITNICK, S. Molecular systematics of *Drosophila hydei* subgroup as inferred from mitochondrial DNA sequences. *J. Mol. Evol.*, v. 43, p. 281-286, 1996.
- SULLIVAN, D. T.; ATKINSON, P. W.; BAYER, C. A.; MENOTTI-RAYMOND, M. The evolution of Adh expression in *repleta* group of *Drosophila*. In: *Ecological and Evolutionary Genetics of Drosophila* (J. S. F. Barker, W. T. Starmer, and R. J. MacIntyre, Eds), p. 407-418. Plenum Press, New York, 2000.
- TATARENKOV, A.; AYALA, F. Phylogenetic relationships among species groups of the *virilis-repleta* radiation of *Drosophila*. *Mol. Phylogen. Evol.*, v. 21, p. 327-331, 2001.
- THROCKMORTON, L. H. The phylogeny, ecology and geography of *Drosophila*. In: King, R.C., *Handbook of Genetics*, v. 3, New York: Plenum Press, p. 421-469, 1975.
- THROCKMORTON, L. H. Pathways of evolution in the genus *Drosophila* and the founding of *repleta* group. In: *Ecological Genetics and Evolution: The cactus-yeast- Drosophila Model System* (J. S. F. Barker and W. T. Starmer, Eds), p. 33-48. Academic press, New York, 1982.
- TIDON-SKLORZ, R; SENE, F. M. Two new species of the *Drosophila serido* sibling set (Diptera, Drosophilidae). *Inheringia Sér Zool.*, v. 90, p. 141-146, 2001.
- TIDON-SKLORZ, R., *et al.* The genus *Drosophila* (Diptera, Drosophilidae) in the Serra do Cipó, state of Minas Gerais, Brazil, *Revta. Bras. Ent.*, v. 38(3/4), p. 627-637, 1994.
- TIDON-SKLORZ, R; SENE, F. M. *Drosophila seriema* n. sp: new member of *Drosophila serido* (Diptera, Drosophilidae) superspecie taxon. *Ann. Ent. Soc. Am.*, v. 88, p. 139-142, 1995.
- TOSI, D.; SENE, F. M. Further studies on chromosomal variability in *Drosophila serido* (Diptera, Drosophilidae) *Revta. Bras. Genet.*, v. 12, p. 729-746, 1989.
- VIEIRA, C., *et al.* Evolution of genome size in *Drosophila*. Is the invader's genome being invaded by transposable elements? *Mol. Biol. Evol.*, v. 19, p. 1154-1161, 2002.

VILELA, C. R. A revision of *Drosophila repleta* species group (Diptera, Drosophilidae) *Rev. Bras. Ent.*, p. 1-114, 1983.

VILELA, C. R.; SENE, F. M. Two new neotropical species group (Diptera, Drosophilidae). *Rev. Bras. Ent.*, v. 30, p. 295-299, 1977.

YANNOPOULOS, G., *et al.* *Hobo* is responsible for induction of hybrid dysgenesis by strains of *Drosophila melanogaster* bearing the male recombination factor 23,5 MRF. *Cell*, v. 48, p. 487-495, 1987.

WASSERMAN, M. Cytological evolution in *Drosophila repleta* species group. In *Ecological Genetics and Evolution: The cactus-yeast- Drosophila Model System* (J. S. F. Barker and W. T. Starmer, Eds), p. 49-64. *Academic press, New York*, 1982.

WASSERMAN, M. Cytological evolution in *Drosophila repleta* species group. In: *Drosophila inversion polymorphism* (C.B Krimbas and J. R. Powell, Eds), p. 455-552 CRC Press, Boca Raton, F.L, 1992.

WRIGHT, S.; FINNEGAN, D. Genome evolution: sex and the transposable element. *Curr. Biol.*, v. 11, R296-R299, 2001.

VANHEE-BROSSOLLET, C.; VAQUERO, C. Do natural antisense transcripts make sense in eukaryotes? *Gene*, v. 211, p. 1-9, 1997.

ZOUROS, E. Genetic differentiation associated with the early stages of speciation in the *mulleri* subgroup of *Drosophila*. *Evolution*, v. 27, p. 601-621, 1973.

ALMEIDA, L.M. Distribuição genômica e relações filogenéticas dos elementos transponíveis *copia*, *micropia* e *Minos* em espécies do grupo *repleta* de *Drosophila*. São José do Rio Preto, 2004. 134p. Tese (Doutorado em Genética) – Instituto de Biociências, Letras, e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”.

No presente estudo foram investigados a distribuição, o número de inserções genômicas, a divergência de seqüências de DNA, a ocorrência de transferência horizontal e as relações filogenéticas dos elementos transponíveis *copia*, *micropia* e *Minos* em espécies do grupo *repleta* de *Drosophila*. Para isso, foram empregadas as técnicas PCR, *Southern blot* e seqüenciamento. Os resultados mostraram que o elemento transponível *micropia* é um antigo componente do genoma das espécies do grupo *repleta*, uma vez que esse é amplamente distribuído e possui baixos números de inserções genômicas. Os elementos *copia* e *Minos*, embora apresentem baixo número de inserções, parecem ter sido introduzidos recentemente no genoma dessas espécies. As análises de divergência de seqüências mostraram que as espécies do grupo *repleta* abrigam duas subfamílias *micropia*, duas subfamílias *Minos* e uma subfamília *copia*, como também que forças seletivas atuam na conservação da região ULR do elemento *copia*. Análises de associações históricas do elemento *copia* com as espécies do grupo *repleta* sugeriram a ocorrência de um evento de transferência horizontal entre *D. buzzatii* e *D. serido* para explicar a história evolutiva desse elemento no cluster *buzzatii*. Por outro lado, análises de divergência de nucleotídeos, de associações históricas e das taxas de substituições em sítios sinônimos dos TEs e do gene *Adh* das espécies hospedeiras permitiram sugerir que o elemento *Minos* foi transferido horizontalmente ao longo do tempo evolutivo. Foram identificados cinco possíveis eventos de transferência: entre *D. saltans* e uma linhagem ancestral dos clusters *mulleri* e *mojavensis* e essa linhagem e *D. hydei*; entre *D. mulleri* e *D. mojavensis*, entre *D. serido* e *D. buzzatii* e, ainda entre *D. emarginata* e *D. spenceri*. Uma hipótese alternativa seria a ocorrência de repetidos eventos de transferência horizontal envolvendo a espécie *D. hydei*, *D. aldrichi*, *D. saltans* e a linhagem ancestral dos clusters *mulleri* e *mojavensis*. Essa proposta está de acordo com os tempos de divergência das espécies hospedeiras (14, 6 e de 4 milhões de anos, respectivamente para as espécies *D. hydei*, espécies dos clusters *mulleri* e *mojavensis* e *D. saltans*). Esses resultados mostram que o fenômeno de transferência horizontal é relativamente freqüente no grupo *repleta* de *Drosophila*.

Palavras-chave: elementos transponíveis, transferência horizontal, retrotransposon *copia*, retrotransposon *micropia*, transposon *Minos*, grupo *repleta*.



ALMEIDA, L.M. Genomic distribution and phylogenetic relationships of transposable elements *copia*, *micropia* and *Minos* in *Drosophila* species of the *repleta* group. São José do Rio Preto, 2004. 134 p. Tese (Doutorado em Genética) – Instituto de Biociências, Letras, e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”.

The present study aimed to investigate the distribution, the number of insertions in the genome, the divergence of DNA sequences, the occurrence of horizontal transfer and the phylogenetic relationships of *copia*, *micropia* and *Minos* transposable elements in *Drosophila* species of the *repleta* group. To do so, we employed PCR, *Southern blot*, RT-PCR and sequencing analyses. The results suggested that the *micropia* element is an old component of the *repleta* group, since it is widely distributed among the species and has low number of insertions in the host genomes. Although, *Minos* and *copia* showed low number of insertions in the host genomes it is suggested that they were recently introduced in the *repleta* group. The analyses showed the existence of two subfamilies of *micropia*, two subfamilies of *Minos* and just one subfamily of *copia* element among the *repleta* species. The low diversity of nucleotides in the ULR region of the retrotransposon *copia* suggested that selective constraints maintain this regulatory region over evolutionary time. In addition, historical associations suggested the occurrence of one event of horizontal transfer between *D. buzzatii* and *D. serido* to explain the *copia* evolution in the *buzzatii* cluster. On the other hand, divergence of nucleotides, historical associations and comparison of substitution rates between *Minos* and the gene host *Adh* suggested that distribution of *Minos* transposon among the *saltans* and the *repleta* group is determined by events of horizontal transmission balanced by vertical inactivation and extinction. Five possibilities of horizontal transfers were detected in the present study: between *D. saltans* and the ancestor of the *mulleri* and the *mojavensis* clusters and this ancestor lineage and *D. hydei*; between *D. mulleri* and *D. mojavensis*; between *D. buzzatii* and *D. serido* and between *D. spenceri* and *D. emarginata*. An alternative hypothesis would be three events of horizontal transfer involving *D. hydei*, *D. aldrichi*, *D. saltans* and the lineage ancestor of the *mulleri* and the *mojavensis* clusters. This hypothesis is in agreement with the time of divergence of the host species (14, 6 and 4 Mya, respectively for *D. hydei*, the *mulleri* and the *mojavensis* clusters and *D. saltans*). These results show that horizontal transfer is a phenomenon relatively frequent in the *repleta* group of species.

Keywords: transposable elements, horizontal transfer, *copia* retrotransposon, *micropia* retrotransposon, *Minos* transposon, *Drosophila repleta* group.



