Universidade Federal do Rio Grande do Sul – UFRGS

Programa de Pós-Graduação em Genética e Biologia Molecular – PPGBM

Elemento de transposição *micropia* e evolução cromossômica do subgrupo *cardini*, grupo c*ardini* do gênero *Drosophila* (Diptera: Drosophilidae)

Juliana Cordeiro

Tese de Doutorado submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau de Doutor em Ciências (Genética e Biologia Molecular).

Orientadora: Dra. Vera Lúcia da Silva Valente

Porto Alegre, agosto de 2009.

Este trabalho foi desenvolvido no Laboratório de *Drosophila* do Departamento de Genética da Universidade Federal do Rio Grande do Sul (UFRGS) sendo financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do Rio Grande do Sul (FAPERGS) e pela Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES).

“Se vi mais longe, foi por estar de pé sobre ombros de gigantes”

Isaac Newton

"Aprender é a única coisa de que a mente nunca se cansa, nunca tem medo e nunca se arrepende."

Leonardo da Vinci

Dedico este trabalho à minha família *et al*., pais, irmãos, tias, e Dani (∞); por tudo o que vivi, senti, aprendi e ainda aprenderei.

AGRADECIMENTOS

Por meio dessas poucas palavras eu quero exprimir todo o meu agradecimento às pessoas que me ajudaram, colaboraram, participaram de um período importante da minha vida, e de alguma forma fizeram a diferença.

Vera, eu quero te agradecer pela orientação profissional, pelo entusiasmo constante, e por tudo o que me ensinaste sendo a pesquisadora de enorme talento e ao mesmo tempo tão humana e compreensiva. O segredo do sucesso do laboratório é o clima de amizade a convivência harmônica que você oferece. Vera Valente Imortal, obrigada por tudo!

Elgion, Cláudia e Rosane, vocês foram importantes em etapas diferentes, atuando de formas diferentes e inspiradoras, essenciais à minha formação.

Elmo e Helen, obrigada por serem sempre tão prestativos.

Paulo e Dani, eu agradeço por toda a convivência que sempre tivemos que moldou o meu caráter e o meu profissionalismo, com e sem espátula multiuso. A amizade de vocês é preciosa.

Gisele, nem sei como agradecer todo o teu esforço e a tua amizade, fins de semana no lab com chocolate e conversa sobre nossos gatos. Você fez parte muito importante na minha vida, pessoal e profissional. O teu entusiasmo me contagia te adoro! Obrigada por tudo!

Carol, CReverton, Hermes (e Hércules), Ju WG, Mary, Mário, vocês são pessoas que quero sempre estar perto, para conversar, rir, aprender, ensinar, comer a pizza de Viamão, ou só para poder pensar melhor na hora de fazer os cálculos da concentração da solução. Amigos são pra essas coisas...

Ricardo da Vegetal, como eu poderia me esquecer de agradecer a você. Sou eternamente grata, sem você metade dessa Tese não sairia.

Dirleane, Gilberto, Ferreira e Gilson, na nossa convivência eu aprendi muito, me diverti muito com o hino de Ajuricaba, com dois presentes em um só amigo secreto, com os ensinamentos sobre a arte do bonsai. Obrigada pelas conversas e pelos momentos de descontração.

Mary, Hermes e Ferreira, obrigada por toda a ajuda cuidando das *cardini*s, obrigada pelas ajudas técnicas.

Hermes, Jonas, Marcão e Sá, obrigada desde o Labster, por toda a convivência e amizade, por tornar "os laboratórios" um ambiente terapêutico.

Bibi, você é a salvadora do nosso laboratório. Obrigada por tudo.

Marcão e Monica, Sá e Jonas vocês nunca negaram ajuda. Sou grata pelo primeiro dia em que fiquei na casa de vocês, e pelos anos que se seguiram, de amizade, ajuda, jogatinas, pizzas e churrascos (acreditem, eu adorava os churrascos!).

Sá, obrigada por toda a amizade, por todos os conselhos, hoje e sempre.

Adriana, Maríndia e Nina durante a nossa convivência aprendi muito com vocês. Obrigada.

Liz, obrigada por todo o ensinamento, por toda a tua paciência, por toda a tua amizade. Obrigada pelo companheirismo (em qualquer situação). Você é uma pessoa admirável.

Adri, Fabiano e Ana, obrigada pelos agradáveis momentos durante a minha adaptação ao laboratório. Obrigada pelas conversas, pelas risadas, pelos aprendizados.

Leti e Nai, vocês me aturaram nos piores momentos. Amo vocês. Irmãs de alma.

Naomi, eu aprendi a viver com você. Obrigada pelos momentos de descontração, principalmente aos “fevereiros” quando um roxinho não significava nada.

Adri, Fê e Gui dizem que irmão a gente não escolhe, mas atura. Se eu tivesse que escolher, escolheria todos os dias e aturaria sempre vocês.

Juquinha, Eli e Peter, obrigada por fazerem parte dessa família, tornando os momentos de descontração ainda mais alegres (Peter, you have to learn portuguese and then read this sentence again).

Graça e Beta, minhas tias do cafezinho aos domingos. Obrigada pelo apoio fornecido à distância, pelo exemplo de simpatia e de vida.

Pedro e Berna, também conhecidos por Pai e Mãe, obrigada pelo esforço incondicional por manter a paz e tranqüilidade na minha vida quando possível. Esse trabalho e o que sou é reflexo da educação e de tudo o que vocês me proporcionaram.

Dani, obrigada por ter feito meus dias mais felizes. Obrigada por tornar tudo mais fácil, acredite! Obrigada por me fazer sorrir, por me fazer sentir saudades.

SUMÁRIO

Resumo 10

Abstract 12

Capítulo I

Prefácio 14

Introdução 15

O gênero *Drosophila* 15

O grupo *cardini* 16

Identificando por códigos de barra - Evolução cromossômica em *Drosophila* 24

Elementos de Transposição (TEs) 29

McClintock e o arco-íris nos grãos de milho – Um breve histórico 29

Variabilidade gerando variabilidade – Dinâmica dos TEs 30

Explorando novos horizontes – Evolução de TEs 35

Organizando o que gera desorganização – Classificação dos TEs 38

O TE estudado - O retroelemento *micropia* 42

Objetivos 45

Capítulo II – The LTR retrotransposon *micropia* in the *cardini* group of *Drosophila* (Diptera: Drosophilidae): a possible case of horizontal transfer 47

Capítulo III – Evolutionary history and classification of *micropia* elements in the *Drosophila* genus.....58

Capítulo IV – Brief report: Chromosomal localization of the retrotransposable element *micropia* in the *cardini* group of the *Drosophila* genus (Diptera: Drosophilidae).....121

Capítulo V – Chromosomal evolution in the *cardini* group of *Drosophila* (Diptera: Drosophilidae): Evolutionary analysis of six species of the *cardini* subgroup, their inversions and new photomaps.....139

Capítulo VI

Discussão Geral.....182

Conclusões.....186

Perspectivas.....187

Referências Bibliográficas.....190

RESUMO

Esta Tese possui dois enfoques como objetivo geral. Primeiramente, nossos dados contribuem para o conhecimento do padrão de evolução do elemento de transposição *micropia* dentro do gênero *Drosophila*, também verificando a atuação desses elementos como fonte de variabilidade genética. Em segundo lugar geramos dados citogenéticos e moleculares inéditos a respeito de espécies do grupo *cardini* de *Drosophila*, bastante freqüente nas assembléias de Drosophilidae na região Neotropical e que havia sido muito pouco estudado do ponto de vista genético até então.

Com isso, no **Capítulo II** verificamos a presença e a alta similaridade de sequências deste retroelemento em diferentes populações de *D. cardinoides*, *D. neocardini* e *D. polymorpha*. Ainda, quando comparadas com a sequência presente em uma espécie do grupo *repleta*, *D. hydei*, as sequências daquelas três espécies também apresentaram alta similaridade (97%). O grupo *repleta* e o grupo *cardini* do gênero *Drosophila* parecem ter divergido há 45 milhões de anos atrás. Portanto, para explicar os dados obtidos sugerimos a atuação de transmissão horizontal entre as espécies analisadas.

Ampliando as análises, no **Capítulo III** a presença de *micropia* foi identificada nas demais espécies do grupo *cardini* com exceção das espécies do grupo que apresentam distribuição geográfica restrita às ilhas caribenhas. As comparações com sequências presentes em outras espécies do grupo *repleta*, assim como nos 12 genomas de espécies do gênero *Drosophila* disponíveis em bancos de dados públicos, verificamos que a história evolutiva do elemento *micropia* provavelmente inclui polimorfismo ancestral e transmissão tanto vertical quanto horizontal com mecanismos de introgressão entre as espécies potencialmente atuando na geração desses padrões.

Com a finalidade de estudar a possível atuação deste retroelemento como fonte de variabilidade genética através da geração de inversões nas espécies do grupo *cardini*, o primeiro passo foi a identificação de potenciais sítios de inserção nos cromossomos politênicos dessas espécies. No **Capítulo IV** estimamos o número de cópias de *micropia* no genoma de seis espécies do grupo *cardini* (*D. cardini*, *D. cardinoides*, *D. neocardini*, *D. neomorpha*, *D. parthenogenetica* e *D. polymorpha*) verificando que varia entre seis e 18 cópias. Ainda, observamos a presença de potenciais cópias em pontos de quebra para inversões cromossômicas em três espécies. No **Capítulo IV** discutimos o significado desses dados com base nos dados já obtidos para outras espécies.

O polimorfismo cromossômico dos cromossomos politênicos e a análise comparada dessas estruturas para as seis espécies do grupo *cardini* previamente citadas, foi estudado no **Capítulo V**. Neste capítulo identificamos a presença de uma inversão nas populações estudadas de *D. cardini* e *D. neocardini*, duas inversões nas populações de *D. cardinoides* e *D. parthenogenetica* e quatro inversões nas populações de *D. polymorpha*. Das 10 inversões identificadas, sete são descritas pela primeira vez neste capítulo. Nenhuma inversão heterozigota foi encontrada na população de *D. neomorpha* aqui analisada, portanto apresentando um padrão homocariotípico. Ainda no **Capítulo V** apresentamos os primeiros fotomapas para as espécies *D. cardini* e *D. parthenogenetica*, assim como a reconstrução dos fotomapas das espécies *D. cardinoides*, *D. neocardini* e *D. polymorpha*. Também foi realizada a comparação par a par de cada cromossomo politênico entre todas as espécies com a finalidade de encontrar regiões similares úteis no estabelecimento de relações de parentesco através deste marcador. No **Capítulo V** esses dados são discutidos baseados no enfoque evolutivo para o grupo *cardini*.

ABSTRACT

This Thesis possesses two approaches as mainly objectives. First, our data contribute for the knowledge of the evolutionary pattern of the transposable element *micropia* in the *Drosophila* genus, and also to the role of these elements as source of genetic variability. Second, we generate cytogenetic and molecular new data regarding the *cardini* group species of *Drosophila* genus; these species are frequent in the Neotropical assemblies of Drosophilidae, and they had been very little studied of the genetic point of view until then.

In **Chapter II** we verify the presence and the high sequence similarity of this retroelement in different populations of *D. cardinoides*, *D. neocardini* and *D. polymorpha*. Furthermore, when these sequences are compared with the sequence present in the genome of a *repleta* group species, *D. hydei*, they show also a high similarity (97%) among them. The *repleta* group and the *cardini* group of the *Drosophila* genus seem to have diverged 45 million years ago. Therefore, to explain the obtained data we suggest the action of horizontal transmission events between the species.

Extending this analysis, in **Chapter III** the presence of *micropia* was identified in other species of the *cardini* group, with exception of the group species with geographic distribution restricted to the Caribbean islands. The comparisons with sequences present in other species of the *repleta* group as well as in the 12 *Drosophila* genomes available in public data bases, we verify that the evolutionary history of the *micropia* element probably includes ancestral polymorphism, vertical and horizontal transmission with introgression mechanisms among species potentially acting in the generation of this evolutionary pattern.

Aiming the analysis of *micropia* as source of genetic variability through the inversions generation in the *cardini* group species, the first step was the identification of potential insertion sites in the polytene chromosomes of these species. In **Chapter IV** we estimate the *micropia* copies number in the genome of six *cardini* group species (*D. cardini*, *D. cardinoides*, *D. neocardini*, *D. neomorpha*, *D. parthenogenetica* and *D. polymorpha*) verifying that it varies between six and 18 copies. Furthermore, we observed the presence of copies in the break points of chromosomal inversions for three species. In **Chapter IV** we argue the meaning of these data on the basis of the data already obtained for other species.

The chromosomal polymorphism observed in the polytene chromosomes and the comparative analysis of these structures for the six species of the *cardini* group previously cited were studied in **Chapter V**. In this chapter we identified the presence of only one chromosomal inversion in the analyzed populations of *D. cardini* and *D. neocardini*, two inversions in the *D. cardinoides* and *D. parthenogenetica* populations, and four inversions in the *D. polymorpha*. From these 10 identified inversions, seven are described for the first time in this chapter. No heterozygous inversion was found in the *D. neomorpha* population analyzed here, therefore presenting a homokaryotipic pattern. Besides this, in **Chapter V** we also present the first reference photomaps for *Drosophila cardini* and *D. parthenogenetica*, as well as the photomaps reconstruction of *D. cardinoides*, *D. neocardini* and *D. polymorpha*. We also performed a pairwise analysis of the polytene chromosomes between the species aiming to find useful regions for the establishment of evolutionary relationships based in this marker. In **Chapter V** these data are argued based on evolutionary approaches for the *cardini* group.

CAPÍTULO I

Prefácio

Este trabalho teve inicialmente a intenção de dar continuidade aos estudos genéticos realizados com espécies do grupo *cardini* do gênero *Drosophila* (Diptera: Drosophilidae). Durante a minha graduação, orientada pelo professor Dr. Paulo Hofmann, trabalhei com a Dra Daniela De Toni e conheci o grupo *cardini* juntamente com a exuberância dos cromossomos politênicos. No meu mestrado tive a oportunidade de estudar a presença de elementos transponíveis (TEs, do inglês *transposable elements*) em algumas espécies do grupo *cardini* cuja distribuição ocorre no sul do Brasil. Apesar dos obstáculos que sempre surgem quando desafiamos nosso conhecimento, com a ajuda do professor Dr. Elgion Loreto e sob a orientação da professora Dra. Vera Valente, conheci o admirável mundo dos TEs: transposons, retrotransposons, a instabilidade do genoma, mutações, controle genético, geração de inversões, reorganização genômica, etc.

No meu projeto de doutorado reuni cromossomos politênicos e um elemento de transposição, estudando a interação entre o grupo *cardini* e o retroelemento *micropia*. Tanto o grupo *cardini* quanto o retroelemento *micropia* possuem uma história recente em pesquisas. Em 1942 o grupo *cardini* começou a ser estudado (Sturtevant, 1942), salientando o elevado grau de polimorfismo inter e intraespecífico. O retroelemento *micropia* , por sua vez, foi descoberto por acaso, nas alças das estruturas de *lampbrush* dos cromossomos Y em *D. hydei* (Huijser *et al*., 1988). Os resultados da minha Tese serão descritos e discutidos a seguir.

Introdução

O gênero *Drosophila*

A família Drosophilidae (Classe Insecta: Ordem Diptera) é apenas uma das mais de 150 famílias que compõem a ordem Diptera (Yeates e Wiegmann, 2005), e encontra-se entre as mais diversas e amplamente distribuídas. Esta família compreende mais de 3500 espécies (Bächli, 2009) conhecidas vulgarmente por mosca da fruta ou mosca do vinagre.

Dentre os 40 gêneros descritos para esta família, o gênero *Drosophila* é o mais diverso com mais de 1400 espécies já descritas, presente nos mais diversos tipos de ambientes com exceção dos Pólos Norte e Sul. Neste gênero, os subgêneros mais representativos são o subgênero *Drosophila* e o subgênero *Sophophora* (revisão em Wheeler, 1981; Bächli, 2009). Segundo Tamura *et al.* (2004), estes dois subgêneros dividiram-se há cerca de 63 milhões de anos.

O subgênero *Sophophora* foi estabelecido por Sturtevant (1939) e compreende cerca de 340 espécies subdivididas em oito grupos (Bächli, 2009): *dentissima, dispar, fima, melanogaster, obscura, populi, saltans* e *willistoni*. Throckmorton (1975) sugeriu, ainda, que a radiação *Sophophora* originou-se nos trópicos do Velho Mundo e, segundo Powell *et al*. (2003), a linhagem que levou aos grupos *saltans* e *willistoni* divergiu da linhagem ancestral aos grupos *melanogaster* e *obscura* há cerca de 20 milhões de anos.

O subgênero *Drosophila* é o maior representante do gênero *Drosophila*, com pelo menos 721 espécies descritas, subdivididas em 43 grupos (Bächli, 2009): *angor, annulimana, antioquia, aureata, bizonata, bromeliae, calloptera, canalinea, carbonaria,*

cardini, *carsoni*, *coffeata*, *dreyfusi*, *flavopilosa*, *funebris*, *guarani*, *guttifera*, *histrio*, *immigrans*, *macroptera*, *melanica*, *mesophragmatica*, *morelia*, *nannoptera*, *nigrosparsa*, *onychophora*, *pallidipennis*, *peruviana*, *picta*, *pinicola*, *polichaeta*, *quadrisetata*, *quinaria*, *repleta*, *robusta*, *rubrifrons*, *simulivora*, *sticta*, *testacea*, *tripunctata*, *tumiditarsus*, *virilis* e *xanthopallescens*. Segundo Powell e DeSalle (1995), esta radiação ocorreu entre 25 e 36 milhões de anos atrás.

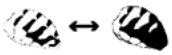
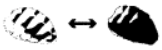
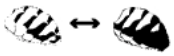
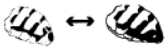
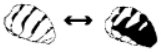
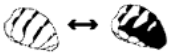
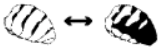
Na região Neotropical, o subgênero *Drosophila* constitui o maior representante da família Drosophilidae apresentando pelo menos 25 grupos diferentes de espécies (Val *et al.*, 1981). Malogolowkin (1953) e Throckmorton (1975) dividiram as espécies do subgênero *Drosophila* em duas linhagens principais: a seção ou radiação *virilis-repleta* e a seção ou radiação *quinaria-tripunctata*.

O grupo *cardini*

O grupo *cardini* está compreendido dentro da radiação *quinaria-tripunctata* do subgênero *Drosophila* e teve sua origem no Novo Mundo (Trockmorton, 1975). Este grupo de espécies está filogeneticamente relacionado aos grupos *calloptera*, *guarani* e *tripunctata* (Throckmorton, 1975; Grimaldi, 1990; Robe *et al.*, 2005). Alguns trabalhos de taxonomia com o grupo *cardini* reportam várias dificuldades em estabelecer critérios baseados em características de morfologia externa para identificar as espécies (Streisinger, 1946; Stalker, 1953; Heed e Wheeler, 1957; Vilela *et al.*, 2002); e, muitas vezes, é necessário recorrer à identificação pela morfologia da genitália interna do macho, o edeago (Heed, 1962; De Toni *et al.*, 2005).

Este grupo foi primeiramente estabelecido por Sturtevant (1942) incluindo *D. cardini*, *D. similis* e outras duas espécies, *D. albirostris* e *D. metzii*, que posteriormente foram classificadas como pertencentes ao grupo *tripunctata* (Vilela, 1984), demonstrando a proximidade taxonômica destes dois grupos. Desde então outras espécies foram incluídas no grupo *cardini*, baseado nos caracteres da morfologia interna e externa (Tabela 1).

Tabela 1: Divisão das espécies do grupo *cardini* nos seus respectivos subgrupos, com seu padrão de coloração abdominal

Subgrupo	Espécie	Padrão de coloração	Primeira descrição
cardini	*D. bedichecki*	?	Heed e Russel, 1971
	D. acutilabella		Stalker, 1953
	D. cardini		Sturtevant, 1916
	D. cardinoides		Dobzhansky e Pavan, 1943
	D. neocardini		Streisinger, 1946
	D. neomorpha		Heed e Wheeler, 1957
	D. parthenogenetica		Stalker, 1953
	D. polymorpha		Dobzhansky e Pavan, 1943
	D. procardinoides		Frydenberg, 1956
dunni	*D. antillea*		Heed, 1962
	D. arawakana		Heed, 1962
	D. belladunni		Heed e Krishnamurthy, 1959
	D. caribiana		Heed, 1962
	D. dunni		Townsend e Wheeler, 1955
	D. nigrodunni		Heed e Wheeler, 1957
	D. similis		Williston, 1896

Fonte: Brisson *et al*. (2006)
? não há informação para esta espécie

O grupo *cardini* tem a seguinte distribuição geográfica (Figura 1), segundo Bächli (2009):

- *D. bedicheki* é descrita para a ilha de Trinidad, porém parece ser uma espécie rara, pois não existem mais registros desta espécie (Brisson *et al.*, 2006);
- *D. acutilabella* tem distribuição restrita ao sul da América no Norte, Panamá, e algumas ilhas ao norte do Caribe;
- *D. cardini*, *D. cardinoides*, *D. neocardini* e *D. polymorpha* possuem a distribuição mais ampla do grupo: desde o sul da América do Norte até o sul o sul da América do Sul (sul do Brasil e norte do Chile e Argentina); *D. cardini* ainda tem ocorrência na ilha do Hawaii e *D. cardinoides* já foi descrita para a ilha de Galápagos;
- *D. neomorpha* e *D. parthenogenetica* têm distribuição descrita para o México, América Central, Colômbia e Brasil. A presença destas espécies no Brasil é recente (De Toni *et al*., 2005), na região norte e sul somente, e especula-se que essa distribuição seja mais ampla devido à baixa frequência de tais espécies nas coletas e à dificuldade de identificação, facilmente confundida com *D. polymorpha* e *D. cardinoides*;
- *D. procardinoides* tem a distribuição restrita a uma área de alta altitude entre Bolívia, Peru e Brasil;
- *D. antillea*, *D. arawakana*, *D. belladunni*, *D. caribiana*, *D. dunni*, *D. nigrodunni* e *D. similis* são espécies pertencentes ao subgrupo *dunni* e possuem uma distribuição restrita às ilhas caribenhas Santa Lúcia, Guadalupe, Jamaica, Martinique, Porto Rico e São Tomas, Barbados, São Vicente e Grenada, respectivamente.

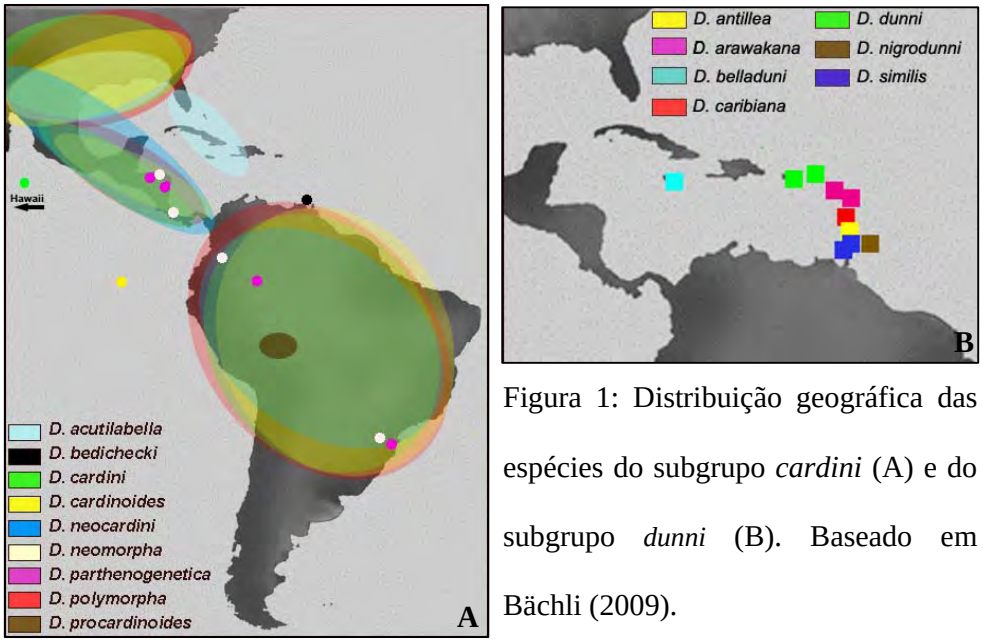


Figura 1: Distribuição geográfica das espécies do subgrupo *cardini* (A) e do subgrupo *dunni* (B). Baseado em Bächli (2009).

Originalmente, as espécies do grupo *cardini* foram divididas nos dois subgrupos (subgrupo *dunni* e subgrupo *cardini*) de acordo com a distribuição geográfica e o padrão da pigmentação das faixas abdominais dessas espécies (Tabela 1; Figura 1). Com exceção de *D. procardinoides*, que é monomórfica quanto à pigmentação abdominal, todas as demais espécies do subgrupo *cardini* são polimórficas para este marcador e apresentam uma distribuição continental (Heed e Krishnamurth, 1959). O padrão monomórfico da pigmentação abdominal de *D. procardinoides* é atribuído ao fato desta espécie apresentar uma distribuição restrita á uma região de altitude entre a Bolívia Brasil e Peru (Figura 1A), diferentemente das demais espécies deste subgrupo (Brisson *et al*., 2006). A distribuição do subgrupo *dunni* apresenta um padrão clinal (Figura 1B; Figura 2) onde as espécies de padrão mais escuro ocupam ilhas ao sul da distribuição do grupo e as espécies mais claras ocupam ilhas ao norte (Heed e Krishnamurth, 1959). Estudos evolutivos verificaram que o padrão clinal da pigmentação dessas espécies é resultado de forças seletivas atuando no

habitat durante o processo de especiação (Hollocher, 1996; Hollocher *et al*., 2000; Wilder *et al*., 2004).

No grupo *cardini*, *D. polymorpha* é a espécie que possui o maior polimorfismo de pigmentação intraespecífico que parece responder a um controle genético com a interação de três lócus (Da Cunha, 1949; Da Cunha *et al*., 1953; Heed e Blake, 1963; Martinez e Cordeiro, 1970) (Da Cunha, 1949; Heed e Russel, 1971; Brisson *et al*., 2005). Esta espécie também apresenta uma distribuição clinal do polimorfismo de pigmentação (Figura 3) com indivíduos mais claros ao sul e mais escuros ao norte da sua disctribuição (Heed, 1963). Esta espécie apresenta um padrão fenotípico sazonal (Machado *et al*., 2001) onde os indivíduos escuros aumentam em número durante os períodos mais frios do ano, por volta de 13ºC, o que parece ter uma explicação fisiológica ainda não totalmente elucidada.

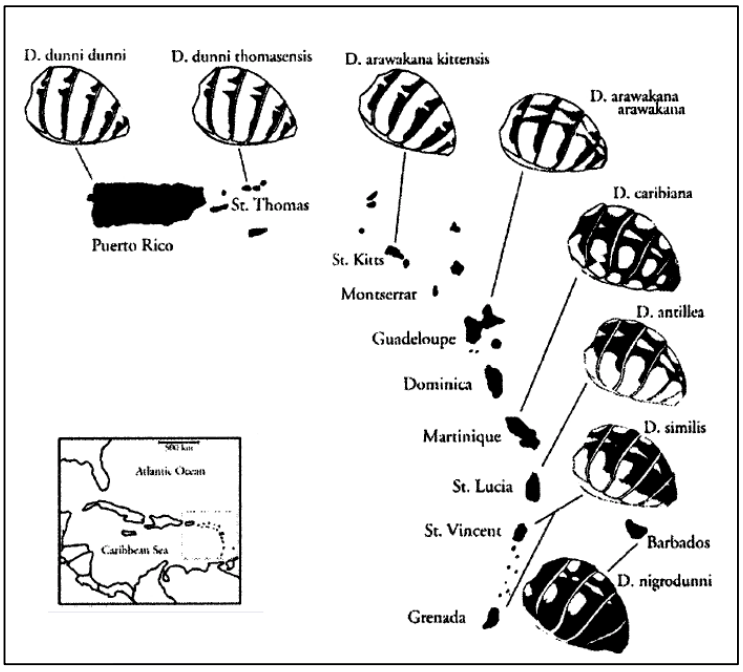


Figura 2: Distribuição dos fenótipos de pigmentação abdominal das espécies do subgrupo *dunni* nas ilhas do Caribe. Modificado de Heed e Krishnamurthy (1959).

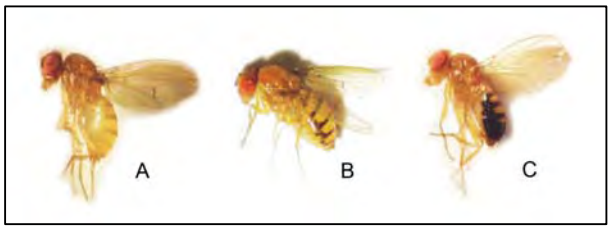


Figura 3: Padrão do polimorfismo de pigmentação em *D. polymorpha*. A. Fenótipo claro. B. Fenótipo intermediário. C. Fenótipo escuro. Modificado de Machado *et al*. (2001).

Recentemente, as relações filogenéticas do grupo *cardini* foram estudadas utilizando-se marcadores genéticos, tanto genes nucleares quanto genes mitocondriais (Figura 4) (Robe *et al*., 2005; Brisson *et al*., 2006). Previamente, hipóteses alternativas foram propostas utilizando dados de intercruzamentos (Heed e Krishnamurthy, 1959), morfologia da genitália interna masculina (Heed, 1962), inversões cromossômicas (Heed e Russel, 1971) e o padrão eletroforético de isoenzimas (Napp e Cordeiro, 1981).

Os dados de Brisson *et al*. (2006), com os genes nucleares *pgd – phosphogluconate dehydrogenase*, *mlc – myosin light chain*, e *sod – Cu, Zn superozidase dismutase* e os genes mitocondriais *COII – cytochrome oxidase II*, *cytB – cytochrome B*, e *16S – 16S ribosonal DNA*, mostram que *D. neocardini* se encontra filogeneticamente mais próxima a *D. cardinoides* (Figura 4), enquanto que os outros marcadores (morfologia da genitália interna masculina, inversões cromossômicas e isoenzimas) mostram uma maior proximidade filogenética de *D. neocardini* com *D. polymorpha*. Isto pode estar refletindo a mistura de características ancestrais presentes em *D. neocardini* no que diz respeito a características de genitália e inversões cromossômicas (Heed, 1962; Heed e Russel, 1971). Outra inconsistência foram as relações entre *D. cardinoides*, *D. parthenogenetica* e *D.*

procardinoides as quais foram agrupadas, porém com baixo suporte estatístico (Brisson *et al*., 2006). Os autores informam que essas espécies possuem uma origem recente e ainda podem estar se diferenciando tanto nos locos nucleares quanto mitocondriais.

De acordo com Brisson *et al*. (2006), as espécies do subgrupo *cardini* se divergiram há 6,6 milhões de anos e formam um grupo parafilético. Os autores explicam que a parafilia do subgrupo *cardini* é resultado da radiação das espécies do subgrupo *dunni* do continente para as ilhas. Todos os trabalhos anteriormente citados, que estudaram as relações filogenéticas do grupo, concordam com o estabelecimento de *D. cardini* como a espécie basal do grupo (Figura 4). Esta informação corrobora os dados citogenéticos, onde *D. cardini* apresenta o número diplóide de cromossomos igual a 12, todos acrocêntricos, enquanto que as demais espécies possuem oito cromossomos, o par XY acrocêntrico e os autossomos metacêntricos (Heed e Russel, 1971). Além disso, esta constituição cromossômica das espécies sugere que, durante a evolução deste grupo, ocorreram vários eventos de fusão e inversões cromossômicas (Heed e Russel, 1971).

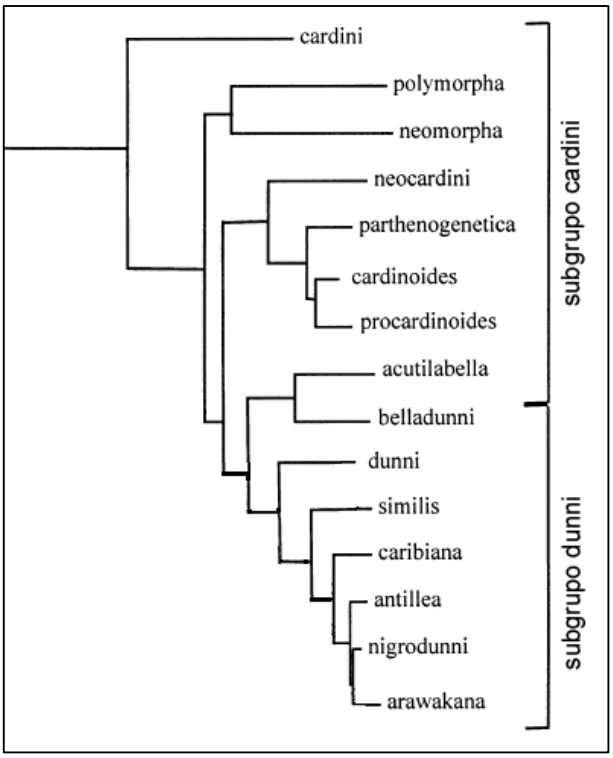


Figura 4: Reconstrução da filogenia do grupo *cardini* baseado em três genes nucleares e três genes mitocondriais. Modificado de Brisson *et al*. (2006).

O trabalho de Heed e Krishnamurthy (1959), com cruzamentos interespecíficos, mostra um padrão interessante; enquanto que as espécies presentes em ilhas intercruzam e geram híbridos férteis, as espécies presentes no continente possuem um intercruzamento mais restrito. A maioria das espécies de ilha não intercruzam com as espécies do continente, porém *D. acutilabella* potencialmente intercruza com algumas espécies do continente e com espécies presentes nas ilhas do Caribe produzindo híbridos férteis. *Drosophila acutilabella* é a única espécie do grupo *cardini* que apresenta distribuição tanto em ilhas quanto no continente (Figura 1A). As inconsistências encontradas por Brisson *et al*. (2006) nas relações filogenéticas entre as espécies do grupo *cardini* para cada marcador molecular estudado podem estar mostrando que as espécies do grupo *cardini* tiveram uma radiação rápida com a diferenciação incompleta dos genes (*incomplete lineage sorting*) nas suas espécies ancestrais.

Além do polimorfismo de pigmentação abdominal, as espécies do grupo *cardini* também são polimórficas para inversões cromossômicas (Da Cunha *et al*., 1953; Heed e Russel, 1971; Rohde e Valente, 1996a; De Toni *et al*., 2001a, 2006). Heed e Russel (1971) realizaram um extensivo trabalho delineando as relações filogenéticas das espécies do grupo *cardini* baseado nas inversões cromossômicas em cromossomos politênicos compartilhadas entre as espécies. Apesar do grande número de inversões analisadas, a dificuldade está em identificar tais inversões nos estudos atuais, pois os primeiros fotomapas de espécies do grupo *cardini* foram publicados anos depois por Rohde e Valente (1996a), para *D. cardinoides* e *D. polymorpha* e por De Toni *et al*. (2001b; 2006) para *D. neocardini* e *D. neomorpha*. Nesses estudos, *D. polymorpha* tem se mostrado a espécie mais polimórfica do grupo *cardini* (Da Cunha *et al*., 1953; Rohde e Valente, 1996a; De Toni *et al*., 2001a) e também a mais versátil, sendo encontrada em abundância nos

diferentes tipos de ambientes (De Toni e Hofmann, 1995; Rohde e Valente, 1996b; Gottschalk *et al*., 2007). Segundo Da Cunha *et al.* (1950, 1959), a quantidade de polimorfismo encontrada em uma espécie reflete a variedade de nichos ecológicos ocupados por suas populações, pois esses polimorfismos representam características que adaptaram os indivíduos da espécie à tais ambientes.

Considerando os grupos do gênero *Drosophila* nativos da região Neotropical, as espécies do subgrupo *cardini* ficam em segundo lugar em abundância, perdendo para as espécies do grupo *willistoni* que possui a maior representatividade (Gottschalk *et al*., 2007, Chaves e Tidon, 2008). As espécies do grupo *cardini*, portanto, mostram-se objetos de estudo promissores e potenciais em pesquisas de variabilidade, sendo visível o alto grau de polimorfismo cromossômico e de pigmentação presentes no grupo.

Identificando por códigos de barra - Evolução cromossômica em *Drosophila*

Os cromossomos politênicos representam um arranjo linear das unidades gênicas de uma espécie que, por serem resultado de sucessivos ciclos de endomitose, amplificando desta forma o número de cópias dos genes, têm sido bastante utilizados nos estudos de biologia molecular e evolução (Krimbas e Powell, 1992). O processo de endomitose ocorre no estágio de interfase da divisão celular e resulta em uma grande estrutura cromossômica que apresenta um padrão de bandas formado naturalmente pela sinapse precisa dos cromômeros paralelos das cromátides politenizadas. Este padrão de bandas é único para cada espécie, tornando-se único também para cada um dos cromossomos (Ashburner, 1976).

Por volta de 1935 foram realizadas as primeiras observações de inversões cromossômicas heterozigotas nos cromossomos politênicos de *D. pseudoobscura*, *D. persimilis* e nos seus híbridos (revisão em Krimbas e Powell, 1992). Desde então, tem sido demonstrado que mais da metade das espécies do gênero *Drosophila* são naturalmente polimórficas para inversões cromossômicas paracêntricas (inversão ocorrida em apenas um braço cromossômico) e pericêntricas (inversão envolvendo os dois braços cromossômicos).

Baseado nas inversões cromossômicas observadas em várias espécies de *Drosophila*, Sturtevant e Dobzhansky (1936) e Dobzhansky (1937) verificaram que seria possível inferir as relações filogenéticas entre as espécies bastando apenas apresentarem certo número de inversões em um dos braços de seus cromossomos politênicos, porém com ausência de indicação temporal dos eventos evolutivos. A metodologia de inferência das relações filogenéticas utilizada por esses autores é baseada na sobreposição de inversões paracêntricas. Através desta metodologia vários autores têm construído e reconstruído filogenias para espécies de diferentes grupos do gênero *Drosophila*: grupos da fauna havaiana (Carson e Kaneshiro, 1976), grupo *guarani* (Salzano, 1954; Kastritsis, 1969), grupo *melanogaster* (Lemeunier e Ashburner, 1976 e 1984; Lemeunier *et al.*, 1986), grupo *repleta* (Wasserman, 1982, 1992; Wasserman e Wilson, 1957 e Diniz, 1998), grupo *saltans* (Bicudo, 1973), grupo *tripunctada* (Kastritsis, 1966), grupo *virilis* (Throckmorton, 1982) e o subgrupo *willistoni* (Rohde *et al.*, 2006).

As análises mais acuradas de trabalhos recentes, que compararam sequências nucleotídicas entre os genomas das 12 espécies de *Drosophila* já disponíveis em banco de dados públicos (Bhutkar *et al*., 2008; Schaeffer *et al.*, 2008), verificaram que durante o processo evolutivo das espécies de *Drosophila* houve uma intensa reorganização das sequências cromossômicas, potencialmente auxiliada pelas inversões paracêntricas e

pericêntricas. O processo pelo qual origina as inverões ainda é muito limitado, entretanto parece haver o envolvimento de eventos de recombinação ectópica entre sequências repetitivas e invertidas nas regiões dos pontos de quebra das inversões (Cáceres *et al*., 1999, 2001; Puig *et al*., 2004). Em *D. melanogaster*, *D. simulans* e *D. yakuba* foi verificado que 59% das inversões cromossômicas parecem estar associadas com duplicações invertidas de genes ou outras sequências não repetitivas (Ranz *et al*., 2007).

Elementos de Transposição

McClintock e o arco-íris nos grãos de milho – Um breve histórico

O estudo com TEs teve seu início no final da década de 40. Naquele período, após anos de pesquisa sobre um lócus mutável no genoma de milho (*Zea mays*), a citogeneticista norte americana Barbara McClintock chegou à conclusão de que, neste organismo, o lócus *Ds* possuía a capacidade de se transpor de um lugar a outro no genoma, auxiliado pelo lócus *Ac*, potencialmente alterando a coloração da camada de aleurona nos grãos de milho. Essas descobertas não foram facilmente aceitas pela comunidade científica da época porque a existência de tais elementos móveis implicaria no dinamismo do genoma, algo inaceitável até então. Porém, 35 anos mais tarde, Barbara McClintock recebeu a menção máxima da pesquisa, o Prêmio Nobel em Medicina, pela descoberta dos elementos de transposição (Feschotte *et al*., 2002; http://profiles.nlm.nih.gov/). A partir de então os TEs foram descobertos em uma variedade de organismos, estando amplamente distribuídos em bactérias, plantas e animais (Biémont e Vieira, 2006).

Basicamente, TEs são definidos como sequências de DNA que possuem a capacidade de se transporem (elementos autônomos), ou de serem transpostos (elementos não autônomos), de uma região do genoma a outra (Capy *et al*., 1998). O nível de dinamicidade do genoma é tal que esses elementos transpassam os limites de espécies sendo capazes de invadir novos genomas (Loreto *et al*., 2008). Atualmente, sabe-se que as conseqüências da mobilização desses elementos são enormes, gerando mutações por meio de deleções e inserções de sequências nucleotídicas, rearranjos cromossômicos por meio de pareamentos ectópicos, alteração no controle de genes, etc., potencialmente proporcionando instabilidade ao genoma hospedeiro (Kidwell e Lisch, 1997; Capy *et al*., 1998; Craig *et al*., 2002). Contudo, essas conseqüências nem sempre têm um caráter nocivo, uma vez que um dos principais impactos no genoma hospedeiro é o aumento da variabilidade genética (Capy *et al.*, 1998; Kidwell e Lisch, 1997, 2001).

Variabilidade gerando variabilidade – Dinâmica dos TEs

Uma porção substancial do genoma eucarioto é composta por sequências referentes à TEs; por exemplo, estima-se que em *D. melanogaster* 22% do genoma é constituído por essas sequências (Kapitonov e Jurka, 2003), em *Zea mays* mais de 50% do genoma (SanMiguel *et al*., 1996) e em humanos 45% do genoma é composto por TEs (*International Human Genome Sequencing Consortium*, 2001). Essas sequências são abundantes e parecem ser antigas no genoma eucarioto (Kapitonov e Jurka, 2003). Além disso, os TEs são filogeneticamente muito próximos aos vírus (Xiong e Eickbush, 1990), sendo que a relação entre parasita e hospedeiro é comumente

utilizada nas relações entre TE e genoma hospedeiro, facilitando o entendimento das interações.

A primeira etapa na evolução de um TE, após ser gerado ou introduzido no genoma hospedeiro, é a amplificação do seu número de cópias, diversificando os sítios de inserção (Kidwell e Lisch, 2001). É nesta etapa que as espécies ficam mais vulneráveis aos efeitos das mobilizações dos TEs. Portanto, a maioria das mutações geradas pela mobilização dos TEs é potencialmente mutagênica para o genoma do hospedeiro (Kidwell e Lisch, 2001). Essas mobilizações possuem mecanismos regulados de forma bastante complexa e diferencialmente para cada TE. Apesar disso, o genoma hospedeiro também participa do controle, através da produção de proteínas ativadoras ou repressoras da transposição, silenciamento epigenético de genes, atuação de RNA de interferência, etc. (Lozovskaya *et al*., 1995; Labrador e Corces, 1997; Hurst e Werren, 2001; Castro e Carareto, 2004; Rij e Berezikov, 2009). Uma vez silenciadas, após longos períodos, essas sequências passam a fazer parte do DNA medianamente repetitivo do genoma hospedeiro e serão transmitidas verticalmente; sendo que, permanecendo assim, elas aparentemente perderão sua função no genoma. Porém, alguns TEs podem se tornar parte integrante do genoma por meio do processo evolutivo chamado de domesticação molecular, realizando funções vitais ao organismo. Uma dessas domesticações parecem ser os mecanismos utilizados pelas enzimas RAG1 e RAG2 durante a recombinação dos segmentos V(D)J em vertebrados. Essas recombinações geram um imenso repertório de receptores de superfície em células T e imunoglobulinas que são necessárias no reconhecimento de diversos patógenos (revisão em Sinzelle *et al*., 2009). O mecanismo catalítico pelo qual RAG1 e RAG2 afetam a recombinação é similar aos mecanismos de transposição de TEs (Lewis e Wu, 1997). O interessante é que, mesmo *in vitro*, essas duas

enzimas são capazes de catalisar a transposição de segmentos de DNA que possuem sinalização para a recombinação (Agrawal *et al*., 1998).

Em *Drosophila* há o exemplo de domesticação molecular dos elementos *Het-A* e *TART*. As espécies de *Drosophila* não produzem a enzima telomerase. Os elementos *Het-A* e *TART* parecem ter sido domesticados pelo genoma por apresentarem a capacidade de promoverem sua auto-replicação em regiões teloméricas, e desta forma, garantem a integridade dessas regiões a cada nova divisão celular. Esses TEs se transpõem repetidamente e se inserem especificamente na porção final dos cromossomos (revisão em Pardue *et al*., 2005). Além de ser um ótimo exemplo de domesticação, os TEs *Het-A* e *TART* representam um exemplo de co-evolução entre TE e genoma hospedeiro resultando em uma relação de mutualismo.

Mobilização dos TEs e suas consequências

A mobilização dos TEs em um genoma pode resultar em diferentes tipos de alterações: inserções/deleções, alterações gênicas, e rearranjos de sequências cromossômicas (Capy *et al*., 1998; Kidwell e Lisch, 2001).

Inserções e deleções - Geralmente a mobilização de um TE para um novo lugar no cromossomo gera a duplicação do sítio alvo. Consequentemente, a saída deste TE do local de origem deixa alterações naquela sequência nucleotídica, inserção de nucleotídeos previamente duplicados ou deleção de nucleotídeos próximos ao sítio de inserção devido a uma excisão imprecisa. Um exemplo relacionado a alterações no sítio de inserção é a análise do elemento *P* em *D. melanogaster* a qual verificou que, apesar deste elemento eliminar 14pb presentes ao redor do sítio alvo durante a sua inserção, os 3pb pertencentes

ao motivo do sítio alvo são restaurados pelas extremidades repetidas do TE, o que favorece a inserção de outras cópias do elemento *P* (Linheiro e Bergman, 2008).

Alterações gênicas – Com a mobilidade dos TEs, eles eventualmente podem ser inseridos em regiões codificantes (éxons), regiões não codificantes (íntrons) ou em sequências regulatórias de um gene e, desta forma, resultar em efeitos diferentes sobre a expressão do gene, dependendo da localização da inserção. Um exemplo deste tipo de interação é a atuação de RNA antisenso originado pelo transposon *Kepler* no genoma de *D. buzzatii* que silencia a atividade do gene *CG13617*, envolvido no desenvolvimento de *Drosophila*, ambos adjacentes ao ponto de quebra da inversão *2j* nesta espécie (Puig *et al*., 2004). Outro exemplo em *Drosophila* é a presença da cópia de um transposon *Tc1-like* inserido na região intergênica de duas enzimas relacionadas com a degradação do hormônio juvenil, que altera a expressão do gene *Jheh3* (Gonzalez *et al*., 2009).

Rearranjos de sequências cromossômicas – Devido à característica dos TEs de apresentarem um número de cópias variável, o pareamento não homólogo dessas cópias pode gerar alterações nos arranjos de sequências gênicas por meio de recombinações ectópicas. TEs têm sido encontrados nos pontos de quebra de inversões cromossômicas em *Drosophila* (Lyttle e Haymer, 1992; Regner *et al.*, 1996; Ledevèze *et al*., 1998). No entanto, a confirmação de que eles podem atuar como promotores de inversões cromossômicas veio com o trabalho de Cáceres *et al*. (1999) onde os autores demonstraram ser o elemento *Galileo* o principal agente de eventos de inversões nos cromossomos de *D. buzzatii*. Não obstante, Casals *et al*. (2006) reportam que nem todas as inversões onde há elementos em seus pontos de quebra foram necessariamente geradas pela recombinação ectópica daqueles elementos. Nesses casos, os TEs presentes nos pontos de quebra atuariam como invasores secundários.

Sabe-se que grande fração do genoma de diferentes organismos é composta por TEs e que, apesar de grande parte das alterações mediadas por eles serem danosas, muitos desses elementos promovem benefícios ao genoma em que estão inseridos (Sinzelle *et al*., 2009). Assim sendo, tem-se verificado que, além de promoverem a reorganização do genoma, os TEs participam ativamente da evolução dos diversos genomas em que estão inseridos (Biémont e Vieira, 2005). Outra atuação dos TEs parece ser o aumento do tamanho do genoma (valor-C), sem ter uma explicação estritamente relacionada à complexidade do organismo, mas sim com o conteúdo de TEs (Bowen e Jordan, 2002; Vieira *et al.*, 2002; Bartolomé *et al*., 2009).

A atividade dos TEs pode ser induzida como uma resposta a estresses ambientais, sendo o estresse caracterizado por uma resposta genética/fenotípica, e não simplesmente fisiológica. A consequência do aumento da atividade dos TEs é a potencial criação de variabilidade genética que, no ambiente estressante, pode gerar novidades que aumentam o *fitness* do indivíduo (Capy *et al*., 2000). González e Petrov (2009) listaram uma série de elementos presentes em *D. melanogaster* que parecem ter um valor adaptativo por terem localização fixada próximo a genes envolvidos em diferentes respostas nos processos biológicos, como aumento de temperatura, resposta a inseticidas, defesa, etc. Porém, em *D. melanogaster* apesar da maioria dos TEs estarem presentes em baixa freqüência, sugerindo que a maioria das inserções é deletéria (González *et al*., 2008), a verificação de que os TEs são fonte de variabilidade adaptativa em *Drosophila* tem sido comprovada.

Explorando novos horizontes – Evolução de TEs

Mesmo existindo a possibilidade da geração *de novo* de um TE em um genoma, através de rearranjos ou outros mecanismos - como, por exemplo, a sequência *Alu* originada a partir de um RNA 7SL (Quentin, 1992), a presença de TEs no genoma de uma espécie pode ser resultado de: transmissão vertical (VT, do inglês *vertical transmission*), onde uma sequência é herdada por descendência; ou transmissão horizontal (HT, do inglês *horizontal transmission*), onde uma sequência invade o genoma de uma espécie reprodutivamente isolada (Capy *et al*., 1998; Kidwell e Lisch, 2001). Esses processos fazem parte das fases do ciclo evolutivo dos TEs, como proposto por Kidwell e Lisch (2001).

A partir da invasão do TE na linhagem germinativa de um novo genoma, geralmente por HT, na ausência de mecanismos que regulem a atividade do novo elemento, o TE passa por um período de alta transposição ampliando o seu número de cópias no genoma invadido e na população através de VT (Le Rouzic e Capy, 2005). TEs estão sujeitos a baixas pressões seletivas o que, com o passar do tempo, permite a geração de sucessivas mutações nas regiões codificantes. Com isso, o elemento pode se tornar inativo, perdendo tanto a capacidade de produzir suas proteínas quanto a sua integridade nucleotídica (Kidwell e Lisch, 2001; Capy *et al*., 1998). Além disso, a atividade do TE pode ser requisitada durante o processo de domesticação molecular, como no exemplo dos elementos *Het-A* e *TART* (Pardue *et al*., 2005) mencionado anteriormente. Tais processos apresentam como conseqüência a divergência nucleotídica dos elementos em relação ao TE original e/ou a perda dessas sequências. Vários autores indicam que, devido aos efeitos potencialmente danosos da recombinação ectópica entre as diferentes cópias de um

TE, esses elementos passam a fazer parte da heterocromatina (Hoogland e Biémont, 1996; Vieira e Biémont, 1996; Biémont *et al*., 1997; Guerreiro *et al*., 2008) sugerindo a inativação dessas sequências (Kidwell e Lisch, 2001; Bowen e Jordan, 2002; Miller e Capy, 2006).

Portanto, a compreensão do padrão evolutivo de um TE envolve a análise de sua distribuição em diversas espécies assim como a análise da conservação nucleotídica de suas sequências. Quando o TE é transmitido de forma vertical, a sua história evolutiva tende a refletir a história evolutiva do grupo de espécies em que ele está sendo analisado. Porém, quando aparecem distorções nesse padrão evolutivo, a existência de eventos de HT pode ser investigada (Herédia *et al*., 2004; Silva *et al*., 2004; Loreto *et al*., 2008).

Além da distribuição descontínua do TE estudado entre as espécies de um mesmo grupo, a inferência de HT geralmente é explicada pela presença de algumas evidências, como alta similaridade entre sequências de TEs presentes em espécies evolutivamente distantes e a incongruência observada entre as filogenias do TE e da espécie hospedeira (Silva *et al*., 2004; Loreto *et al*., 2008). Um requisito extremamente necessário para a confirmação de que a HT de sequências nucleotídicas ocorreu é a existência de uma sobreposição geográfica, ecológica e temporal das espécies envolvidas no evento. Porém, apesar de fortes evidências favorecendo a existência de eventos de HT em uma análise, é necessário serem descartadas todas as demais possibilidades de VT, que são: a) existência de um polimorfismo ancestral para as sequências do TE analisado com distribuição independente de cópias para as espécies descendentes, b) diferentes taxas de substituição nucleotídica das cópias do TE nas diferentes espécies devido a pressões

seletivas diferenciadas - isto pode levar à menor divergência entre as sequências do TE em relação à divergência entre as sequências de genes nucleares e c) perda estocástica de sequências do TE em algumas espécies, explicando a distribuição descontínua (Silva *et al.*, 2004; Loreto *et al.*, 2008; Lisch, 2008).

O aumento do número de cópias de um TE ao longo do cromossomo aumenta a probabilidade da ocorrência de recombinação não homóloga durante a meiose, produzindo um grande número de rearranjos cromossômicos potencialmente levando a danos durante a gametogênese. Com a existência de seleção contra tais rearranjos, é esperada a localização abundante de TEs na heterocromatina, evitando assim os efeitos deletérios causados pela mobilização dos TEs (Dimitri e Junakovic, 1999).

Vários trabalhos têm mostrado que a ocorrência de HT em genomas eucariotos não é tão rara quanto se conjeturava. Foram descritos casos de HT para os retrotransposons *gypsy* (Herédia *et al.*, 2004; Ludwig *et al.*, 2008), *Penelope* (Morales-Hojas *et al.*, 2006), *copia* (Jordan *et al.*, 1999; Sanchez-Garcia *et al.*, 2005; Almeida e Carareto, 2006), *gtwin* (Ludwig e Loreto, 2007; Ludwig *et al.*, 2008), *Tom*, *297*, *17.6* e *rover* (Vidal *et al.*, 2009), jockey (Mizrokhi e Mazo, 1990; Sanchez-Garcia *et al.*, 2005) e também para os transposons *P* (Clark e Kidwell, 1997; Loreto *et al.*, 2001; Castro e Carareto, 2004), *mariner* (Brunet *et al.*, 1994; 1999), *minos* (Almeida e Carareto, 2005), *hobo* (Daniels *et al.*, 1990), entre muitos outros elementos. Bartolomé *et al.* (2009) verificaram que os retroelementos parecem contribuir com 89% dos casos de HT detectados.

Organizando o que gera desorganização – Classificação dos TEs

Baseado nos mecanismos de transposição dos TEs basicamente existe duas grandes classes que os divide (Finnegan, 1989; Wicker *et al*., 2007):

• Classe I – retroelementos ou elementos que possuem um intermediário de RNA durante a sua transposição, também chamados de retrotransposons. Utilizam a enzima transcriptase reversa no processo de transposição. Os elementos dessa classe também eram conhecidos por elementos copia-e-cola, por terem o comportamento de deixarem uma cópia no local de origem. Este mecanismo é chamado de transposição replicativa. Fazem parte desta classe os elementos com longas repetições terminais (LTR, do inglês *long terminal repeats*), como o elemento *gypsy* e *copia*, os elementos *DIRS*-like (*DIRS*, *Ngaro* e *VIPER*), os elementos *Penelope*-like (PLE), os LINEs (*long intersperced nuclear elements - jockey* e *L1* e os SINEs (*small intersperced nuclear element*).

• Classe II – transposons ou elementos que possuem um intermediário de DNA durante a sua transposição, e utilizam a enzima transposase durante este processo. Nessa classe os elementos eram conhecidos por corta-e-cola, por geralmente serem excisados do local de origem ao se transporem. O mecanismo é chamado de transposição conservativa. Fazem parte desta classe os elementos com repetições terminais invertidas (TIR, *terminal inverted repeats*), como os elementos das famílias *Tc1-mariner*, *hAT*, *Mutator*, *P*, *CACTA* e os elementos *Crypton*, *Helitron* e *Maverick.*

Devido ao crescente número de novos TEs descritos, o sistema de apenas duas classes proposto por Finnegan (1989) foi aprimorado por Capy *et al*. (1998), criando o nível de subclasse que separa os elementos baseado nas suas estruturas: para a classe I, separa os elementos com longas repetições terminais dos elementos sem LTR; para a

classe II separa os elementos com e sem assinatura DDE (dois ácidos aspárticos e um ácido glutâmico) na transposase.

Posteriormente, com a descoberta de uma variedade de TEs que não se encaixava nesse sistema de classificação - por terem um intermediário de RNA, mas sem manter a cópia no local de origem por exemplo, Wicker *et al*. (2007) sugeriram um novo modelo de classificação, mantendo a divisão inicial em classes como proposta por Finnegan (1989), resultando na seguinte estruturação:

• Classes: como anteriormente, divide os TEs baseado na presença ou ausência de um intermediário de RNA durante a transposição;

• Subclasse: distingue os elementos de mecanismo copia-e-cola dos corta-e-cola;

• Ordem: separa os elementos baseado nas diferenças dos mecanismos de inserção, consequentemente considera as características enzimáticas do processo;

• Superfamília: separa os elementos baseado na estrutura das proteínas, na estrutura dos domínios não codificantes, na presença e no tamanho dos sítios alvos de duplicação (TSD, do inglês *target site duplication*);

• Família: separa os elementos de acordo com a similaridade de sequências nucleotídicas. No intuito de facilitar a utilização do sistema proposto, os autores estabeleceram que, na análise de sequências maiores de 80pb, as mesmas podem ser consideradas como pertencentes à mesma família se apresentarem 80% de similaridade em pelo menos 80% da sequência do seu domínio codificante, das regiões terminais repetidas, ou de ambos;

• Subfamílias: neste nível, a classificação utiliza as relações filogenéticas dos elementos de uma mesma família como critério de divisão dos TEs;

• Inserção: este último nível taxonômico é voltado para os anotadores dos projetos genoma dos diferentes organismos. Ele descreve uma cópia em particular do TE correspondendo a um específico evento de transposição.

Uma das características importantes dos TEs é a existência de elementos autônomos e não autônomos (ou defectivos). Enquanto que os elementos autônomos codificam as enzimas necessárias a sua própria transposição, os elementos não autônomos, por terem perdido parte ou todo o seu domínio codificante, dependem das enzimas fornecidas *in trans* pelos elementos autônomos. Neste caso, no sistema de classificação proposto por Wicker *et al*. (2007) cópias autônomas e não autônomas de um mesmo elemento, pertencente a uma certa família, podem ser classificadas como diferentes subfamílias. Ainda, os elementos não autônomos podem ter homologia de sequência suficiente que possibilite classificá-los em diferentes subfamílias (Wicher *et al*., 2001). Sequências autônomas e não autônomas de um mesmo elemento geralmente apresentam a região 5'UTR altamente conservadas, pois essa região é necessária durante sua transposição (Kidwell e Lisch, 1997; Wicker *et al*., 2007).

A Figura 5 mostra esquematicamente a estrutura organizacional dos elementos pertencentes às duas classes. Outras características moleculares, além das indicadas na Figura 5, são utilizadas para identificar os TEs (revisão em Wicker *et al*., 2007).

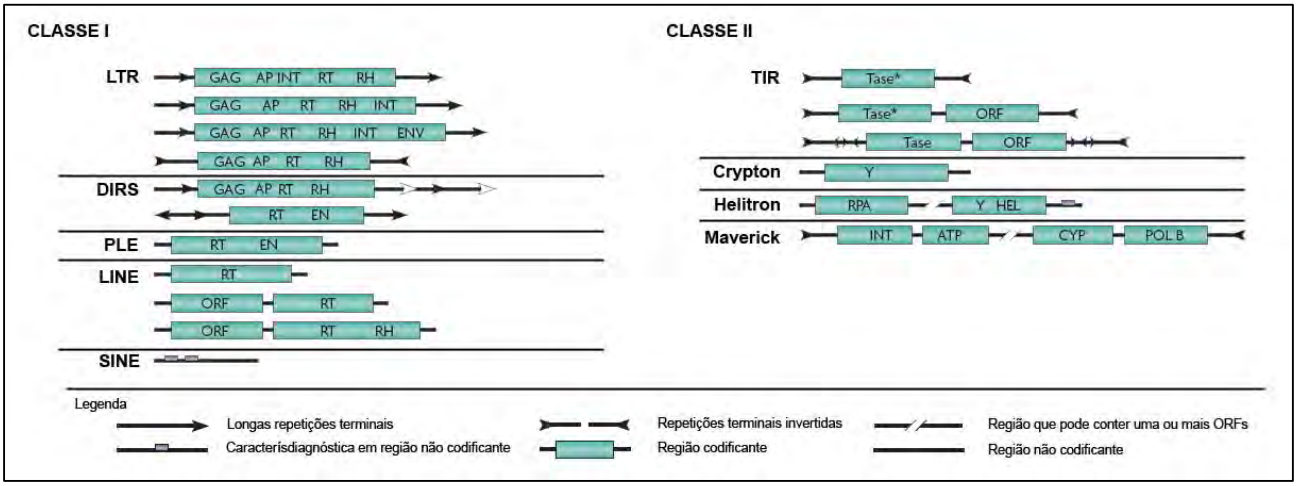


Figura 5: Estruturas dos elementos de Classe I e Classe II. GAG: proteína do capsídeo. AP: proteinase. INT: integrase. RT: transcriptase reversa. RH: RNase H. Env: proteína do envelope viral. En: endonuclease. ORF: nesta figura significa região de leitura de função desconhecida. Tase: transposase (*com motivo DDE). YR: recombinase. RPA: proteína A de raplicação (encontrado apenas em plantas). HEL: helicase. ATP: ATPase de empacotamento. CYP: cisteína protease. PolB: DNA polimerase B. Modificado de Wicker *et al.* (2007).

Este sistema de classificação hierárquico proposto por Wicker *et al.* (2007) recebeu críticas principalmente em relação à definição de família de TEs. Seberg e Petersen (2009) argumentam que o sistema não reflete a filogenia dos TEs uma vez que ele pode gerar uma mistura de grupos monofiléticos, parafiléticos e polifiléticos quando estipula a simples similaridade nucleotídica das sequências para agrupá-las em famílias. Devido à característica intrínseca dos TEs de se mobilizarem dentro do genoma de uma espécie e entre genomas de diferentes espécies (Loreto *et al.*, 2008), eles estão sujeitos a mecanismos (mutações, recombinações, deleções...) que elevam a sua taxa de mutação (Craig *et al.*, 2002). Devido a essas características, a filogenia de TEs muitas vezes envolve uma série de interações gerando um padrão evolutivo bastante complexo (Herédia *et al.*, 2004). Apesar das críticas, e por não existir outra classificação para TEs oficialmente

aceita pela comunidade científica, atualmente alguns autores vêm utilizando o sistema proposto por Wicker *et al.*(2007) (Charles *et al*., 2009; Piednoël e Bonnivard, 2009; Staginnus *et al*., 2009).

O TE estudado - O retroelemento *micropia*

O elemento *micropia* pertence à classe I de elementos de transposição, portanto é um retroelemento, da ordem dos elementos com LTR e superfamília *Ty*3 (Finnegan, 1989; Wicker *et al*., 2007). É constituído por duas ORFs e tem a seguinte composição (Figura 6): a primeira ORF refere-se ao gene GAG que produz uma poliproteína que é processada em três proteínas maduras de produção do capsídeo, com similaridade com proteínas virais; a segunda ORF constitui-se do gene que codifica as enzimas necessárias à transposição do elemento, proteinase (AP), transcriptase reversa (RT), RNase H (RH) e integrase (INT). *Micropia* parece ainda ter uma terceira ORF, porém de função não reconhecida (Huijser *et al*., 1988; Lankenau *et al*., 1988).

Figura 6: Esquema básico do retroelemento com LTR *micropia*. GAG: proteínas do capsideo. AP: proteinase. RT: transcriptase reversa. RH: RNase H. INT: integrase. Modificado de Wicker *et al*. (2007).

Micropia foi primeiramente encontrado nas alças das estruturas de *lampbrush* presentes no cromossomo Y de *D. hydei* durante o estágio de prófase da divisão meiótica dos espermatócitos primários (Hennig *et al.*, 1983; Huijser *et al.*, 1988). O nome *micropia*

vem da técnica empregada na sua descoberta (experimentos de microclonagem) e da sua similaridade com o retroelemento *copia*. Sua caracterização molecular mais completa foi feita para as cópias presentes em *D. hydei* (clones *micropia*-dhMiF2 e *micropia*-dhMiF8) (Huijser *et al.*, 1988) e *D. melanogaster* (clones *micropia*-Dm11 e *micropia*-Dm2) (Lankenau *et al*., 1988; 1990). Na sua extensão, o elemento *micropia* possui duas regiões abertas para leitura que codificam proteínas similares a proteínas de nucleocapsídeo, transcriptase reversa, integrase RNAseH e protease de retrovírus (Lankenau *et al.*, 1988; Huijser *et al.*, 1988).

Sequências similares a *micropia* já foram identificadas em espécies do subgênero *Sophophora* (grupos: *melanogaster*, *willistoni*, e *saltans*) e no subgênero *Drosophila* (grupos: *immigrans, funebris* e *repleta*) (Lankenau, 1993; Almeida *et al.*, 2001; Almeida e Carareto, 2004), ambos subgêneros do gênero *Drosophila*. Almeida e Carareto (2004) identificaram a existência de duas possíveis subfamílias do elemento *micropia* em espécies do grupo *repleta*. Esses dados sugerem uma herança baseada em transmissão vertical deste elemento dentro do gênero *Drosophila*.

Drosophila hydei e *D. neohydei* são espécies irmãs e possuem o mesmo padrão de bandeamento entre os seus cromossomos politênicos. Nessas espécies o elemento *micropia* compartilha alguns sítios de inserção, indicando que este elemento apresentou atividade transposicional após a divergência dessas duas espécies (Hennig *et al*., 1983). Em *D. hydei micropia* possui 11 cópias nos cromossomos X e autossomos e cerca de 50 a 100 cópias no cromossomo Y (Lankenau *et al*., 1994). Outro estudo com diferentes espécies do grupo *repleta* mostrou que elas parecem ter entre sete e dezessete cópias no seu genoma (Almeida e Carareto, 2004). Já para *D. melanogaster* é estimado que a espécie possua entre 16 e 32 cópias do retroelemento *micropia* em seu genoma (Lankenau, 1993).

Trabalhos anteriores reportam que *micropia* expressa tanto RNA senso quanto antisenso, indicando um controle mediado por RNA de interferência (Lankenau, 1993; Lankenau *et al*., 1994; Almeida e Carareto, 2004). O transcrito senso que possui 5kb, referente ao tamanho total do elemento, é expresso tanto nas células somáticas quanto germinativas de machos e fêmeas de *D. hydei*. O elemento produz dois transcritos antisenso, de 1kb e de 1,6kb de extensão, e a presença desses transcritos parece estar limitada às células germinativas dos machos. O transcrito de 1kb parece ser 30 vezes mais abundante nas células germinativas do que os transcritos senso de 5kb. A origem do transcrito antisenso de 1kb parece envolver uma região promotora localizada entre as sequências que dão origem às proteínas RNase H e integrase. A região de iniciação da transcrição parece estar dez nucleotídeos distante de uma sequência idêntica ao promotor da proteína β2-tubulina testículo específica em *D. hydei*. Este promotor não está presente nas sequências *micropia*-Dm11 e *micropia*-Dm2, extraídas de *D. melanogaster*, o que é consistente com a inexistência da expressão dos transcritos antisenso nessa espécie. O controle da expressão do transcrito senso de 5kb parece ser feito através do direcionamento do transcrito antisenso de 1kb às regiões referentes à RNase H e parte da transcriptase reversa, contribuindo para a degradação do RNA e/ou simplesmente para a interferência da sua tradução (Lankenau, 1990; 1993; Lankenau *et al*., 1994; Almeida e Carareto, 2004).

Até agora o elemento *micropia* parece não estar envolvido na geração de mutações, levando a crer que o mecanismo acima descrito, ou algum outro mecanismo de controle de transposição de TEs, parece estar reprimindo efetivamente este elemento (Lankenau, 1993).

Objetivos

Com base nas informações fornecidas sobre o grupo *cardini* assim como sobre o retroelemento *micropia* esta Tese teve como objetivo geral contribuir para o conhecimento do padrão de evolução do elemento de transposição *micropia* dentro do gênero *Drosophila* e dos TEs como fonte de variabilidade genética. Ainda tivemos o objetivo de gerar dados citogenéticos e moleculares inéditos a respeito de espécies do grupo *cardini* de *Drosophila*, bastante freqüente na região Neotropical e que havia sido muito pouco estudado do ponto de vista genético até então. Desta forma, delineamos os diferentes objetivos específicos:

- Analisar *in silico*, a similaridade de seqüências do retroelemento *micropia* dentro do grupo *cardini*, traçando a sua história evolutiva, comparativamente com a do grupo *repleta* e com as 12 espécies que possuem genomas disponíveis em bancos de dados públicos, gerando o cenário evolutivo deste retroelemento no gênero *Drosophila* (Capítulos II e III);

- Inferir o número de cópias do retroelemento *micropia* nos cromossomos politênicos de seis espécies do grupo *cardini* de *Drosophila* por hibridação *in situ*, tentando relacionar seus locais de inserção com pontos de quebra para inversões cromossômicas (Capítulo IV); e,

- Caracterizar o polimorfismo de inversões cromossômicas de seis espécies do grupo *cardini*, construir os fotomapas de referência dos cromossomos politênicos de *Drosophila cardini* e *Drosophila parthenogenetica*, bem como reconstruir os fotomapas de referência anteriores das espécies *D. polymorpha*, *D. cardinoides* e *D. neocardini*, melhorando a definição dos seus padrões de bandas e interbandas para poder analisar comparativamente os cromossomos politênicos das seis espécies sob o ponto de vista evolutivo (Capítulo V).

CAPÍTULO II

Trabalho publicado na revista científica Genetica 134(3): 335-344 (2008)

The LTR retrotransposon *micropia* in the *cardini* group of *Drosophila* (Diptera: Drosophilidae): a possible case of horizontal transfer[1]

Juliana Cordeiro[*§]; Lizandra J. Robe[§]; Elgion L. S. Loreto[#]; Vera L. S. Valente[*]

[*] Programa de Pós Graduação em Biodiversidade Animal, Departamento de Biologia, Centro de Ciências Naturais e Exatas, Universidade Federal de Santa Maria, Santa Maria, RS, Brazil

[§] Programa de Pós Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

[#] Departamento de Biologia, Universidade Federal de Santa Maria, Santa Maria, RS, Brazil

Programa de Pós Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

Corresponding author: Juliana Cordeiro, Laboratório de *Drosophila*, Departamento de Genética, UFRGS, Caixa Postal 15053, CEP: 91501-970. Phone: +55-51-33166729; Fax: +55-51-33167311; e-mail: jlncdr@gmail.com

[1] A parte inicial deste trabalho, referente à coleta dos dados de *micropia*, foi apresentada na dissertação de mestrado da doutoranda, Juliana Cordeiro.

Genetica (2008) 134:335–344
DOI 10.1007/s10709-008-9241-2

The LTR retrotransposon *micropia* in the *cardini* group of *Drosophila* (Diptera: Drosophilidae): a possible case of horizontal transfer

Juliana Cordeiro · Lizandra J. Robe · Élgion L. S. Loreto · Vera L. S. Valente

Received: 13 September 2007 / Accepted: 19 January 2008 / Published online: 8 February 2008
© Springer Science+Business Media B.V. 2008

Abstract The presence of the *micropia* retroelement from the *Ty1-copia* family of LTR retroelements was investigated in three species of the *Drosophila cardini* group. *Southern* blot analysis suggested the existence of at least four *micropia* copies in the genomes of *D. cardinoides*, *D. neocardini* and *D. polymorpha* populations. The high sequence similarity between dhMiF2 and Dm11 clones (*micropia* retroelements isolated from *D. hydei* and *D. melanogaster*, respectively) with *micropia* sequences amplified from *D. cardini* group genome supports the hypothesis that this retroelement plays an active role in horizontal transfer events between *D. hydei* and the *D. cardini* group.

Keywords *Drosophila cardini* subgroup · *Drosophila hydei* · Horizontal transfer · *Micropia* · Retroelements · Transposable elements

Introduction

The *Drosophila cardini* group belongs to the subgenus *Drosophila* and comprises 16 species that inhabit the tropical and subtropical regions in the Americas. The group is conventionally subdivided into two subgroups: the *dunni* subgroup, which includes eight species found in the Caribbean islands and that are intraspecifically monomorphic for abdominal pigmentation; and the *cardini* subgroup, which assembles the other eight species distributed across the Americas, from Mexico to southern Brazil, and that exhibit high intraspecific polymorphism for abdominal pigmentation, except for *D. procardinoides* (Heed and Krishnamurthy 1959; Heed and Russel 1971; Brisson et al. 2006). The species analyzed in this work, *Drosophila cardinoides*, *D. neocardini*, and *D. polymorpha*, are members of the *D. cardini* subgroup that occur in South America (Vilela et al. 2002).

Recently, Brisson et al. (2006) presented the phylogenetic relationships within the *D. cardini* species group. The investigation was based on three nuclear genes (*pgd—phosphogluconate dehydrogenase*, *mlc—myosin light chain*, and *sod*—Cu, Zn *superoxidase dismutase*) and three mitochondrial genes (*COII—cytochrome oxidase II*, *cytB—cytochrome B*, and *16S–16S ribosomal DNA*). Alternative phylogenetic hypotheses had previously been proposed by Heed (1962) using male genitalia data, by Heed and Russel (1971) using chromosomal inversions and intercrossing, by Napp and Cordeiro (1981) using isoenzymes, and by Robe et al. (2005) using the genes *Amd* (α-methyl dopa) and *COII* as molecular markers. According to the total evidence hypothesis presented by Robe et al. (2005) and Brisson et al. (2006), *D. neocardini* is phylogenetically closer to *D. cardinoides* than to *D. polymorpha*, whereas according to Heed (1962), Heed and Russel (1971) and to Napp and Cordeiro (1981), *D. neocardini* and *D. polymorpha* bear closer phylogenetic affinities. It is assumed that the *D. cardini* species subgroup is a paraphyletic group that diverged 6.6 million years ago (Brisson et al. 2006).

Electronic supplementary material The online version of this article (doi:10.1007/s10709-008-9241-2) contains supplementary material, which is available to authorized users.

J. Cordeiro · É. L. S. Loreto
Programa de Pós Graduação em Biodiversidade Animal, Departamento de Biologia, Centro de Ciências Naturais e Exatas, Universidade Federal de Santa Maria, Santa Maria, RS, Brazil

J. Cordeiro (✉) · L. J. Robe · É. L. S. Loreto · V. L. S. Valente
Programa de Pós Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, Porto Alegre, RS CEP: 91501-970, Brazil
e-mail: jlnedr@yahoo.com.br

The *D. cardini* species group has been the subject of research on the polymorphic characteristics such as abdominal pigmentation (Heed 1963; Heed and Blake 1963; Hollocher et al. 2000a, b), and chromosomal inversions (Da Cunha et al. 1953; Rohde and Valente 1996; De Toni et al. 2001). Yet, few studies have included the *D. cardini* species group in the search for mobile elements. For some of these species the presence of *gypsy* (Loreto et al. 1998; Herédia et al. 2004) and *mini-me* retroelements (Wilder and Hollocher 2001), and the absence of *P*, *I*, *mariner* (Loreto et al. 1998) and *hobo* (Daniels et al. 1990; Loreto et al. 1998) have been reported.

The *micropia* retroelement was first discovered in *D. hydei*, a member of *D. repleta* group from the subgenus *Drosophila* (Huijser et al. 1988). The *micropia* retroelement belongs to the *Ty1-copia* family of LTR retroelements (Capy et al. 1997), and is present in the genome of some subgenus *Sophophora* species (*melanogaster*, *willistoni*, and *saltans* groups), and some subgenus *Drosophila* species (*immigrans*, *funebris* and *repleta* group) (Lankenau 1993; De Almeida et al. 2001; De Almeida and Carareto 2004). In structural terms, a far-stretching characterization has been made only for *micropia* in *D. hydei* (dhMiF2 and dhMiF8 clones) (Huijser et al. 1988) and *D. melanogaster* (Dm11 and Dm2 clones) (Lankenau et al. 1988, 1990). Recently, two subfamilies of *micropia* were identified in the *D. repleta* group (De Almeida and Carareto 2004). These data—side by side with the fact that *micropia* has been detected in species from six different groups of the genus *Drosophila*—suggest that this element was already present in the *Drosophila* species genome well before the diversification of the genus.

The establishment of a TE phylogenetic distribution is important to study its evolutionary history. The distribution pattern may be the result of vertical transmission (VT), in which the element is present in the ancestral genome, or of horizontal transmission (HT), when the elements' phylogenetic history is incompatible with the host species ones (Silva et al. 2004). Considering that the *D. cardini* group has a particular ecology and speciation process that could provide peculiar TE evolutionary patterns, we isolated sequences similar to *micropia* in the genomes of several populations belonging to three species of the *D. cardini* group, aiming to contribute to the current knowledge of *micropia* evolutionary dynamics in the *Drosophila* genus. The analyses carried out with the addition of the *micropia* sequences available support the premise of HT between species of the *D. repleta* and *D. cardini* groups.

Materials and methods

Drosophila species and populations

Table 1 lists the species and populations used in this study, with their locations, year of collections and GenBank *micropia* accession numbers. For all the analyzed stocks,

Table 1 *Drosophila cardini* group populations investigated in this study with their respective source and GenBank *micropia* accession numbers

Species	Population	Localization	Coordinate	Year	GenBank Acc. Nos.
D. cardinoides	CMS[a]	Morro Santana	S 30° 04′ W 51° 08′	2002	EU149929
	CUF[a]	Universidade Federal de Santa Catarina	S 27° 36′ W 48° 31′	2001	EU149930
	CCE[a, b]	Florianópolis downtown	S 27° 35′ W 48° 33′	2001	EF090263
D. neocardini	NAR[a]	Arvoredo Island	S 27° 17′ W 48° 21′	1999	EU149931
	NCE[a, b]	Florianópolis downtown	S 27° 35′ W 48° 33′	2001	EF090264
	NMC[a]	Morro da Cruz	S 27° 35′ W 48° 31′	2001	EU149933
	NUF	Universidade Federal de Santa Catarina	S 27° 36′ W 48° 31′	2001	–
	NJO[a]	Joinville city	S 26° 17′ W 49° 01′	2004	EU149932
D. polymorpha	PCI[a]	Centro de Instruções de Santa Maria	S 29° 32′ W 53° 42′	2003	EU149934
	PMS[a]	Morro Santana	S 30° 04′ W 51° 08′	2002	EU149935
	PTU[a]	Parque do Turvo	S 27° 20′ W 53° 40′	1999	EU149937
	PCE[a, b]	Florianópolis downtown	S 27° 35′ W 48° 33′	2001	EF090265
	PMC	Morro da Cruz	S 27° 35′ W 48° 31′	2001	–
	PML	Morro da Lagoa da Conceição	S 27° 35′ W 48° 28′	1999	–
	PMA	Mangue do Itacorubi	S 27° 34′ W 48° 31′	2003	–
	PRE[a]	Joaquina's beach	S 27° 38′ W 48° 27′	2003	EU149936
	PUF	Universidade Federal de Santa Catarina	S 27° 36′ W 48° 31′	2001	–

[a] Only these population were used for phylogenetic analysis

[b] These populations had the *micropia* fragment cloned

isofemale strains were initially established to reduce intrapopulational variability. These strains have been maintained in the laboratory by means of mass crosses and reared in corn flour culture medium, in a controlled chamber (17 ± 1°C, 60% r.h.).

Southern blot analysis

Seventy adult flies per sample were macerated in a 1.5 ml microcentrifuge tube using liquid nitrogen. Genomic DNA was extracted according to a phenol:chlorophorm protocol (Sassi et al. 2005). DNA samples (7 µg) were digested with restriction enzymes BglII and HindIII (Invitrogen). The BglII restriction enzyme has no cleavage site in *D. hydei micropia* sequence, and the HindIII restriction enzyme has one cleavage site in nucleotide 2,628 of this sequence (Fig. 1). As positive control, we used 5 ng of the *micropia* probe. DNA fragments were electrophoresed on 0.8% agarose gels and transferred to nylon membranes (Hybond N + ®, GE Healthcare). Probes were labeled using the chemioluminescent hybridization system Gene Images® (GE Healthcare), according to manufacturer's instructions. Membranes were hybridized at 60°C, adding 100 µg/ml sonicated salmon sperm DNA (Invitrogen) in the prehybridization and hybridization solutions. Each membrane was washed for 15 min at 60°C, first at 1 × SSC and 0.1% SDS, and then with 0.5 × SSC and 0.05% SDS. For signal detection, the Gene Images CDP-Star® detection module (GE Healthcare) was used, and each film underwent 1 min exposure. The probe used was a 3.1 kb fragment obtained by EcoRI digestion of the dhMiF2 plasmid (NCBI accession number X13304) (Fig. 1) (Huijser et al. 1988). Comparing the incomplete dhMiF2 clone with dhMiF8 and Dm11 clones (Huijser et al. 1988; Lankenau et al. 1988, 1989) we inferred the

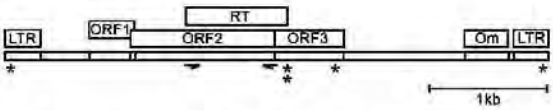


Fig. 1 Restriction map of *D. hydei micropia* retroelement clone, dhMiF2, indicating the restriction sites for enzymes EcoRI (*, nucleotide 2; 3,053 and 4,805) and HindIII (**, nucleotide 2,628). *micropia*-dhMiF2 is approximately 4,809 bp long. Arrows indicate the annealing position of the primers used (forward primer anneals in nucleotide 1,777 and reverse primer anneals in nucleotide 2,589 in the complementary strand). The region of 3.1 kb comprising the left LTR and the three ORF's, as obtained by EcoRI cleavage, was utilized as the *Southern* blot probe. This scheme was based on Huijser et al. (1988), Lankenau et al. (1988), and on the data available from the NCBI website (accession number X13304—dhMiF2 *micropia* element). LTR: long terminal repeats. RT: reverse transcriptase. ORF: open reading frame. Om: other motifs

micropia-dhMiF2 length to be roughly 4.8 kb, with LTRs flanking ORFs that code proteins that are similar to nucleocapsid, protease, reverse transcriptase, RNAse H, and also to human retroviruses integrase (Huijser et al. 1988; Lankenau et al. 1989).

Sequence and phylogenetic analysis

Polymerase chain reactions (PCR) were performed in 25 µl volumes using approximately 25 ng template DNA, 5 pMol of each primer, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$ and 1 unit Taq DNA polymerase in 1× Polymerase Buffer (all from Invitrogen). The PCR cycling was performed at a 64°C annealing temperature, with the use of the primers MIC1777 (5′ CTCCCCTTTTGCCAG TCCT 3′) and MIC2570 (5′ TTGAGCTAGCGTCGGTG TG 3′), drawn according to dhMiF2 sequence (accession number X13304). These primers anneal into *micropia*'s reverse transcriptase (Fig. 1) generating a fragment of about 810 bp. The dhMiF2 plasmid was used as positive control. The amplicons were separated by 1.5% agarose gel electrophoresis and purified using QIAquick® DNA-Gel Extraction System (Qiagen) according to the supplier's specifications. One population of each species (Table 1) had its fragment cloned in TOPO-TA Cloning® Vector (Invitrogen) in a pilot research. The recombinant plasmids obtained were purified using QIAprep® Miniprep Kit (Qiagen) according to manufacturer's instructions and sequenced in a MegaBace 500 automatic sequencer with M13 standard primers. These same populations also had their *micropia* fragments directly sequenced using the same *micropia* PCR primers. The sequences obtained with both methods were compared and few or no nucleotide differences were found between sequences. Considering this result and the fact that in almost all the cases the electropherogram of the directly obtained sequences reflected the few polymorphic sites, we decided to adopt a direct sequencing strategy for the remaining populations' strains. In all cases, both strands were sequenced. Ambiguities and compressions were resolved when the contigs were assembled in Staden Package Gap 4 program (Staden 1996). The GenBank accession numbers at National Center for Biotecnology Information website (NCBI—http://www.ncbi.nlm.nih.gov/) of the sequences obtained are indicated in Table 1. For the phylogenetic analysis, we used the *micropia* sequences of *D. hydei* (clone dhMiF2—accession number X13304) and of *D. melanogaster* (clone Dm11—accession number X14037) obtained from the nucleotide sequence database in the NCBI website.

All these sequences were aligned using the ClustalX 1.81 (Jeanmougin et al. 1998), by the program's default

parameters (The alignments are available in the site: http://www.ufsm.br/labdros/links/micropia1.aln). The phylogenetic relationships between the *micropia* sequences were estimated using three different methods:

(1) Neighbor-joining (NJ) (Saitou and Nei 1987), using the Mega 3.1 program (Kumar et al. 2004), with the Tamura three-parameter nucleotide substitution model (Tamura 1992);
(2) Maximum parsimony (MP), using the branch-and-bound search algorithm (Hendy and Penny 1982), in PAUP 4.0b10 program (Swofford 2003);
(3) Maximum likelihood (ML), with the PAUP 4.0b10 program using the HKY nucleotide substitution model (Hasegawa et al. 1985), without gamma or invariable sites correction, as proposed in the ModelTest program (Posada and Crandall 1998).

Confidence values for each clade in the three trees were calculated by the bootstrap test (Felsenstein 1985) with 1,000 replications. In the phylogenetic analyses conducted, the gaps in the sequences were considered as lost data by the Mega 3.1 program, according to the pairwise deletion method, and as a new character state by the PAUP 4.0b10 program. The two deletions generated by the alignment in *D. melanogaster* (23 and 10 bp long), the only deletion in *D. hydei* (2 bp), and the only 133 bp long deletion in *D. polymorpha* were adjusted, so that phylogenetic programs could interpret each of these as a result of one single *indel* event.

The matrix of divergence was constructed through the Tamura three-parameter model, both for *micropia* and a nuclear gene sequence, *Amd*. The *Amd* sequences were obtained from the NCBI website (accession numbers in GenBank AY699258 for *D. cardinoides*, AY699260 for *D. neocardini*, AY699259 for *D. polymorpha*, AF293712 for *D. hydei*, and X00695 for *D. melanogaster*). To check whether *micropia* sequences present a coding potential, the amino acid sequences were obtained using the GeneDoc program, version 2.6.001 (Nicholas and Nicholas 1997). The similarity/identity of *micropia* amino acid sequences as well as of *Amd* amino acid sequences presented by the different species were also calculated using that program.

The aligned sequences were used to obtain the values of synonymous substitution per synonymous sites (d*S*) and of nonsynonymous substitution per nonsynonymous sites (d*N*), using the p-distance analysis and the Nei and Gojobori method (Nei and Gojobori 1986), as implemented in Mega 3.1 program (Kumar et al. 2004). Additionally, a selection *Z*-test was carried out to check whether d*N* and d*S* values differed significantly, and thus infer the probable kind of selection pressure acting on the sequences.

Results

Inference of the insertion pattern of *micropia* in *Drosophila cardini* group

Figure 2a and b show blots suggesting an estimation of the *micropia* insertion sites of *D. cardinoides*, *D. neocardini*, and *D. polymorpha* populations that had their genomic DNA digested with the BglII restriction enzyme. Populations of the three species generated a strong band in the blots, with as little as 1 min film exposure, suggesting that there are sequences in the genomes that could be highly similar to the *D. hydei micropia* element used as probe. The species analyzed presented some insertions site number variation, from four to six, in which *D. neocardini* populations presented the stronger signs of hybridization, and *D. cardinoides* populations showed the weaker ones. There was a small population insertion site polymorphism within *D. neocardini* and *D. polymorpha* (as given by the arrows in Fig. 2), different from the two *D. cardinoides* populations that showed a constant pattern of insertion sites. Figure 1S (supplementary material), as an example, shows the banding patterns presented by *D. polymorpha* populations when their genomic DNA was digested with the HindIII restriction enzyme, for which the dhMiF2 clone has an internal cleavage site around nucleotide 2,628 (Fig. 1). In this case, HindIII is expected to generate roughly twice the number of bands generated by BglII, which seems to have occurred.

Sequence and phylogenetic analysis

All the analyzed populations presented a unique fragment amplified by primers MIC1777 and MIC2570. The amplicons correspond to the central part of the reverse transcriptase domain of the *micropia* retroelement (Fig. 1). This fragment was amplified in all populations, even using a high annealing temperature (64°C). All *D. cardinoides* and *D. neocardini* populations presented the fragment of the approximately expected size (850 bp), while all the *D. polymorpha* populations presented a smaller fragment (approximately 700 bp). The characterization of the fragments sequenced from different populations from each species (Table 1), using the BlastN program, showed the existence of nearly 97% similarity with dhMiF2, and nearly 83% with Dm11 sequences.

The *D. cardini* group *micropia* sequences were aligned with to *D. hydei* and *D. melanogaster* homologous sequences in order to construct a phylogenetic matrix of 711 nucleotides in length and 14 strains. Of these, only the dhMiF2 and *D. polymorpha micropia* sequences presented stop codons along their nucleotide sequences (one and two

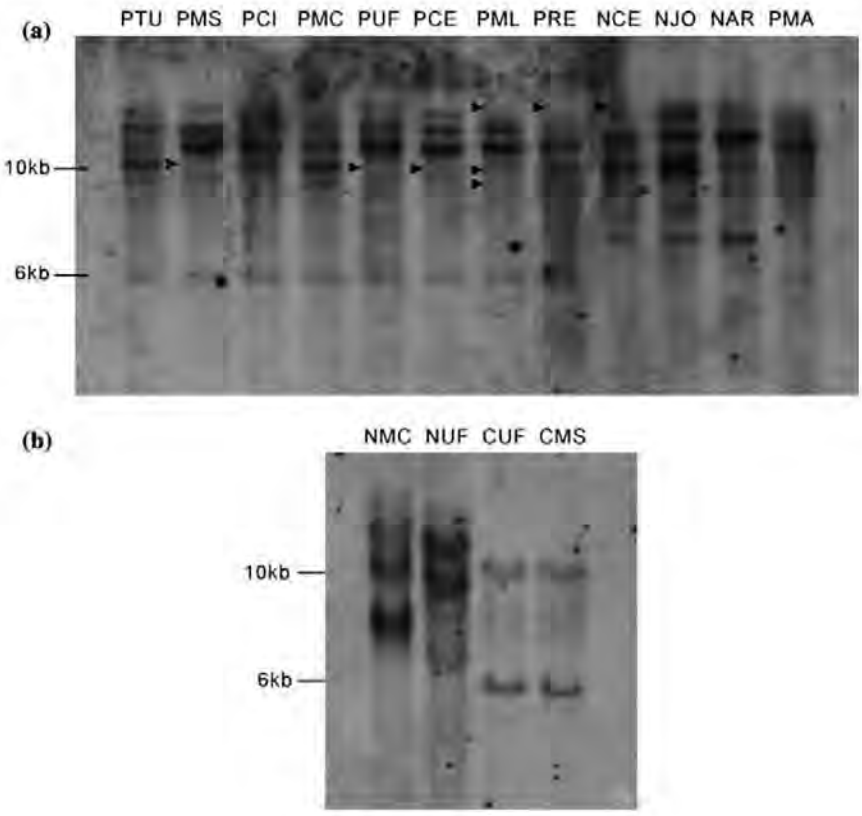


Fig. 2 *Southern* blot hybridization showing *micropia* element insertion site pattern in the genome of different populations of three *D. cardini* group species: *D. cardinoides*, *D. neocardini*, and *D. polymorpha*. The genomic DNA was cleaved using the restriction enzyme BglII, (**a**) and (**b**), and was hybridized using the *micropia* probe. Arrows indicate the interpopulation polymorphic sites. Observe the existence of faint bands. Abbreviations are as given in Table 1

stop codons, respectively). The phylogenetic tree constructed with the NJ analysis for the *micropia* sequences using the Dm11 sequence as outgroup and the bootstrap values obtained by the NJ, MP and ML methods are illustrated in Fig. 3. As can be observed, the three methods produced results in reasonable agreement with each other. The only discordant point concerns the relationships between the species of the *D. cardini* group, which are recovered as politomy by MP and ML, but not by NJ (Fig. 3).

The matrix of divergence for the nucleotide and amino acid sequences for *micropia* and *Amd*, a nuclear gene (see Materials and methods), are presented in Table 2. The comparisons between the *D. cardini* group *micropia* sequences and dhMiF2 suggest the action of a selection pressure higher or equal to that suffered by the nuclear gene (for example, the average nucleotide distance for *micropia* is 0.157 and for *Amd* is 0.215). In the analysis of deduced amino acid sequences, dhMiF2 and *D. polymorpha* presented frameshift, and thus the relationships between their divergences and the divergences of the other species could not be calculated.

In scenarios that suggest the occurrence of HT, shorter divergence values for TE's in relation to nuclear genes may be expected. The occurrence of VT or HT can be tested by comparing the divergence values between the TE that is being analyzed and a nuclear gene (Ludwig and Loreto 2006). In this case, it is possible to infer the occurrence of HT for the TE only when divergence ratio for the TE versus the nuclear gene is significantly below one (<1) (tested by a qui-square test). Figure 4 shows the divergence ratio, d*micropia*/d*Amd*, for each pair of sequences analyzed. The values were significantly below one for the comparisons between dhMiF2 and the *D. cardini* species group, as well as within *D. cardini* species group.

By conducting the selection *Z*-test for each pair of sequences, it was possible to infer the action of a purifying selection over all the comparisons involving dhMiF2 and the sequences of the *D. cardini* group, and in those involving Dm11 and these sequences (at a 0.05 significance level). Yet, for all the comparisons within the *D. cardini* species group, the null hypotheses could not be rejected. This may be a byproduct of the low divergence values presented by these sequences. As expected, the selection *Z*-test strongly indicated a purifying selection pattern for *Amd*.

The d*N*/d*S* ratios for the *micropia* and *Amd* sequences were calculated in order to effectively compare the

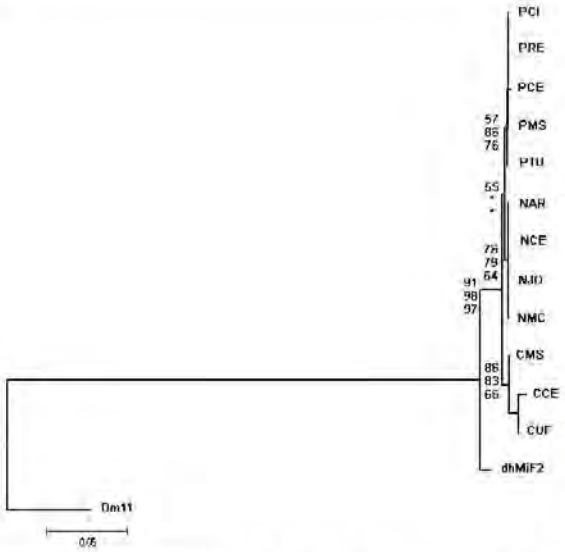


Fig. 3 Neighbor-joining tree of *micropia* retroelement in the *D. cardini* group using the dhMiF2 and Dm11 sequences in the analyses. The tree was rooted using the Dm11 sequence as outgroup. The genetic distances were calculated by the Tamura three-parameter model. The bootstrap values of the nodes are percentages for 1,000 replicates, and are indicated next to the respective branch. Below the bootstrap values for Neighbor-joining there are the bootstrap values for Maximum parsimony and Maximum likelihood, respectively. * Clade not recovered

selection strength acting over the TE and the nuclear gene. The inference of purifying selection is possible when the d*N*/d*S* ratios are under 1. Table 3 reveals that all d*N*/d*S* ratios calculated for *micropia* and *Amd* reach the purifying selection requirement, although the values presented for *Amd* are always under those obtained for *micropia* in the interspecies relationships within the *D. cardini* group, and between these species and dhMiF2/*D. hydei*. For *micropia*, the average values obtained for d*N* and d*S* were 0.051 and 0.331, respectively, which results in d*N*/d*S* = 0.154. On the other hand, for *Amd* the average values obtained for d*N* and d*S* were 0.066 and 0.575, respectively, and therefore d*N*/d*S* = 0.115. Thus, the action of a stronger purifying selection could not offer the best explanation for the lower divergence values obtained for *micropia*, when compared to *Amd*.

Discussion

Ample literature data on *Drosophila* TE's has been produced for the retrotransposons *412* (Hoogland et al. 1997; Cizeron et al. 1998), *1731* (Montchamp-Moreau et al. 1993), *copia* (Vieira et al. 1999; Stacey et al. 1986), *gypsy* (Vieira et al. 1999; Herédia et al. 2004), *micropia* (De Almeida and Carareto 2004; De Almeida et al. 2001; Lankenau 1993), among others. This retrotransposon variability in the genomes of different species has been explained by the presence of these elements in the genome of the ancestral species and/or by the invasion of new genomes by means of HT (Jordan et al. 1999; Herédia et al. 2004; Ludwig and Loreto 2006). Our data add to the existing information on *micropia* representation and evolution within the genus *Drosophila*.

The number of insertion sites inferred for *micropia* in the genomes of the *D. cardini* species group is small as compared to those found for *D. repleta* species group, which is between seven and 17 copies (De Almeida and Carareto 2004), and *D. melanogaster*, which is between 16 and 32 copies (Lankenau 1993). As regards the *D. cardini* group, the estimated number of *micropia* insertions probably ranges from four to six in *D. polymorpha*, *D. neocardini* and *D. cardinoides* (Fig. 2a and b). In this study, the variation within the species might correspond to a small range of insertion site polymorphism for different populations. Although *D. cardinoides* seems to be

Table 2 Comparative analysis between divergence values for *Amd* and *micropia* nucleotide sequences calculated by the Tamura three-parameter model (below and left in the diagonal), and identity/similarity of amino acids for the same sequences (above and right in the diagonal) for the *D. cardini* species group, *D. hydei*, and *D. melanogaster*

	1 dhMiF2/*D. hydei*		2 *D. polymorpha*		3 *D. neocardini*		4 *D. cardinoides*		5 Dm11/*D. melanogaster*	
	Amd	*Micropia*	*Amd*	*Micropia*	*Amd*	*Micropia*	*Amd*	*Micropia*	*Amd*	*Micropia*
1			0.87/0.93	[a]	0.86/0.94	[a]	0.86/0.93	[a]	0.86/0.93	[a]
2	0.272	0.026			0.93/0.96	[a]	0.93/0.95	[a]	0.84/0.90	[a]
3	0.279	0.024	0.082	0.005			0.95/0.96	0.93/0.95	0.83/0.91	0.80/0.87
4	0.260	0.030	0.069	0.013	0.109	0.010			0.83/0.90	0.72/0.80
5	0.313	0.357	0.253	0.364	0.262	0.371	0.247	0.370		

[a] Not calculated. *Drosophila polymorpha* and dhMiF2 *micropia* sequences presented frameshift

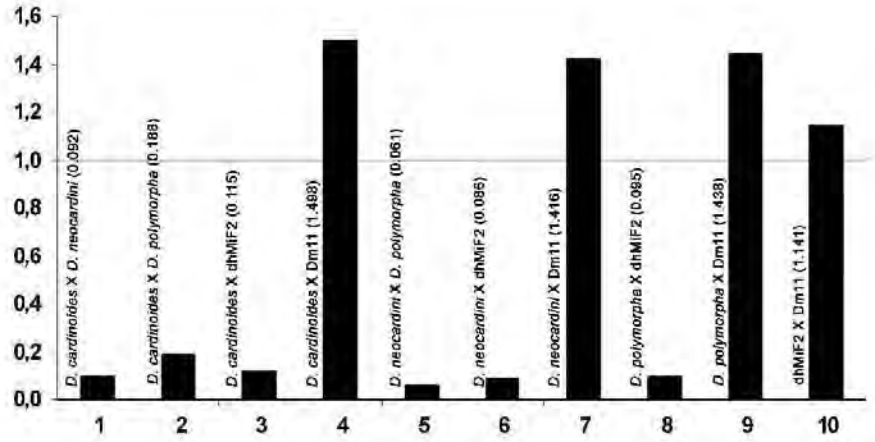


Fig. 4 Comparative analysis between divergence values (Tamura three parameter model) for *micropia* and *Amd* sequences. The horizontal line marks the 1:1 ratio that would be expected if *micropia* evolves at the same rate as *Amd*. Vertical bars represent the ratio of *micropia* distances to *Amd* distances (d*micropia*/d*Amd*) for each pairwise comparison. In all cases, except comparison 10, a significant difference between d*micropia* and d*Amd* was detected at $P < 0.005$ through a qui-square test

Table 3 d*N*/d*S* values for *micropia* (left and below) and *Amd* (right and above) sequences calculated by p-distance model, according to the Nei and Gojobori model

	1 dhMiF2/*D. hydei*	2 *D. polymorpha*	3 *D. neocardini*	4 *D. cardinoides*	5 Dm11/*D. melanogaster*
1	–	0.084	0.111	0.109	0.130
2	0.162	–	0.047	0.048	0.162
3	0.178	0.500	–	0.019	0.164
4	0.267	0.579	0.600	–	0.179
5	0.156	0.164	0.154	0.159	–

monomorphic in relation to its insertion pattern, it is possible that different *micropia* insertion sites would be found when analyzing a greater number of natural populations.

Nevertheless, the TE insertion site number should be interpreted with care, as such results are dependent on conservation of the restriction map used as reference. As the dhMiF2 clone for the *micropia* retroelement does not have a restriction site for BglII, it should be expected that all hybridization bands visualized in the blots might be at least longer than 4.8 kb, as seen in Fig. 2a and b, where all populations presented hybridization bands longer than 5.5 kb. On the other hand, *micropia* presents an internal site for the restriction enzyme HindIII near nucleotide 2,628. For this enzyme, the banding pattern presented by the genomes of *D. polymorpha* were almost the same as expected, which should be twice the number of hybridization bands, when compared with the BglII pattern (Fig. 1S, supplementary material). In both instances, strong hybridization signals were seen right after 1 min film exposure, even using a heterologous probe. These data suggest that the *D. cardini* species group might contain nucleotide sequences that are similar to the *micropia* sequence isolated from *D. hydei*.

In general, although the differences in the *micropia* insertion pattern observed between different populations of *D. neocardini* and *D. polymorpha* were small, they could be attributed to two hypotheses: (a) *micropia* may have been present in the genomes of these species for a long time, with some of the copies having been lost; or (b) *micropia* sequences may have been recently introduced, in which case differential transposition explains the differential pattern. This last hypothesis is supported by the fact that *micropia* is transcriptionally active in *D. repleta* species group (Lankenau et al. 1994; De Almeida and Carareto 2004), being putatively able to transpose. The interspecific variability presented by the *D. cardini* group species, concerning *micropia* hybridization pattern, could also be attributed to recent mobilization events. Our PCR and sequencing data seem to corroborate the hypothesis of a recent invasion by *micropia* in the *D. cardini* group, since there is barely any sequence polymorphism between populations (Fig. 3). Besides that, all *D. polymorpha* populations presented the 133-bp deletion, showing the low degree of *micropia* polymorphism within this species.

Another factual possibility to be addressed is the coexistence of different subfamilies of one TE within the same host genome. Such event might be explained by independent invasions (Silva et al. 2004), and/or by the presence of ancestral polymorphism (Herédia et al. 2004). The presence of two *micropia* subfamilies was hypothesized by De Almeida and Carareto (2004) for 26 species of *D. repleta* group. Our data show a low degree of interpopulation variability concerning the *micropia* sequences (Fig. 3). Nevertheless, we cannot rule out the possibility of

divergent sequence representatives of the *micropia* subfamilies that could not be amplified by the primers used. It is possible that the faint bands observed in Fig. 2 reflect the presence of other *micropia* subfamilies dispersed in the genome.

The sequences amplified in the *D. cardini* group presented a high similarity index to the dhMiF2 sequence (97%), which contributes to the hypothesis that recent invasion could explain the data set here obtained. In addition, this high similarity indicates a *micropia* HT to the genomes of the *D. cardini* species group, maybe from *D. hydei* or other members of the *D. repleta* group or *vice versa*. However, we favor the first hypothesis, since our data suggest recent acquisition of *micropia* within the *D. cardini* species genomes. The lack of sequences from the same region in other species hampers the refinement of these conclusions. De Almeida and Carareto (2004) analyzed *micropia* sequences in the *D. repleta* group that correspond to nucleotides 2,813 to 3,198 different from this study.

The average divergence between dhMiF2 and *D. cardini* group *micropia* was close to 0.027, while between dhMiF2 and Dm11 this distance reached 0.357. However, the divergence values for *Amd* were much higher than those for *micropia*, except for the comparisons involving Dm11 (Table 2). If we were to explain these values admitting *micropia* VT, we would have to infer a higher selective pressure to the TE than to the nuclear gene, which is not expected. Although active elements suffer a selective pressure to maintain its sequences, specially in the reverse transcriptase domain, *Amd* is expected to be even more restricted, since it plays essential roles in the *Drosophila* development and viability, acting from the formation of the eggs' vitelline membrane and embryo cuticle to neurotransmitter synthesis (Wang et al. 1996). Alternatively, we could accept the incongruity of these divergences as evidence that corroborates the HT hypothesis for *micropia* retroelement between the species studied.

Comparing the value of d*N*/d*S*, which measures the type and level of selection acting on the sequences, *micropia* presented an average value of 0.154. This value is significantly higher than that obtained for *Amd* (0.115). So, the hypothesis that *micropia* sequences suffer greater selection pressure than the nuclear gene may now be rejected, reinforcing the HT hypothesis.

Incongruities found between the phylogeny patterns of the host species *versus* that of the TE is another evidence that could favor HT. Phylogenetic relationships in the subgenus *Drosophila* have been inferred by Robe et al. (2005) using two genetic markers, *Amd* and the mitochondrial gene *COII*. In that study, the *D. cardini* group was placed in the *immigrans-tripunctata* radiation, while the *D. repleta* group was positioned in the *virilis-repleta* radiation, showing the distant phylogenetic relationship between these two groups. These relationships are also supported by the findings of Malogolowkin (1953) and Throckmorton (1975). Therefore, the great similarity of the *micropia* sequences between *D. hydei* and the *D. cardini* group should not be expected, and one HT event may be postulated.

Considering the divergence and phylogenetic data, additional HT events might be suggested within the *D. cardini* group. For example, assuming that the phylogeny presented by Brisson et al. (2006) for the *D. cardini* group is accurate and that the *micropia* phylogeny suggested here is correct (Fig. 3), the incongruities and the divergences for *micropia* shown in Table 2 would be explained if we infer two independent HT events within the *D. cardini* group: one from *D. neocardini* to *D. cardinoides*, and another from *D. neocardini* to *D. polymorpha*. Alternatively, these results could reflect three independent HT events between *D. hydei* and each of the three *D. cardini* group species. However, as neither the species phylogeny nor the *micropia* phylogeny is perfectly resolved, the postulated HT events are mainly speculative. As the phylogenies of Heed (1962), Heed and Russel (1971) and Napp and Cordeiro (1981) are in perfect agreement with *micropia* phylogeny other hypothesis such as *Amd* ancestral polymorphism could provide potential explanations for the observed incongruent patterns, concerning its phylogeny and distance values. Considering that the *D. cardini* group is probably a result of rapid radiation (Throckmorton 1975), these data could reflect incomplete lineage sorting for the *Amd* gene in a way that the alleles sampled considerably predate the speciation events. This pattern likely also affect other genes, given the difficulty in the phylogenetic relationships establishment within the *D. cardini* group (see Introduction).

According to Vilela et al. (2002), the *D. cardini* subgroup occurs in Central and South America exclusively, where *D. polymorpha*, *D. neocardini* and *D. cardinoides* are distributed from the south of Central America to southern Brazil. Concerning *D. hydei*, the species originally inhabits the Mexican deserts and has acquired a cosmopolitan status (Val et al. 1981). It is thus possible to state that the likely contact site for the supposed HT event was Central and/or South America, as no occurrence of the *D. cardini* species group is reported outside the area previously mentioned (Bächli 2007). Moreover, the speculative HT events within the *D. cardini* species group may have been facilitated by the occurrence of hybridizations between species (Heed and Russel 1971). Considering the evolution rate calibration as 1.6% divergence per million years ago (Sharp and Li 1989), which is indicated for neutral evolution sequences and using only d*S* values, the HT event between *D. hydei* and *D. cardini*

group might have taken place at least 1.94 million years ago, i.e., well after the divergence of *D. cardini* species subgroup here studied (6.6 million years ago, according to Brisson et al. 2006).

Lately, increasing evidence of LTR retroelement HT events in *Drosophila* has been produced (Jordan et al. 1999; Herédia et al. 2004; De Almeida and Carareto 2006; Ludwig and Loreto 2006). A well characterized study of HT in the genus *Drosophila* addressed the *gypsy* retroelement that apparently is capable to invade genomes as a result of *env* gene expression (Mejlumian et al. 2002), which is not the case of *micropia*. However, mites, bacteria, parasitoid wasps, and even viruses have been appointed as vectors that could mediate retroelement HT (Houck et al. 1991; Heath et al. 1999). *Micropia* apparently has not fallen out of this rule, due to its intrinsic characteristics, to the opportunity generated by the wide geographic distribution of the host species, and to their coexistence within one given ecological niche. Indeed, the results presented here may be a first step towards the understanding of the *micropia* evolutionary pattern. Future studies including more species will improve this scenario.

Acknowledgements This work was supported by grants from CAPES, CNPq, and FAPERGS research funding. The authors thank MSc Adriana Ludwig for statistical support and discussion, Daniel F. Dombrosky for valuable help and the members of the *Drosophila* Laboratories from Universidade Federal do Rio Grande do Sul, Universidade Federal de Santa Catarina and Universidade Federal de Santa Maria. We thank also the two anonymous reviewers' valuable suggestions, Dr. Claudia M. A. Carareto and Dr. Dirk-Henner Lankenau for kindly providing the dhMiF2 plasmid, and Félix Nonnemacher for English revision.

References

Bächli G (2007) Taxodros database 2006/08: The database on taxonomy of Drosophilidae. URL: http://www.taxodros.unizh.ch/. Last accessed in April 2007

Brisson JA, Wilder J, Hollocher H (2006) Phylogenetic analysis of the *cardini* group of *Drosophila* with respect to changes in pigmentation. Evolution 60:1228–1241

Capy P, Bazin C, Higuet D, Langin T (1997) Dynamics and evolution of transposable elements. Landes Bioscience, Texas

Cizeron G, Lemeunier F, Loevenbruck C, Brehm A, Biémont C (1998) Distribution of the retrotransposable element *412* in *Drosophila* species. Mol Biol Evol 15:1589–1599

Da Cunha AB, Brncic D, Salzano FM (1953) A comparative study of chromosomal polymorphism in certain South American species of *Drosophila*. Heredity 2:193–202

Daniels SB, Chovnick A, Boussy IA (1990) Distribution of *hobo* transposable elements in the genus *Drosophila*. Mol Biol Evol 7:589–606

De Almeida LM, Carareto CMA (2004) Identification of two subfamilies of *micropia* transposable element in species of the *repleta* group of *Drosophila*. Genetica 121:155–164

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

De Almeida LM, Castro JP, Carareto CMA (2001) *Micropia* transposable element occurrence in Drosophila species of the saltans group. DIS 84:114–117

De Toni DC, Herédia FO, Valente VLS (2001) Chromosomal variability of *Drosophila polymorpha* populations from Atlantic Forest remnants of continental and insular environments in the State of Santa Catarina, Brazil. Caryologia 4:329–337

Felsenstein J (1985) Confidence limits on phylogenies: an approaching using bootstrap. Evolution 39:783–791

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Heath BD, Butcher RD, Whitfield WG, Hubbard SF (1999) Horizontal transfer of *Wolbachia* between phylogenetically distant insect species by a naturally occurring mechanism. Curr Biol 9:313–316

Heed WB (1962) Genetic characteristics of island populations. Univ Texas Publ Stud Genet 6205:173–206

Heed WB (1963) Density and distribution of *Drosophila polymorpha* and its color alleles in south America. Evolution 17:502–518

Heed WB, Blake PR (1963) A new color allele at the *e* locus of *Drosophila polymorpha* from northern south America. Genetics 48:217–234

Heed WB, Krishnamurthy ND (1959) Genetic studies on the *cardini* group of *Drosophila* in the West Indies. Univ Texas Publs 5914:155–179

Heed WB, Russel JS (1971) Phylogeny and population structure in island and continental species of the *cardini* group of *Drosophila* studied by inversion analysis. Univ Texas Publs 7103:91–130

Hendy MD, Penny D (1982) Branch and bound algorithms to determine minimal evolution trees. Math Biosci 59:277–290

Herédia FO, Loreto ELS, Valente VLS (2004) Complex evolution of *gypsy* in drosophilid species. Mol Biol Evol 21:1–12

Hollocher H, Hatcher JL, Dyreson EG (2000a) Evolution of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. Evolution 54:2046–2056

Hollocher H, Hatcher JL, Dyreson EG (2000b) Genetic and developmental analysis of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. Evolution 54:2057–2071

Hoogland C, Vieira C, Biémont C (1997) Chromosomal distribution of the *412* retrotransposon in natural populations of *Drosophila simulans*. Heredity 79:128–134

Houck MA, Clark JB, Peterson HR, Kidwell MG (1991) Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. Science 253:1125–1128

Huijser P, Kirchhoff C, Lankenau D-H, Hennig W (1988) Retrotransposon-like sequences are expressed in Y chromosomal lampbrush loops of *Drosophila hydei*. J Mol Biol 203:689–697

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with ClustalX. Trends Biochem Sci 23:403–405

Jordan IK, Maayunina LV, McDonald JF (1999) Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. PNAS 22:12621–12625

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform 5:150–163

Lankenau D-H (1993) The retrotransposon family *micropia* in *Drosophila* species. In: McDonald J (eds) Transposable elements and evolution. Kluwer Publishers, Amsterdam pp 232–241

Lankenau D-H, Huijser P, Hennig W (1989) Characterization of the long terminal repeats of *micropia* elements microdissected from

the Y-chromosomal lampbrush loops "threads" of *Drosophila hydei*. J Mol Biol 209:493–497

Lankenau D-H, Huijser P, Jansen E, Miedema K, Hennig W (1988) *micropia*: a retrotransposon of *Drosophila* combining structural features of DNA viruses, retroviruses and non-viral transposable elements. J Mol Biol 2:233–246

Lankenau D-H, Huijser P, Jansen E, Miedema K, Hennig W (1990) DNA sequence comparison of *micropia* transposable elements from *Drosophila hydei* and *Drosophila melanogaster*. Chromosoma 99:111–117

Lankenau S, Corces GV, Lankenau D-H (1994) The *Drosophila micropia* retrotransposon encodes a testis-specific antisense RNA complementary to reverse transcriptase. Mol Biol Evol 17:1542–1557

Loreto ELS, Basso da Silva L, Zaha A, Valente VLS (1998) Distribution of transposable elements in Neotropical species of *Drosophila*. Genetica 101:153–165

Ludwig A, Loreto ELS (2006) Evolutionary pattern of the *gtwin* retrotransposon in the *Drosophila melanogaster* subgroup. Genetica 130:161–168

Malogolowkin C (1953) Sobre a genitália dos drosofilídeos. IV. A genitália masculina no subgênero *Drosophila* (Diptera, Drosophilidae). Rev Bras Biol 13:245–264

Mejlumian L, Pelisson A, Bucheton A, Terzian C (2002) Comparative and functional studies of *Drosophila* species invasion by the *gypsy* endogenous retrovirus. Genetics 160:201–209

Montchamp-Moreau C, Ronsseray M, Jacques M, Lehmann M, Anxolabéhère D (1993) Distribution and conservation of sequences homologous to the *1731* retrotransposon in *Drosophila*. Mol Biol Evol 10:791–803

Napp M, Cordeiro AR (1981) Interspecific relationships in the *cardini* group of *Drosophila* studied by electrophoresis. Revta Bras Genet 4:537–547

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Nicholas KB, Nicholas HB (1997) GeneDoc: a tool for editing and annotations of multiple sequence alignements. Distributed by the author. http://www.psc.edu/biomed/genedoc

Posada C, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818

Robe LJ, Valente VLS, Budnik M, Loreto ELS (2005) Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: A nuclear versus mitochondrial gene approach. Mol Phylogenet Evol 3:623–640

Rohde C, Valente VLS (1996) Cytological maps and chromosomal polymorphism of *Drosophila polymorpha* and *Drosophila cardinoides*. Braz J Genet 19:27–32

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

Sassi AK, Herédia FO, Loreto ELS, Valente VLS, Rohde C (2005) Transposable elements *P* and *gypsy* in natural populations of *Drosophila willistoni*. Genet Mol Biol 28:734–739

Sharp PM, Li W-H (1989) On the rate of DNA sequence evolution in *Drosophila*. J Mol Evol 28:398–402

Silva JC, Loreto ELS, Clark JB (2004) Factors that affect the horizontal transfer of transposable elements. Curr Issues Mol Biol 6:57–72

Stacey SN, Lansman RA, Brock HW Grigliatti TA (1986) Distribution and conservation of mobile elements in the genus *Drosophila*. Mol Biol Evol 3:522–534

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

Swofford DL (2003) PAUP*. Phylogenetic analysis using parsimony (*and other methods), Version 4. Sinauer Associates, Sunderland, Massachusetts

Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C biases. Mol Biol Evol 9:678–687

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

Val FC, Vilela CR, Marques MD (1981) Drosophilidae of the Neotropical region. In: Ashburner M, Carson HL, Thompson JN Jr (eds) The genetics and biology of *Drosophila*. 3a. Academic Press, London, pp 123–168

Vieira C, Lepetit D, Dumont S, Biémont C (1999) Wake up of transposable elements following *Drosophila simulans* worldwide colonization. Mol Biol Evol 16:1251–1255

Vilela CR, da Silva AFG, Sene FM (2002) Preliminary data on the geographical distribution of *Drosophila* species within morpho-climatic domains of Brazil. Rev Bras Entomol 2:139–148

Wang D, Marsh JL, Ayala FJ (1996) Evolutionary changes in the expression pattern of a developmentally essential gene in three *Drosophila* species. Proc Natl Acad Sci USA 93:7103–7107

Wilder JA, Hollocher H (2001) Mobile elements and the genesis of microsatellites in dipterans. Mol Biol Evol 18:384–392

CAPÍTULO III

Trabalho a ser submetido à revista científica Journal of Molecular Evolution

Evolutionary history and classification of *micropia* elements in the *Drosophila* genus

Juliana Cordeiro[1]; Lizandra Jaqueline Robe[2]; Francine Cenzi de Ré[3]; Elgion Lucio da Silva Loreto[1, 2, 3]; Vera Lucia da Silva Valente[1, 2, 4]

[1] Programa de Pós Graduação em Genética e Biologia Molecular; Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

[2] Programa de Pós-Graduação em Biodiversidade Animal, Departamento de Biologia Universidade Federal de Santa Maria, Santa Maria, RS, Brazil

[3] Curso de Ciências Biológicas, Departamento de Biologia, Centro de Ciências Naturais e Exatas, Universidade Federal de Santa Maria, Santa Maria, RS, Brazil

[4] Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul; Porto Alegre; Rio Grande do Sul; Brazil

Corresponding author: Juliana Cordeiro, Departamento de Genética, UFRGS, Caixa Postal 15053, CEP: 91501-970. Phone: +55-51-33166729; Fax: +55-51-33167311; e-mail: jlncdr@gmail.com

Abstract

This study analyses the distribution and the evolutionary history of the *micropia* retrotransposon in species from the *cardini* and *repleta* groups, as also in the 12 *Drosophila* genomes available at Flybase site. A phylogenetic study was performed with the 71 *micropia* related sequences found on the genomes of *D. sechellia*, *D. melanogaster*, *D. simulans*, *D. erecta*, *D. yakuba* and *D. willistoni*; including other 33 sequences found on the *cardini* and *repleta* species groups. Using a series of criteria to support horizontal transfer inferences (as incongruence between the transposable element and host phylogeny, discontinuous distribution of the element, comparison of the synonymous divergences between the transposable element and a nuclear host gene, among other considerations) we identified, at least, 17 horizontal transfer events between the species. Based on nucleotide and amino acid sequence similarities, we also identified the existence of three different families of *micropia* related sequences in the *Drosophila* genus. The evolutionary history of these sequences probably includes the effects of ancestral polymorphism, vertical and horizontal transmission, with some of these last possibly related to introgression mechanisms.

Running title: *micropia* element in the *Drosophila* genus

Key words: *micropia* classification – *Drosophila* genus – *cardini* group – *repleta* group – horizontal transfer

Introduction

Transposable elements (TE) are mobile sequences, found in practically all life forms, that can spread within a genome or even between genomes, acting like "intragenomic parasites" (Craig et al. 2002). Despite the significant host fitness cost generated by the mobilization of these elements, several authors have explained their abundance in these genomes by means of their participation on fundamental features such as chromatin organization, chromosome structure, genome size, regulatory functions and coevolutionary processes (Capy et al. 1998; Cáceres et al. 1999; Kidwell and Lisch 2001; Pooma et al. 2002; Vieira et al. 2002; Lippman et al. 2004; Puig et al. 2004). However, the abundance of TEs in a host genome seems to be determined by their intrinsic ability to replicate by transposition and to keep out of extinction (Kidwell and Lisch 2001; Silva et al. 2004; Loreto et al. 2007).

One of the most notable TE properties is their capability to cross species boundaries, invading new genomes, by means of horizontal transfer (HT) process. It was shown by several authors that this event is more common than ever thought (review in Loreto et al. 2007), becoming an important step in the TEs' "life cycle" (Kidwell and Lisch 2001), and therefore, to the host genome evolution. However, considering all the interactions between TEs and the host genome, the evolutionary history of these mobile elements is considerably more complex than suggested by the classification attempts (Finnegan 1989; Capy et al. 1998; Wicker et al. 2007).

Recently, it was shown that the long terminal repeats (LTR) retrotransposons perform approximately 90% of the HT events detected across the *Drosophila melanogaster*, *D. simulans* and *D. yakuba* genomes (Bartolomé et al. 2009); and the *micropia* retroelement is among these elements. The *micropia* TE was discovery by chance

in the lampbrush loops of the *Drosophila hydei* Y chromosomes (Huijser et al. 1988; Lankenau 1993). Concerning molecular characteristics, there are four best characterized *micropia* elements (*micropia*-dhMiF2 and *micropia*-dhMiF8, both extracted from the *D. hydei* genome, and *micropia*-Dm11 and *micropia*-Dm2, both from the *D. melanogaster* genome) (Huijser et al. 1988; Lankenau et al. 1988; 1990), which positioned this family as belonging to the *Ty*3 superfamily of retrotransposable elements with LTRs (Lankenau 1993; Capy et al. 1988). Sequences similar to *micropia* were identified in the genome of species from the *Sophophora* subgenus (*melanogaster*, *willistoni*, and *saltans* groups), and the *Drosophila* subgenus (*cardini*, *immigrans*, *funebris* and *repleta* group) (Lankenau 1993; Almeida et al. 2001; Almeida and Carareto 2004; Cordeiro et al. 2008). Almeida and Carareto (2004) identified two subfamilies of *micropia* in the genomes of some *repleta* group species. These findings, together with the previous data, suggest a heritance based on vertical transfer (VT) of this element in the *Drosophila* genus. However, recently Cordeiro et al. (2008) found evidences in the *cardini* group that strongly suggests a HT of *micropia* between *D. hydei* and *D. neocardini*, showing that the evolutionary pattern of *micropia* is likely not only based on VT.

Previous works showed that *micropia* expression gives rise to a complex set of sense and antisense RNAs (Lankenau 1993; Lankenau et al. 1994; Almeida and Carareto 2004), and the principal transcript is a 5kb full-length sense product expressed in both somatic and germ line tissues of males and females of *D. hydei*. In addition, *micropia* encodes two different antisense RNAs, 1kb and 1.6kb long, which have their expression limited to the testis. These antisense transcripts, produced by the copies present on Y chromosome, are potentially directed against the entire RNaseH and some domains of the reverse transcriptase, contributing to the *micropia* full-length RNA degradation and/or to

translational interference (Lankenau 1993; Lankenau et al. 1994; Almeida and Carareto 2004). The transposing activity of this retrotransposon family might be quite effectively repressed by the mechanism above described; in a way that *micropia* has never been observed to cause any mutation (Lankenau 1993; Almeida and Carareto 2004). However, the *micropia* elements isolated from *D. melanogaster* does not seem to express such RNA repression mechanism (Lankenau 1993).

This study analyses 104 *micropia* related sequences obtained from eight species of the *cardini* group, four species from the *repleta* group and six species from the 12 available *Drosophila* genomes. Here we delineate the most likely evolutionary history of these sequences on the *Drosophila* genus, possibly involving three different families of *micropia* that underwent a series of HT events.

Materials and methods

Species and sequences

A total of 41 *Drosophila* species were analysed in this study. Sixteen species were utilyzed in the *in silico* search and 25 species in the *in vivo* search. Table 1 and the Supplementary Material Table 1S lists the species used in the *in vivo* study, with their collection sites coordinates or Stock Center number. The *in silico* search recovered *micropia* sequences that were previously deposited in GenBank (Benson et al. 2008) for *D. cardinoides* (EF090263), *D. neocardini* (EF090264), *D. polymorpha* (EF090265), *D. hydei* (dhMiF2 sequence - X13304) and *D. melanogaster* (Dm11 sequence - X14037). Besides, a search was also conducted in the 12 available *Drosophila* genomes (*Drosophila* 12 Genomes Consortium 2007).

Table 1: *Drosophila* species investigated in this study with their respective source and GenBank *micropia* accession numbers.

Group	Species	Coordinate or Stock Center number	Localization	GenBank Acc. Nos.
cardini	*D. cardini*[A]	23º14'S 46º58'W	Serra do Japi-SP-BR	FJ748690
				FJ748691
				FJ748692
	D. neomorpha[B]	26º17'S 49º00'W	Joinville-SC-BR	FJ748695
				FJ748696
				FJ748697
	D. parthenogenetica	27º35'S 48º22'W	Florianopolis-SC-BR	FJ748698
				FJ748699
				GQ339587
				GQ339588
				GQ339589
				GQ339590
	D. procardinoides	15181-2241.00	Coroico-BR	FJ748700
				FJ748701
				FJ748702
	D. acutilabella	15181-2171.09	Everglades-Florida-USA	FJ748684
				FJ748685
				FJ748686
				FJ748687
				FJ748688
repleta	*D. buzzatii*	37º19'S 59º09'W	Tandil-AR	FJ748689
				GQ339579
				GQ339580
				GQ339581
				GQ339582
	D. mercatorum[C]	27º35'S 48º22'W	Florianopolis-SC-BR	FJ748693
				GQ339583
				GQ339584
				GQ339585
				GQ339586
				FJ748694
	D. zottii[D]	26º17'S 49º00'W	Joinville-SC-BR	FJ748703
				GQ339578

Capital letters refer to the fly collectors: [A]Dra Claudia Rohde, [B]Dra Daniela Cristina De Toni, [C]Dr Marco Silva Gottschalk, [D]MSc Jonas da Silva Doge.

PCR, Dot Blot, cloning and sequencing

Genomic DNA from each of the 25 species analysed *in vivo* was prepared according to Sassi et al. (2005). Polymerase Chain Reactions (PCR) were performed with the primers described in Cordeiro et al. (2008). The following conditions for a 25µl PCR reaction were used: 25ng template DNA, 20pMol of each primer, 0.2mM of each nucleotide, 1.5mM $MgCl_2$ and 1 unit Taq DNA polymerase in 1x Polymerase Buffer (all from Invitrogen). Amplifications parameters were 95°C for 2min, 35 cycles at 95°C for 30s, 57°C for 30s and 72°C for 1min, followed by an extension cycle at 72°C for 10min. *Drosophila hydei* genomic DNA was used as positive control. The above described annealing temperature varied, nevertheless, between 50°C and 60°C in an attempt to amplify the *micropia* element in all species. In order to confirm the homology of the amplified fragments, a Southern Blot was conducted with the obtained amplicons. The PCR products were eletrophorezed using a 1% agarose gel and transferred to nylon membranes (Hybond N+®, GE Healthcare), as in Southern Blot methodology. The 812bp PCR *micropia* fragment from *D. hydei* was used as probe (from nucleotide 1,777 to 2,589 of dhMiF2 sequence) (see Figure 1 in Cordeiro et al. 2008). The probe was labeled and the signal was detected using the *Gene Images^{TM} AlkPhos Direct^{TM} labelling and detection system* (GE Healthcare), according to manufacturer's instructions. The membranes were hybridized at 55°C and exposed to the revealing film for 5 minutes. In the Dot Blot methodology, for denaturation, DNA samples (3µg in 10ul) were boiled for five minutes, freezed for other five minutes and applied onto a nylon membrane (Hybond N+®, GE Healthcare). As positive control, we used 5ng (in 10ul) of the *micropia* probe. The probe sequence, the labeling of the probe, the signal detection and the hybridization temperature were performed as above mentioned. The revealing film underwent 3 minutes exposure.

To isolate the amplified fragments, the amplicons were separated by 1.5% agarose gel electrophoresis and purified using *illustra GFX™ PCR DNA and Gel Band Purification kit* (GE Healthcare) according to the supplier's specifications. The fragments were cloned using the *TOPO-TA Cloning® Vector* system (Invitrogen) or *pGEM®-T Easy Vector* system (Promega). The obtained recombinant plasmids underwent a new PCR reaction using the universal M13 primers at a 55°C annealing temperature. The amplicons so obtained were purified by incubations with Exonuclease I (ExoI) and Shrimp Alkaline Phosphatase (SAP) (both from GE Healthcare) and directly sequenced in a MegaBACE™500 (GE Healthcare) automatic sequencer. In all cases, both strands were sequenced, and ambiguities and compressions were resolved through assemblage in the Staden Package Gap 4 program (Staden 1996). The GenBank accession numbers at National Center for Biotecnology Information website (NCBI – http://www.ncbi.nlm.nih.gov/) of the obtained sequences are indicated in Table 1.

Genome search

Initially, a BLAST search was performed in each of the 12 *Drosophila* genomes (*Drosophila* 12 Genomes Consortium 2007) available in the Flybase website (http://flybase.org/blast/ - Last accessed on February 2009) searching in the up to date releases and using the default parameters. The queries consisted of a 722bp fragment from the reverse transcriptase region of dhMiF2, Dm11 and each of the fragments sequenced in this work (see Table 1). The retrieved sequences matched the following criteria: more than 400bp length; scores > 200; and *E*value < $1e^{-40}$. For each chosen sequence 2kb from both side of the hit were saved for the analysis. Table 2S (see Supplementary Material) lists the species from which *micropia* sequences were obtained in the *in silico* search.

Phylogenetic analysis, codon usage bias and dS estimates

The sequences obtained according to the described methodologies were aligned using ClustalW (Jeanmougin et al. 1998) as contained in MEGA4 program (Tamura et al. 2007). The evolutionary model that best fits to the data was searched through the Akaike Information Criterion test (AIC test) (Akaike 1974) as performed by PAUP4.0b10 (Swofford 2003) and ModelTest3.7 (Posada and Crandall 1998) or MrModelTest2.2 (Nylander 2004), in which cases the HKY model (Hasegawa et al., 1985) with a gamma correction (α = 2.1450) was chosen The sequences were analyzed by four phylogenetic methods: (1) Neighbor-Joining (NJ) (Saitou and Nei 1987), using MEGA4 program, following the Tamura-3 parameters model (Tamura 1992) with gamma correction. This model was chosen because it is the one available in MEGA4 that is more similar with the HKY+G model proposed by the AIC test and ModelTest3.7 program; (2) Maximum Parsimony (MP) in PAUP4.0b10, through an heuristic search with Nearest-neighbor interchange (NNI) applied to 100 random stepwise addition trees, with maxtrees set to 20,000; (3) Maximum Likelihood (ML) in PhyML2.4.4 (Gindon and Gascuel 2003) according to the HKY + G evolutionary model (Hasegawa et al. 1985); and (4) Bayesian inferences in MrBayes3.1.2 (Huelsenbeck and Ronquist 2001) with the trees saved every 100 of the 3,000,000 generations of Markow Chain Monte Carlo (MCMC) and a burnin of 7,500, following the HKY + G model. In the NJ, MP and ML analyses, bootstrap tests (Felsenstein 1985) with 1,000 replications were utilized to measure the support of the clades. For the Bayesian analyses the posterior probability (PP) of each clade on the 50% majority rule consensus tree was calculated (Hall, 2001). In all the four phylogenetic analysis, a sequence related to reverse transcriptase of a Bracovirus from the Hymenoptera *Glyptapentelis flavicoxis* (EF710649) was utilized as outgroup because a BLASTx search,

executed using dhMiF2 and Dm11 as query, recovered it as the best external hit.

The nucleotides and amino acid divergence matrices were constructed using MEGA4 program. The Tamura 3-parameters substitution model with gamma correction (α=2.1450) was used to calculate the divergence between the nucleotide sequences. The JTT (Jones et al. 1992) amino acid substitution model with gamma correction (α=3.128) was used to calculate the divergence between the amino acid sequences, since it is the MEGA4 model that better resembles the LG model (Le and Gascuel 2008) indicated by ProtTest2.1 program (Abascal et al. 2005). To estimate the number of synonymous substitution per synonymous sites (dS) values and the number of synonymous sites (S) we used Nei-Gojobori equation (1986) as computed in MEGA4 software using p-distance. Based on the MB tree, the sequences that fit the following criterion were grouped to calculate the dS values: monophyletic sequences found in the genome of the same species presenting a PP > 0.70. The average number of synonymous and nonsynonymous substitution per codon of the aligned nucleotide sequence were calculated by the SNAP program (Korber 2000) available at http://www.hiv.lanl.gov/. The effective number of codon (ENC; Wright 1990) and the codon bias index (CBI; Bennetzen and Hall 1982) were calculated in the DnaSP5.0 software (Librado and Rozas 2009).

The alpha–metildopa (*Amd*) nuclear gene was used in the comparisons of the dS values with *micropia* because it is the only nuclear gene available for all species here analyzed. Moreover, among different nuclear genes, *Amd* was shown to present one of the lowest codon usage bias (Vidal et al. 2009) which is a pre-requisite to avoid HT overestimation. The *Amd* sequences of the species were obtained from the NCBI website (accession numbers in GenBank: AY699258 for *D. cardinoides*, AY699260 for *D. neocardini*, AY699259 for *D. polymorpha*, AF324947 for *D. buzzatii*, AF293712 for *D.*

hydei, AF293730 for *D. willistoni*, X04695 for *D. melanogaster*); and from the species genomes (XM_01974256 for *D. erecta*, XM_002039039 for *D. sechellia*, XM_002079852 for *D. simulans* and XM_002090630 for *D. yakuba*). For *D. zottii*, *D. cardini*, *D. acutilabella*, *D. neomorpha*, *D. parthenogenetica* and *D. procardinoides* the *Amd* sequences were obtained by one of the authors (FCR).

Results

Discontinuous distribution of *micropia* in *Drosophila* genus

We have analyzed the presence of *micropia* related sequences in the genome of *Drosophila* species in three ways: by *in silico* search, and through PCR and Dot Blot. The *in silico* search showed that *micropia* is present in the genome of: *D. erecta*, *D. melanogaster*, *D. sechellia*, *D. simulans*, *D. yakuba* (all from the *melanogaster* subgroup) and *D. willistoni* (*willistoni* group); but absent in the genomes of: *D. ananassae* (*melanogaster* group), *D. pseudobscura*, *D. persimilis* (both from the *obscura* group), *D. virilis* (*virilis* group), *D. grimshawi* (picture wing group) and *D. mojavensis* (*repleta* group). Supplementary Material Table 2S summarizes the main information about the sequences found in the *in silico* search. This discontinuous distribution was also confirmed by the PCR and Dot Blot results, as can be seen in the Supplementary Material Table 1S and Figure 1S. We only were able to amplify, clone and sequence *micropia* copies present in the eight species shown in Table 1, despite the fact that we also tested 17 other species (Table 2S). The *cardini* group showed an interesting distribution pattern, where only the mainland species (from Florida, USA to south Brazil) (Heed and Russel 1971) have *micropia* related sequences in their genome. The other four species; which seem to be

devoid of *micropia* (Figure 1S); belong to the *dunni* subgroup, endemic to the Caribbean islands (Heed and Russel 1971).

Phylogenetic analysis

For the *micropia* phylogenetic analysis, we used a total of 104 sequences: 66 obtained through genome search; five from GenBank; and 33 from *in vivo* search in species from the *cardini* and *repleta* groups. From all of these, 58 sequences are putatively active, maintaining its encoding potential (55.77% of the total number of the sequences) (see Figure 1). We analyzed these sequences under four phylogenetic methods, which showed similar topologies. In Figure 1 the Bayesian phylogenetic reconstruction is shown, with the posterior probability indicated in the branches, whereas in Table 2 the bootstrap values for the recovered clades in each of the other three phylogenetic methods are presented.

In general, two main clades were recovered: clade 1, grouping the 61 sequences from the *melanogaster* subgroup (29 from *D. sechellia*, 12 from *D. yakuba*, 10 from *D. simulans*, nine from *D. melanogaster*, and the unique sequence from *D. erecta*); clade 2, clustering the sequences found in the four studied members from the *repleta* group (*D. hydei* with one sequence, *D. zottii* with two sequences, *D. buzzatii* with five sequences and *D. mercatorum* with six sequences), with those found in the eight species from the *cardini* group (*D. neocardini*, *D. polymorpha* and *D. cardinoides* each with one sequence; *D. cardini*, *D. neomorpha* and *D. procardinoides* with three sequences; *D. acutilabella* with five sequences; *D. parthenogenetica* with six sequences), and with those encountered in the *D. willistoni* genome (six sequences). Within each of these clades contrasting patterns

emerge, in a way that in the clade 1 the sequences recovered from the different species are mainly intermingled, whereas in the clade 2 each species sequences appear mostly clustered. Exceptions are provided, nevertheless, by *D. yakuba* in clade 1 and *D. parthenogenetica*, *D. cardini* and *D. willistoni* in clade 2. Although, in both clusters a general subdivision can be envisioned (see clades 3 and 4 within main clade 1, and clades 5 and 6 within main clade 2), polytomic clades detach as the prevalent pattern, rendering the obtained phylogeny incongruent with the phylogeny of the species (Figure 2).

Figure 1 (see next page): Phylogenetic reconstruction obtained using Bayesian analysis and the HKY+G model, as proposed by the AIC test for the 104 *micropia* related nucleotide sequences. The posterior probability of each clade is indicated near its respective internal branch; and under each clade there is an arbitrarily defined number representing the clade that follows and which can be used to interpret Table 2 and information in the text. See legend in the figure for identification about potential coding sequences and about the taxonomic grouping of the species from which the sequences belong.

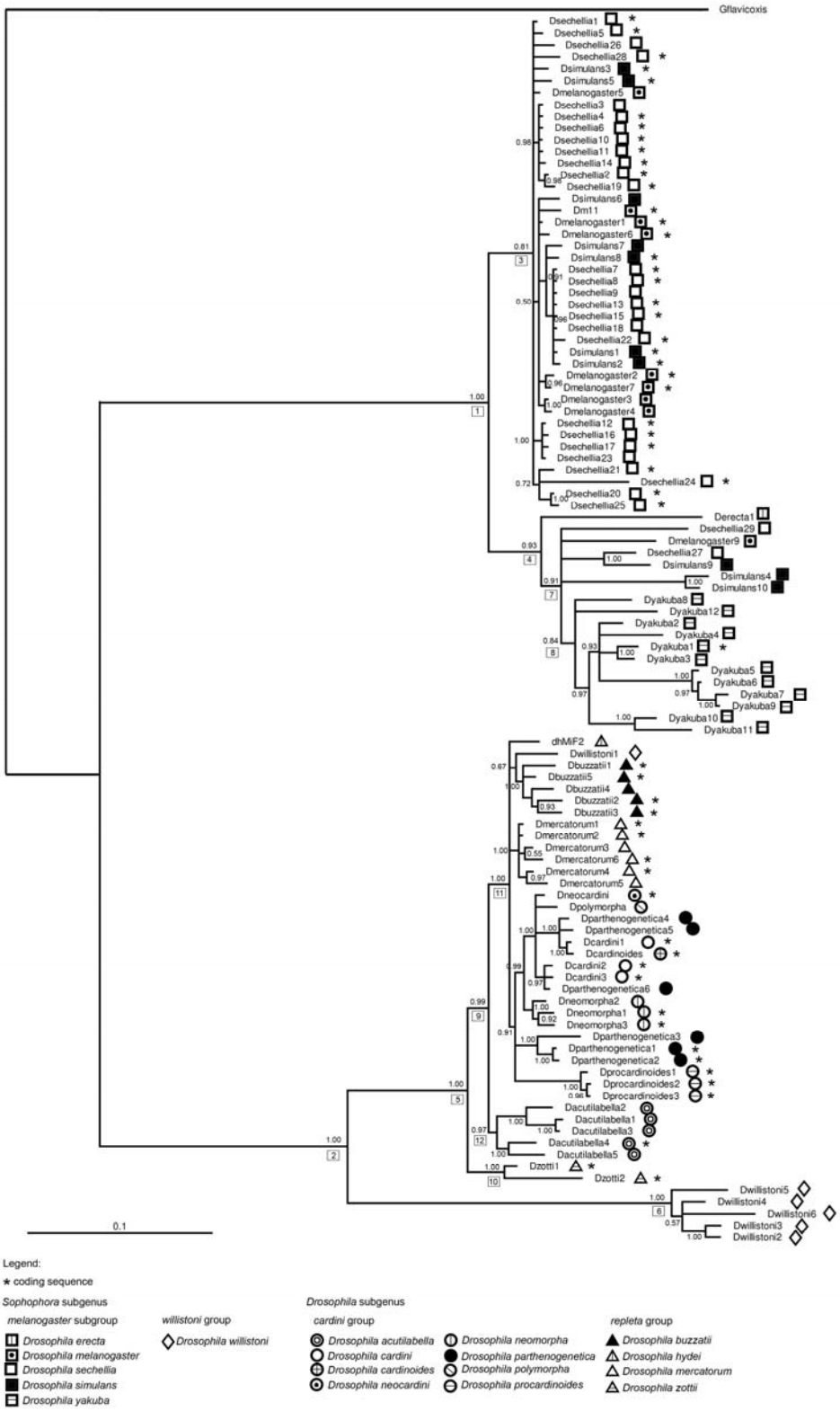

Gflavicoxis
Dsechellia1
Dsechellia5
Dsechellia26
Dsechellia28
Dsimulans3
Dsimulans5
Dmelanogaster5
Dsechellia3
Dsechellia4
Dsechellia6
Dsechellia10
Dsechellia11
Dsechellia14
Dsechellia2
Dsechellia19
Dsimulans6
Dm11
Dmelanogaster1
Dmelanogaster6
Dsimulans7
Dsimulans8
Dsechellia7
Dsechellia8
Dsechellia9
Dsechellia13
Dsechellia15
Dsechellia18
Dsechellia22
Dsimulans1
Dsimulans2
Dmelanogaster2
Dmelanogaster7
Dmelanogaster3
Dmelanogaster4
Dsechellia12
Dsechellia16
Dsechellia17
Dsechellia23
Dsechellia21
Dsechellia24
Dsechellia20
Dsechellia25
Derecta1
Dsechellia29
Dmelanogaster9
Dsechellia27
Dsimulans9
Dsimulans4
Dsimulans10
Dyakuba8
Dyakuba12
Dyakuba2
Dyakuba4
Dyakuba1
Dyakuba3
Dyakuba5
Dyakuba6
Dyakuba7
Dyakuba9
Dyakuba10
Dyakuba11
dhMiF2
Dwillistoni1
Dbuzzatii1
Dbuzzatii5
Dbuzzatii4
Dbuzzatii2
Dbuzzatii3
Dmercatorum1
Dmercatorum2
Dmercatorum3
Dmercatorum6
Dmercatorum4
Dmercatorum5
Dneocardini
Dpolymorpha
Dparthenogenetica4
Dparthenogenetica5
Dcardini1
Dcardinoides
Dcardini2
Dcardini3
Dparthenogenetica6
Dneomorpha2
Dneomorpha1
Dneomorpha3
Dparthenogenetica3
Dparthenogenetica1
Dparthenogenetica2
Dprocardinoides1
Dprocardinoides2
Dprocardinoides3
Dacutilabella2
Dacutilabella1
Dacutilabella3
Dacutilabella4
Dacutilabella5
Dzotti1
Dzotti2
Dwillistoni5
Dwillistoni4
Dwillistoni6
Dwillistoni3
Dwillistoni2
0.1
Legend:
* coding sequence
Sophophora subgenus
melanogaster subgroup
Drosophila erecta
Drosophila melanogaster
Drosophila sechellia
Drosophila simulans
Drosophila yakuba
willistoni group
Drosophila willistoni
Drosophila subgenus
cardini group
Drosophila acutilabella
Drosophila cardini
Drosophila cardinoides
Drosophila neocardini
Drosophila neomorpha
Drosophila parthenogenetica
Drosophila polymorpha
Drosophila procardinoides
repleta group
Drosophila buzzatii
Drosophila hydei
Drosophila mercatorum
Drosophila zottii

Table 2: Bootstrap values presented for each of the clades shown in Figure 1 by the trees reconstructed using different phylogenetic reconstruction methods.

Clades	Phylogenetic reconstruction methods		
	NJ	ML	MP
1	99	100	100
2	97	992	99
3	--	837	94
4	96	932	63
5	82	995	100
6	99	100	100
7	--	72	74
8	--	40	--
9	--	--	75
10	--	--	75
11	--	--	76
12	--	--	81

--: Clade not recovered.

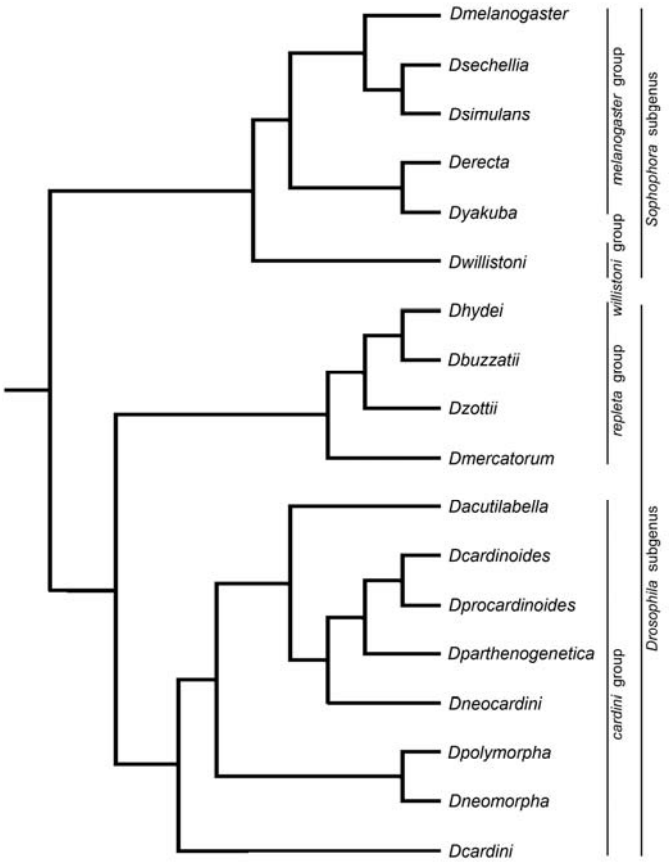


Figure 2: Representation of the evolutionary relationships of the species used in the phylogenetic analysis of *micropia* related sequences. This schematic tree is based on Robe et al. (2005), Brisson et al. (2006), and *Drosophila* 12 Genome Consortium (2007). *Drosophila zottii* position was inferred though a NJ analysis with the available *Amd* sequences.

Synonymous divergence values

To better examine the phylogenetic incongruence causes, we performed comparisons between dS values obtained for *micropia* and those for a nuclear gene, alpha-methyldopa (*Amd*). Although active elements might suffer a selective pressure to maintain its sequences, especially in the reverse transcriptase domain, *Amd* is expected to be even more restricted, since it plays essential roles in the *Drosophila* development and viability (Wang et al. 1996).

In recent events of HT transmission, it is expected a significantly lower dS value for the TE than for the nuclear gene, reflecting the lower divergence time of the TE sequences in relation to the divergence time of the species (Silva and Kidwell 2000; Loreto et al. 2008; Bartolomé et al. 2009). From the 1,378 pairwise dS comparisons a total of 358 significantly lower *micropia* dS values; statistically tested using the Fisher's exact test; were encountered (Table 3S), indicating that the source of the phylogenetic incongruences could be related to HT events.

In general, the significant lower dS values were concentrated in the pairwise comparisons executed within three main groups: those relating different sequences that compose clade 3, comparing sequences from *D. melanogaster*, *D. sechellia* and *D. simulans*; those from clade 7, comparing the sequences D.melanogaster9, D.sechellia29 and 27, D.simulans4, 9 and 10 with the sequences from *D. yakuba*; and those from clade 5, comparing D.willistoni1 and all sequences from the *repleta* and *cardini* species groups, and these last against each other and within themselves. It is noticeable that the highest significant values of Fisher's exact test were presented in the comparisons within clade 3, followed by those comparing sequences from clade 11 (Table 3S), that is, the most polytomic clades.

Different causes could lead to these low TE dS values, for example the sites required in splicing mechanisms or those involved in RNA secondary structure and other aspects related to functional RNAs (Parmely et al. 2006; Xing and Lee 2006). Codon usage bias is also a factor that may be responsible for low dS values for the TE sequences, since there is a negative correlation between codon usage bias and the dS values (Vidal et al. 2009). So, this possibility needs to be considered as a source of HT overestimation. However, *micropia* shows lower codon bias indices than *Amd* (Figure 3A), and also higher effective number of codons (Figure 3B). Taken together, these data validate the use of dS comparisons to infer HT events for *micropia*.

Analyzing the dN and dS substitutions along the *micropia* and *Amd* codon sequences it is possible to see that the dS substitutions are evenly distributed through the *micropia* codons (Figure 4A), which is at contrast with *Amd*, that presents peaks and valleys of dS values (Figure 4B). This pattern allows to infer that, whichever is the source of the low synonymous divergences, it is acting homogenously along the approximately 700 bp of the analyzed *micropia* sequence, a pattern not expected through the action of selective pressures related to both, coding or non coding features (RNA interference, for example), but expected for HT. As a whole, *micropia* presents an average dN/dS value which is higher than that presented by *Amd* (0,167[2] and 0,134, respectively), reinforcing the hypothesis of a lower selective pressure for the TE.

[2] It is important to detach that this value is greater than that of *Amd*, even with only coding sequences being retained in the analysis, a matrix which is highly skewed for related sequences.

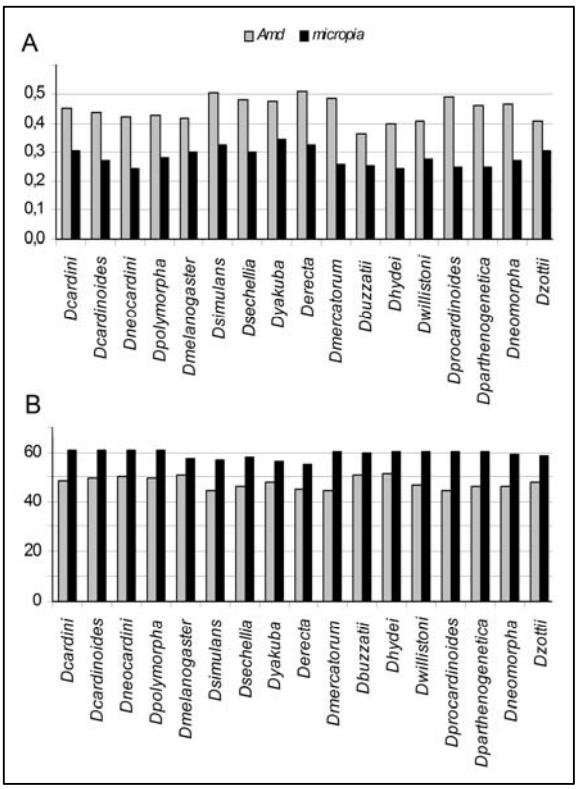


Figure 3: Comparisons of (A) codon bias index, CBI, and (B) effective number of codon, ENC, between *micropia* and *Amd* gene for each species. ENC varies between 21 (maximum bias) and 61 (minimum bias) and CBI varies between 0 (no bias) and 1 (maximum bias).

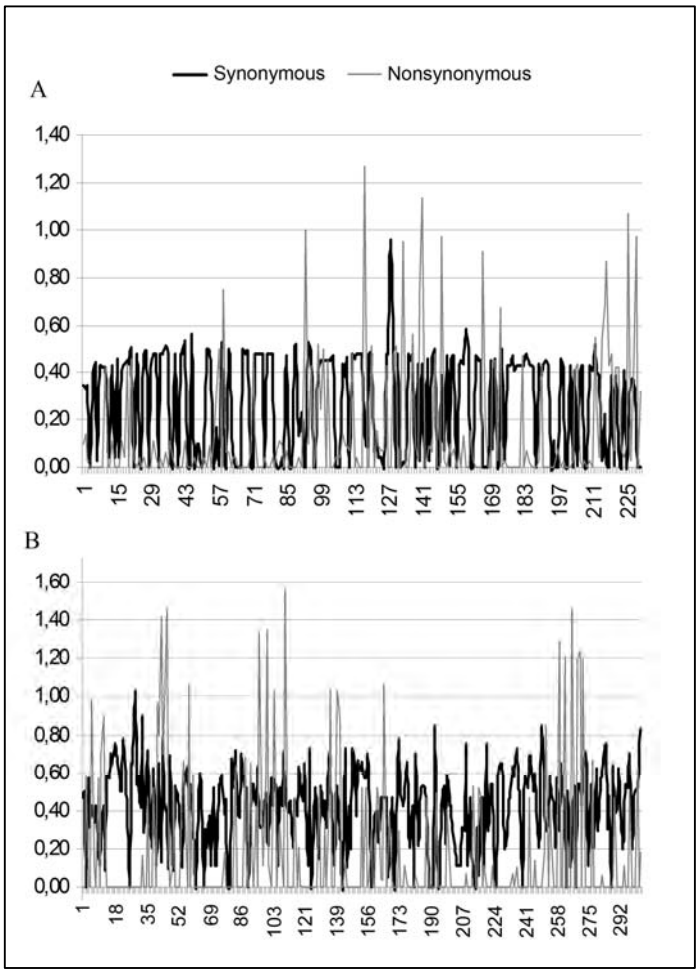


Figure 4: Distribution of the synonymous and nonsynonymous substitutions per site of each codon position within the *micropia* putative encoding sequence alignment (A) and the *Amd* matrix (B) as performed by SNAP (Korber, 2000; as available at www.hiv.lanl.gov). These values were calculated only for the *micropia* coding sequences (see Figure 1).

Micropia new classification

Analyzing the phylogenetic relationships of the *micropia* sequences studied here (Figure 1), a general subdivision in two groups is clearly observed: the sequences are divided between those grouped in clade 1 and those constituting clade 2. To clarify if this subdivision is reflecting the presence of different families or subfamilies of *micropia* related sequences in the species of the *Drosophila* genus, we evaluated the similarity values of the nucleotide and amino acid sequences. In this way, we grouped the sequences from clade 1 and compared them with the clustered sequences from clade 2. This same analysis was performed between the grouped sequences from clade 3 and the sequences from clade 4 and between the sequences from clade 5 and clade 6 (Table 4). The pattern that emerges is that the sequences from clade 1 (that are exclusively present in the *melanogaster* subgroup), is similar to the sequences from clade 2 in an average of 56.6% and 73.4% of the nucleotide and amino acid sequences, respectively. Moreover, within clade 2, the sequences from clade 6 (belonging to species from the *willistoni* group) are similar with those composing clade 5 (belonging mainly to species from the *Drosophila* subgenus, that is, the *cardini* and *repleta* species groups) in an average of 71% and 79.8% of the nucleotide and amino acid positions, respectively.

Table 4: Nucleotide and amino acid similarity values for *micropia* sequences grouped in clades (see Figure 1).

	Nucleotide (%)	Amino Acid (%)
clade1 x clade2	56.6	73.4
clade3 x clade4	89.3	87.4
clade5 x clade6	71.0	79.8

Discussion

Are there three different *micropia* families from *Drosophila* genus?

Except for the very well accepted criteria used to classify TEs in classes and subclasses proposed by Finnegan (1989), in general, there is no consensus in the criteria adopted to achieve family and subfamily definition. Several authors used different strategies to identify new TE families and subfamilies, whether based on nucleotide and/or amino acid sequence similarities (Lohe et al. 1995; Clark and Kidwell 1997; Capy et al. 1998; Heredia et al. 2004; Wicker et al. 2007; Bao et al. 2009; Smith et al. 2009). Given the abundance and diversity of TEs, Wicker et al. (2007) proposed a classification for eukaryotic TEs based on nucleotide similarity, with relevance not only to TE specialists, but also to genome annotators. This unified classification system is necessary in such a moment that more and more large and complex eukaryotic genomes are being sequenced and new repetitive sequences are being identified. Although this new attempt to TE classification might have some inconsistencies (Seberg and Petersen, 2009), due to characteristics generated by their behavior to recombine and the complex structure of the TE populations, it is the most parsimonious classification system for TEs until now proposed.

As reviewed by Capy et al. (1998), the *micropia* TE is a family of the *Ty*3 superfamily, from LTR retrotransposons subclass, belonging to class I of RNA-intermediated elements. Based on Wicker's et al. (2007) rule; which defines a TE family as a group of sequences longer than 80bp that "share 80% of sequence identity in at least 80% of their coding or internal domain, or within their terminal repeat region, or in both"; to establish the new classification of *micropia* sequences, we propose here the re-classification of the former *micropia* family into three new families. The first family,

herein named *nanopia*, is present on species of the *melanogaster* subgroup (clade 1 in Figure 1); the second family, the *micropia* family [since *micropia* was first described in *D. hydei* (Huijser et al. 1988; Lankenau et al. 1988)], is present mostly on the species from *Drosophila* subgenus (clade 5); and the third family, herein named *picopia*, is encountered in *D. willistoni* and maybe other related species. *micropia* was named after the microdissection experiments and the similarity of the sequences with the *copia* elements (Huijser *et al*., 1988). The proposed *nanopia* and *picopia* names are not related with metrical values. This *micropia* families division is supported by our analysis of nucleotide and amino acid similarities (Table 4) and follows the hierarchical step by step procedure implemented by Wicker et al. (2007), besides presenting a strong connection with *micropia* related sequences phylogenetic history and proposing only monophyletic groups (Seberg and Petersen, 2009).

Evolutionary history of *micropia* in *Drosophila* genus

Previous works have proposed HT events between *Drosophila* species based on different kind of inferences (Silva and Kidwell 2000; Heredia et al. 2004; Almeida and Carareto 2006; Sánchez-Garcia et al. 2005; Bartolomé et al. 2009). Here, we adopted the following criteria to infer HT evens, as indicated by Loreto et al. (2008): (a) analysis of the TE patchy distribution in the species; (b) incongruence in the comparisons between the TE phylogeny and the species relationships described elsewhere; (c) analysis of the sequences similarity thought dS values comparisons between the TE and a nuclear gene, with support of the dN/dS index and the codon usage bias analysis. In our evolutionary inferences, VT and ancestral polymorphisms appear to be a strong feature, since faint signals were recovered in Dot Blot and in Southern Blot both, for species not presenting detectable

micropia related sequences (Figure 1S) or for species with *micropia* characterized elements [Figure 2 in Cordeiro et al. (2008)]. Nevertheless, in the attempt of outlining the evolutionary scenario of *micropia*, *nanopia* and *picopia* retroelements in the species here studied, our results show that, beyond VT and ancestral polymorphism, there might have happen successive invasions of *micropia* related sequences in the species genome after the speciation events.

The *nanopia* sequences detected in *D. erecta* and *D. yakuba* seem to be representatives of ancient elements likely present in the *Sophophora* subgenus most recent common ancestral. This prelude is taken from their dS values, which are not at high contrast with those presented for *Amd* (Table 3S) in the comparisons involving both these species sequences, in those comparisons involving these sequences and some of the *D. melanogaster* cryptic species sequences (clade 3 in Figure 1), or in those comparisons involving these and both, the *D. willistoni picopia* sequences and the *Drosophila* subgenus *micropia* sequences (clade 2 in Figure 1). Nevertheless, some extremely similar representatives of this group of sequences were also encountered within the *D. melanogaster* cryptic species (Dsechellia27 and 29, Dmelanogaster9, Dsimulans4, 9 and 10 sequences) (clade 7 in Figure 1). Considering the evolution rate calibration as 1.6% divergence per MYA (Sharp and Li 1989), which is indicated for neutral evolution sequences and using only dS values, it is possible to infer that the *D. erecta* and *D. yakuba* *nanopia* sequences diverged from the *D. melanogaster* cryptic species sequences that comprise clade 7 at about 5.3MYA and 3.3MYA, respectively, which is well after the species split (12.8MYA, Tamura et al. 2004). So, we conceive that these *D. melanogaster* cryptic species *nanopia* sequences were independently acquired through at least one HT event involving a *D. yakuba* and a *D. melanogaster* cryptic species ancestral lineage.

Although the *D. melanogaster* cryptic species sequences comprising clade 7 present a polytomic relationship (Figure 1) and have lower dS values in the intra-group comparisons, the difference with *Amd* is not statistically significant most of the times (Table 3S), so that VT could not be discarded.

Within clade 3 (Figure 1), nevertheless, the picture is different. In a first sight, it is possible to see that these *nanopia* sequences, as detected in *D. melanogaster*, *D. sechellia* and *D. simulans*, are extremely similar (clade 3 in Figure 1; Table 3S), which could be an indication that these species have acquired their sequences very recently. According to Tamura et al. (2004), *D. melanogaster* shared the most recent common ancestor with *D. simulans* and *D. sechellia* 5.4MYA. Applying the above described molecular clock hypothesis, it is possible to infer that the *nanopia* sequences within these species diverged 0.5MYA, i.e., well after the speciation events. So, HT events could be invoked as a possible source of incongruence. Although it was identified that these three species can generate fertile hybrids (Lachaise et al. 1986), introgression may have played an essential role in the acquisition of new sequences. Other authors have identified the same highly frequencies of HT events between these species, both for transposons and for retrotransposons (Sanchez-Garcia et al. 2005; Ludwig and Loreto 2007; Ludwig et al. 2008; Bartolomé et al. 2009) and some of them also pointed to introgression as a possible source mechanism (Silva and Kidwell 2000; Ludwig et al. 2008). It is well known that *D. simulans* and *D. melanogaster* are widespread species; but all the other species from *melanogaster* subgroup have their distribution limited to the African continent and *D. sechellia* has an even more restricted distribution, known only for the Seychelles islands (Lemeunier *et al.* 1986; Bächli, 2009). So, at least one of the inferred HT events might

have occurred in Africa, where there is a main contact of the species, mainly regarding *D. sechellia*.

In summary, comparing the TE (Figure 1) and the species phylogeny (Figure 2), and utilizing the dS values information, the conservative scenario that emerges for clade 1 is the presence of a *nanopia* copy in the *melanogaster* subgroup ancestor, with this being inherited both, by the *yakuba/erecta* lineage (giving rise to clade 4 with at least one HT) and by a *D. melanogaster* ancestral lineage (giving rise to clade 3, with at least two independent HT events). More than that, conservatively speaking, we hypothesize that the ancestral *nanopia* copy was likely present in the *Sophophora* ancestral lineage, since when comparing the sequences from clade 1 with those comprising clade 2, no significant incongruences were found as concerns the dS values. Nevertheless, this scenario also requests some *nanopia* stochastic losses, since we were not able to find any *nanopia* sequence in the *D. ananassae* (*ananassae* subgroup of *melanogaster* group), *D. persimilis* and *D. pseudoobscura* (*obscura* group) genomes through our *in silico* search.

Clade 2 from Figure 1 presents an even more complex pattern, with significant incongruences with the species phylogeny (Figure 2): the *picopia* sequences from *D. willistoni* (clade 6 - *Sophophora* subgenus) clustered with the *micropia* sequences mainly found in species from the *cardini* and *repleta* group (clade 5 - *Drosophila* subgenus), with a strong support in all phylogenetic analysis (Figure 1; Table 2). In general, the majority of the *D. willistoni picopia* sequences (clade 6) seem to have a higher evolving rate. But even so, significant deviations between the TE and *Amd* dS values are detected in most of the comparisons involving clades 5 and 6. More than that, a *D. willistoni micropia* sequence, Dwillistoni1, is clustered within the clade grouping the *cardini* and *repleta* species groups copies, and presents lower dS values in the comparisons with these *Drosophila* subgenus

micropia sequences than with the *D. willistoni picopia* sequences (clade 6 in Figure 1) (willi1 cluster in Table 3S). The *Drosophila* subgenus, that embraces the *cardini* and *repleta* group species, diverged from the *Sophophora* subgenus, containing the *willistoni* species group, at about 62.9MYA (Tamura et al. 2004), whereas the sequences from clade 5 and 6 diverged from each other at an estimated 19MYA. Therefore, we conceive that the *cardini* and *repleta* groups both acquired a *micropia* related sequence from *D. willistoni*, or some other related species, with a possible second HT being postulated in the reverse sense, that is, from some species of the *cardini/repleta* group back to *D. willistoni*. As the *willistoni* and *saltans* group diverged at about 20MYA (Powell et al. 2003), almost the same time since the split of the *micropia* and *picopia* sequences (clades 5 and 6) (19MYA), the donor species of the *repleta* and *cardini* groups sequences could be some species from the *saltans* group, which also possess *micropia* and/or *picopia* sequences (Almeida et al. 2001). However, as the sequences obtained for the *saltans* group species embraces another nucleotide region; this hypothesis has to be further elucidated.

Within the *cardini* and the *repleta* groups, *micropia* has a patchy distribution, since neither the species from the *dunni* subgroup (*cardini* group) nor *D. mojavensis* (*repleta* group) seem to have detectable copies of *micropia* in their genomes. Almost all the dS comparisons concerning species from clade 5 were highly significant (Table 3S). Moreover, clade 5 does not present any phylogenetic subdivision between the *cardini* and *repleta* groups species, which is unexpected since these species groups appear to have diverged at about 45MYA (L. J. Robe, unpublished data). So, in an attempt to explain all these results, we propose the occurrence of independent HT events from member(s) of the *willistoni/saltans* lineage to at least one member from the *repleta* and, maybe, from the *cardini* group. Since *D. zottii* is the early offshoot within clade 5, it was probably the first

receptor species (4.1MYA), followed by *D. acutilabella* (3.7MYA), that may have received its copy again from a *willistoni/saltans* species, or from *D. zottii* itself. Afterwards, possibly through the action of introgression mechanisms, *micropia* spread to other members of both of these groups in an almost simultaneous fashion (about 2MYA), since both of them have species that also can interbreed and produce fertile hybrids (Heed 1962; Wasserman 1992). So, the evolutionary history of *micropia* sequences within clade 2 appears to be governed mainly by HT events, with at least 14 events being inferred[3].

All the species from clade 2 have a Neotropical restrict distribution, with the exception of the *D. hydei*, *D. mercatorum* and *D. buzzatii* (all from *repleta* group) that are widespread species (Fontdevila *et al.* 1981; Val et al. 1981; Bächli 2009). *Drosophila willistoni*, from *willistoni* group, and the species *D. polymorpha*, *D. cardinoides*, *D. neocardini*, *D. cardini*, *D. neomorpha* and *D. parthenogenetica*, all from the *cardini* subgroup, have the same widely distribution pattern ranging from south USA through South America (Heed and Russel 1971; Dobzhansky and Powell 1975; Bächli 2009). *Drosophila procardinoides* and *D. acutilabella* are the exceptions within the *cardini* subgroup: the first is endemic to Bolivia, Peru and north of Brazil, whereas the second occurs only in south of USA, Mexico and some Caribbean islands (Brisson et al. 2006; Bächli 2009). Thus, the statement made by Cordeiro et al. (2008), that indicated the Central and/or South America as the likely contact site for the HT events, is also validated here.

HT mechanisms

[3] *Drosophila polymorpha* does not present significant lower dS values for *micropia* in the comparisons involving *D. acutilabella*, *D. cardinoides*, *D. neomorpha*, *D. parthenogenetica* and *D. procardinoides*. However, this result probably reflects the fact that this species presents a 133bp deletion in its *micropia* sequence, so that a reduced sample of synonymous sites is compared.

The *micropia* related sequences contain one open reading frame, flanked by LTRs, that putatively encodes protease, integrase, reverse transcriptase, RNseH, and a capsid protein similar to that from the vertebrates' retroviruses (Lankenau et al. 1989; 1994). According to the International Committee on Taxonomy of Viruses (ICTV) (Boeke et al. 1999), *micropia*, is a metavirus classified as belonging to the Metaviridae family, as does it might be for *nanopia* and *picopia* TE families. The retroviruses and retroelements belonging to this family do not have a functional coding *env* gene, which gives the infectious properties to retroviruses such as *gypsy* (Heredia et al. 2004). However, the hypothesis that the elements without *env* gene can use the Env protein provided by other TEs to produce their infectious particles (Vidal et al. 2009) can not be discarded. Different vectors that could lead to HT on genes have been described, such as process mediated by virus, parasites and parasitoids, intracellular symbiotic bacteria, and others (Silva et al. 2004; Loreto et al. 2008). Considering these possibilities, the sequences here studied seems to have used some of these strategies; however, introgression mechanisms detach, in a first sight, as the most likely source of most HT events here inferred, at least as the *melanogaster*, *cardini* and *repleta* intra-groups transfers are concerned.

Acknowledgements

This work was supported by grants from CAPES, CNPq, and FAPERGS research funding. The authors thank MSc Ricardo Weber for valuable help, to Dr Ronaldo Medeiros Golombieski and MSc Paloma Rubin for help in sequencing the DNA templates, and the members of the *Drosophila* Laboratory from Universidade Federal do Rio Grande do Sul,

especially to BSc Gisele de Souza da Silva for unconditional help.

References

Abascal F, Zardoya R, and Posada D (2005) ProtTest: Selection of best-fit models of protein evolution. Bioinformatics 21(9):2104-2105

Akaike H (1974) A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19:716–723

Almeida LM, Carareto CMA (2004) Identification of two subfamilies of *micropia* transposable element in species of the *repleta* group of *Drosophila*. Genetica 121:155-164

Almeida LM, Carareto CMA (2006) Sequence heterogeneity and phylogenetic relationships between the *copia* retrotransposon in *Drosophila* species of the *repleta* and melanogaster groups. Genet Sel Evol 38:535-50

Almeida LM, Castro JP, Carareto CMA (2001) *micropia* transposable element occurrence in *Drosophila* species of the *saltans* group. Drosoph Inf Serv 84:114-117

Bächli G (2009) Taxodros: The database on taxonomy of Drosophilidae. Available at http://www.taxodros.unizh.ch. Accessed in July 30, 2009

Bao W, Jurka MG, Kapitonov VV, Jurka J (2009) New superfamilies of eukaryotic DNA transposons and their internal divisions. Mol Biol Evol 26(5):983-993

Bartolomé C, Bello X, Maside X (2009) Widespread evidence for horizontal trnafer of transposable elements across *Drosophila* genomes. Genome Biol. 10(2):R22

Bennetzen JL, Hall BD (1982) Codon selection in yeast. J Biol Chem 257:3026-3031

Boeke JD, Eickbush TH, Sandmeyer SB and Voytas DF (1999) Metaviridae. In Murphy FA, (ed) Virus Taxonomy: ICTV VII Report. Springer-Verlag, NY. 123–135

Brisson JA, Wilder J, Hollocher H (2006) Phylogenetic analysis of the *cardini* group of *Drosophila* with respect to changes in pigmentation. Evolution 60:1228-1241

Cáceres M, Ranz JM, Barbadilla A, Long M, Ruiz A (1999) Generation of a widespread *Drosophila* inversion by a transposable element. Science 285:415-418.

Capy P, Bazin C, Higuet D, Langin T (1998) Dynamics and evolution of transposable elements. Texas: Landes Bioscience 197p

Clark JB, Kidwell MG (1997) A phylogenetic perspective on *P* transposable element evolution in *Drosophila*. Proc Natl Acad Sci USA 94:11428-11433.

Cordeiro J, Robe LJ, Loreto EL, Valente VL (2008) The LTR retrotransposon *micropia* in the *cardini* group of *Drosophila* (Diptera: Drosophilidae): a possible case of horizontal transfer. Genetica134(3):335-344.

Craig N, Craigie R, Gellert M, Lambowitz A (2002) Mobile DNA II. Washington, DC: ASM Press.

Dobzhansky T, Powell JR (1975) The *willistoni* group of sibling species of *Drosophila*. In Handbook of genetics. King RC (ed) 3vol, Plenum, New York, pp 589-622.

Drosophila 12 Genomes Consortium. Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, Iyer VN *et al*. (2007) Evolution of genes and genomes on the Drosophila phylogeny. Nature 450(7167):203-18.

Felsenstein J (1985) Confidence limits on phylogenies: an approaching using bootstrap. Evolution 39:783-791

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

Fontdevila A, Ruiz A, Alonso G, Ocaiia R (1981) The evolutionary history of *Drosophila buzzatii*. I. Natural chromosomal polymorphism in colonized populations of the Old World. Evolution 35:148-157

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Sys Biol 52(5): 696-704

Hall BG (2001) Phylogenetic trees made easy - a how to manual for molecular biologists, 2ed. Sinauer Associates, Massachusetts

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol. 22:160-74

Heath BD, Butcher RD, Whitfield WG, Hubbard SF (1999) Horizontal transfer of *Wolbachia* between phylogenetically distant insect species by a naturally occurring mechanism. Curr Biol 9:313-316

Heed WB (1962) Genetic characteristics of island populations. Univ Texas Publ Stud Genet 6205:173-206

Heed WB, Russel JS (1971) Phylogeny and population structure in island and continental species of the *cardini* group of *Drosophila* studied by inversion analysis. Univ. Texas Publs 7103:91-130

Heredia FO, Loreto ELS, Valente VLS (2004) Complex evolution of *gypsy* in drosophilid species. Mol Biol Evol 21:1-12

Houck MA, Clark JB, Peterson HR, Kidwell MG (1991) Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. Science 253:1125-1128

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755

Huijser P, Kirchhoff C, Lankenau D-H, Hennig W (1988) Retrotransposon-like sequences are expressed in Y chromosomal lampbrush loops of *Drosophila hydei*. J Mol Biol. 203:689-697

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with ClustalX. Trends Biochem Sci. 23:403-405

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comp Appl Biosci 8:275-282

Jordan IK, Maayunina LV, McDonald JF (1999) Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. Proc Natl Acad Sci USA 22:12621-12625

Kidwell MG, Lisch DR (2001) Transposable elements, parasitic DNA, and genome evolution. Evolution 55:1-24

Korber B (2000) HIV Signature and Sequence Variation Analysis. Computational Analysis of HIV Molecular Sequences. In: Rodrigo AG, Learn GH (eds). Dordrecht, Netherlands: Kluwer Academic Publishers. 55-72

Lachaise D, David JR, Lemeunier F, Tsacas L, Ashburner M (1986) The reproductive relationships of *Drosophila sechellia* with *D. mauritiana, D. simulans* and *D. melanogaster* from the Afrotropical region. Evolution 40:262–271

Lankenau D-H (1993) The retrotransposon family *micropia* in *Drosophila* species. In: McDonald J (ed) Transposable Elements and Evolution. Amsterdam: Kluwer Publishers. pp 232-241

Lankenau D-H, Huijser P, Hennig W (1989) Characterization of the long terminal repeats of *micropia* elements microdissected from the Y-chromosomal lampbrush loops “threads” of *Drosophila hydei*. J Mol Biol. 209:493-497

Lankenau D-H, Huijser P, Jansen E, Miedema K, Hennig W (1988) *micropia*: a retrotransposon of *Drosophila* combining structural features of DNA viruses, retroviruses and non-viral transposable elements. J Mol Biol. 2:233-246

Lankenau D-H, Huijser P, Jansen E, Miedema K, Hennig W (1990) DNA sequence comparison of *micropia* transposable elements from *Drosophila hydei* and *Drosophila melanogaster*. Chromosoma 99:111-117

Lankenau S, Corces GV, Lankenau D-H (1994) The *Drosophila micropia* retrotransposon encodes a testis-specific antisense RNA complementary to reverse transcriptase. Mol Biol Evol. 17:1542-1557

Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol 25(7):1307-1320

Lemeunier F, David JR, Tsacas L, Ashburner M (1986) The melanogaster species group. In: Ashburner M, Carson HL, Thompson Jr JN (eds) The genetics and biology of *Drosophila*. New York: New York Academic Press. 147-256.

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452

Lippman Z, Gendrel A-V, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, Carrington KD, Doerge RW, Colot V, Martienssen R (2004) Role of transposable elements in heterochromatin and epigenetic control. Nature 430:471-476

Lohe AR, Moriyama EM, Lidholm DA, Hartl DL (1995) Horizontal transmission, vertical inactivation, and stochastic loss of mariner like transposable elements. Mol Biol Evol 12:62–72

Loreto EL, Carareto CM, Capy P (2008) Revisiting horizontal transfer of transposable elements in *Drosophila*. Heredity 100:545-554

Ludwig A, Loreto ELS (2007) Evolutionary pattern of the *gtwin* retrotransposon in the *Drosophila melanogaster* subgroup. Genetica 130:161–168.

Ludwig A, Valente VLS, Loreto ELS (2008) Multiple incasions of errantivirus in the genus *Drosophila*. Ins Mol Biol 17:113-124

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol. 3:418-426

Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala.

Parmley JL, Chamary JV, Hurst LD (2006) Evidence for purifying selection against synonymous mutations in mammalian exonic splicing enhancers. Mol Biol Evol 2:301-309

Pooma W, Gersos C, Grotewold E (2009) Transposon insertions in the promoter of the *Zea mays* a1 gene differentially affect transcription by the Myb factors P and C1. Genetics 161(2):793-801.

Posada C, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818

Powell JR, Sezzi E, Moriyama EN, Gleason JM, Caccone A (2003) Analysis of a shift in codon usage in *Drosophila*. J Mol Evol 57:214-225

Puig M, Cáceres M, Ruiz A (2004) Silencing of a gene adjacent to the breakpoint of a widespread *Drosophila* inversion by a transposon-induced antisense RNA. Proc Natl Acad Sci USA 101:9013-9018.

Robe LJ, Valente VLS, Budnik M, Loreto ELS (2005) Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: A nuclear versus mitochondrial gene approach. Mol Phylogenet Evol. 3:623-640

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406-25

Sánchez-Gracia A, Maside X, Charlesworth B (2005) High rate of horizontal transfer of transposable elements in *Drosophila*. Trends Genet. 21(4):200-3.

Sassi AK, Heredia FO, Loreto ELS, Valente VLS, Rohde C (2005) Transposable elements *P* and *gypsy* in natural populations of *Drosophila willistoni*. Genet Mol Biol. 28:734-739

Seberg O, Petersen G (2009) A unified classification system for eukaryotic transposable elements should reflect their phylogeny. Nat Rev Genet 10(4):276

Sharp PM, Li W-H (1989) On the rate of DNA sequence evolution in *Drosophila*. J Mol Evol. 28:398-402

Silva JC, Kidwell MG (2000) Horizontal transfer and selection in the evolution of P elements. Mol Biol Evol 17(10):1542-57.

Silva JC, Loreto ELS, Clark JB (2004) Factors that affect the horizontal transfer of transposable elements. Curr Issues Mol Biol. 6:57-72

Smith M, Bringaud F, Papadopoulou B (2009) Organization and evolution of two *SIDER* retroposon subfamilies and their impact on the *Leishmania* genome. BMC Genomics 10:240.

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

Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts

Tamura K, Subramanian S, Kumar S (2004) Temporal patterns of fruit fly (Drosophila) evolution revealed by mutation clocks. Mol Biol Evol 21(1):36-44.

Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C biases. Mol Biol Evol. 9:678-687

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

Val FC, Vilela CR, Marques MD (1981) Drosophilidae of the Neotropical region. In: Ashburner M, Carson HL and Thompson Jr JN (ed) The Genetics and Biology of *Drosophila*. 3a. Academic Press, London. pp 123-168.

Vidal N, Ludwig A, Loreto ELS (2009) Evolution of *Tom, 297 17.6* and *rover* retrotransposons in Drosophilidae species. Mol Genet Genomics, doi 10.1007/s00438-009-0468-0.

Vieira C, Nardon C, Arpin C, Lepetit D, Biémont C (2002) Evolution of genome size in Drosophila. Is the invader's genome being invaded by transposable elements? Mol Biol Evol 19(7):1154-61.

Wang D, Marsh JL, Ayala FJ (1996) Evolutionary changes in the expression pattern of a developmentally essential gene in three *Drosophila* species. Proc Natl Acad Sci USA 93:7103-7

Wasserman M (1992) Cytological evolution of the *Drosophila repleta* species group. In: Krimbas CB, Powell JR (eds) *Drosophila* inversion polymorphism. 455-552

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8(12):973-82.

Wright F (1990) The 'effective number of codons' used in a gene. Gene 87:23-29.

Xing Y, and Lee C, (2006) Can RNA selection pressure distort the measurement of Ka/Ks? Gene 370: 1-5.

Supplementary Material

Table 1S: *Drosophila* species investigated in this study which did not present any amplification/hybridization signal of *micropia* related sequences.

Group	Species	Coordinate or Stock Center number	Localization
cardini	*D. dunni*	15182-2301.00	St Thomas-Caribe
	D. similis	15182-2321.00	St George-Grenada
	D. nigrodunni	15182-2311.01	Barbados-Caribe
	D. arawakana	15182-2260.00	St Kitts-Caribe
tripunctata	*D. bandeirantorum*[A]	30º03'S 51º07'W	Porto Alegre-RS-BR
	D. mediopictoides	15220-2371.00	Boquete-Panama
	D. mediopunctata[A]	30º03'S 51º07'W	Porto Alegre-RS-BR
	D. mediodiffusa	15220-2351.04	Maricao-Porto Rico
	D. paraguayensis[B]	27º35'S 48º22'W	Florianopolis-SC-BR
	D. paramediostriata[A]	30º03'S 51º07'W	Porto Alegre-RS-BR
	D. tripunctata	15220-2401.12	Iowa River-Iowa-EUA
guarani	*D. guaru*[C]	26º17'S 49º00'W	Joinville-SC-BR
	D. ornatifons[C]	26º17'S 49º00'W	Joinville-SC-BR
guaramunu	*D. maculifrons*[D]	27º35'S 48º22'W	Florianopolis-SC-BR
	D. griseolineata[D]	27º35'S 48º22'W	Florianopolis-SC-BR
funnebris	*D. funnebris*	15120-1911.01	Mexico City-MX
immigrans	*D. immigrans*[C]	26º17'S 49º00'W	Joinville-SC-BR

Capital letters refer to the fly collectors: [A]Dr. Luciano Basso da Silva, [B]Dr. Marco Silva Gottschalk, [C]MSc. Jonas da Silva Doge, [D]Dra. Daniela Cristina De Toni.

Table 2S: *Drosophila* species analyzed *in silico* with their respective scaffold retrieved from the genome search.

Group	Species	Scaffold/Chromosome/Trace	Adopted Nomenclature
willistoni	*D. willistoni*	scf2_1100000004965	Dwillistoni1
		scf2_1100000004954	Dwillistoni2, 3
		scf2_1100000004953	Dwillistoni4
		scf2_1100000004961	Dwillistoni5
		scf2_1100000004959	Dwillistoni6
melanogaster	*D. melanogaster*	3R	Dmelanogaster1, 2
		3RHet	Dmelanogaster3, 4
		U	Dmelanogaster5, 6, 7
		2R	Dmelanogaster9
	D. simulans	chrU_M_2288	Dsimulans1
		chr2h_Mrandom_009	Dsimulans2
		2R	Dsimulans3, 4
		chrU_M_786	Dsimulans5
		chrU_M_6079	Dsimulans6
		X	Dsimulans7
		chrU_M_978	Dsimulans8
		2h_Mrandom_005	Dsimulans9
		U_M_2475	Dsimulans10
	D. sechellia	scafold 1	Dsechellia1, 2, 3
		scaffold 7	Dsechellia4, 5
		scaffold 0	Dsechellia6, 7, 8, 9
		scaffold 9599	Dsechellia10
		scaffold 2140	Dsechellia11
		scaffold 62	Dsechellia12
		scafold 89	Dsechellia13
		scaffold 2	Dsechellia14
		scafold 1078	Dsechellia15
		scafold 1972	Dsechellia16
		scafold 935	Dsechellia17
		scafold 28	Dsechellia18
		scafold 6	Dsechellia19
		scafold 443	Dsechellia20
		scafold 72	Dsechellia21
		scafold 15	Dsechellia22
		scafold 1378	Dsechellia23
		scafold 12503	Dsechellia24
		scafold 688	Dsechellia25
		scafold 218	Dsechellia26
		scafold 66	Dsechellia27
		scafold 13	Dsechellia28
		scaffold 57	Dsechellia29
	D. erecta	scaffold 4929	Derecta1
	D. yakuba	V2_chrUn_1883	Dyakuba1
		V2_chrUn_975	Dyakuba2
		V2_chrUn_709	Dyakuba3
		V2_chr2h_random_005	Dyakuba4
		V2_chrUh_007	Dyakuba5, 6, 7
		V2_chrUn_080	Dyakuba8
		V2_chrUn_3220	Dyakuba9
		V2_chrX_random_046	Dyakuba10, 11, 12

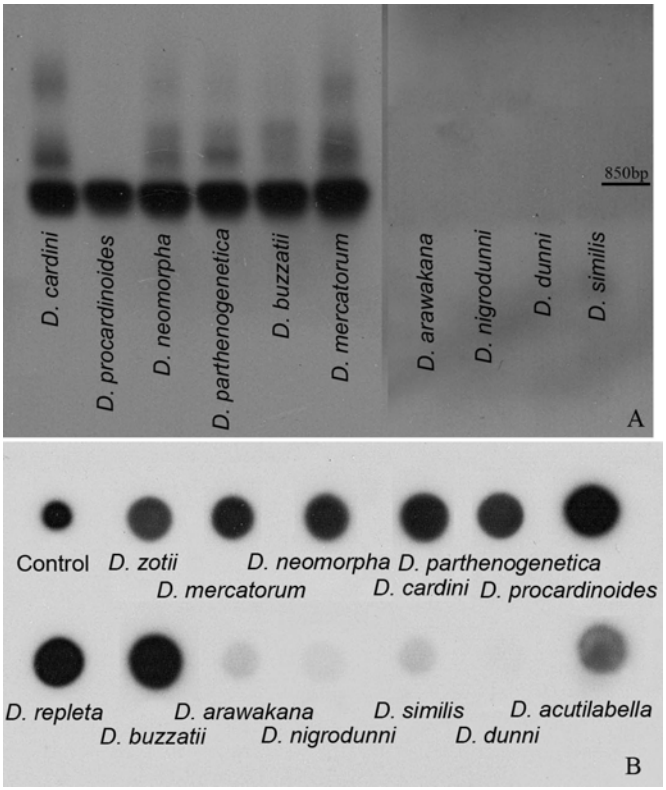


Figure 1S: (A) Southern Blot on the PCR of some species from the *cardini* and *repleta* groups (see Table 1 and Table 2S). (B) Dot Blot on genomic DNA confirming the pattern seen on the PCR/Southern Blot. In both cases, the probe used was an 812bp PCR fragment from *D. hydei* dhMiF2 sequence. Control: 5ul (in 10ul) of *micropia* probe.

Table 3S: Comparative analysis of the dS values between *micropia* and *Amd* sequences. Only the sequences from different species were compared. The comparisons suggesting the occurrence of HT were statistically tested by Fisher's exact test. (*) P<0.05, (**) P<0.01, (***) P<0.001. See footnote for sequences grouping explanation.

Sequences comparisons		dS *micropia*	dS *Amd*	Fisher's exact test
Dsechellia1 X	Derecta1	0.195	0.257	
Dsechellia1 X	Dsimulans1	0.018	0.080	**
Dsechellia1 X	Dsimulans2	0.018	0.080	**
Dsechellia1 X	Dsimulans3	0.000	0.080	***
Dsechellia1 X	sim1	0.210	0.080	
Dsechellia1 X	Dsimulans5	0.012	0.080	**
Dsechellia1 X	Dsimulans6	0.024	0.080	*
Dsechellia1 X	Dsimulans7	0.026	0.080	*
Dsechellia1 X	Dsimulans8	0.016	0.080	**
Dsechellia1 X	Dsimulans9	0.219	0.080	
Dsechellia1 X	Dm11	0.019	0.160	***
Dsechellia1 X	Dmelanogaster1	0.006	0.160	***
Dsechellia1 X	mel1	0.015	0.160	***
Dsechellia1 X	mel2	0.006	0.160	***
Dsechellia1 X	Dmelanogaster5	0.000	0.160	***
Dsechellia1 X	Dmelanogaster6	0.013	0.160	***
Dsechellia1 X	Dmelanogaster9	0.204	0.160	
Dsechellia1 X	yak1	0.223	0.280	
Dsechellia1 X	Dwillistoni1	0.766	0.798	
Dsechellia1 X	willi1	0.868	0.798	
Dsechellia1 X	dhMiF2	0.761	0.775	
Dsechellia1 X	Dmercatorum	0.758	0.647	
Dsechellia1 X	Dbuzzatii	0.777	0.742	
Dsechellia1 X	Dzottii	0.772	0.697	
Dsechellia1 X	Dacutilabella	0.776	0.592	
Dsechellia1 X	Dcardini1	0.796	0.639	
Dsechellia1 X	Dcardini2	0.799	0.639	
Dsechellia1 X	Dcardini3	0.799	0.639	
Dsechellia1 X	Dneomorpha	0.786	0.616	
Dsechellia1 X	parth1	0.774	0.624	
Dsechellia1 X	Dparthenogenetica4	0.796	0.624	
Dsechellia1 X	Dparthenogenetica5	0.791	0.624	
Dsechellia1 X	Dparthenogenetica6	0.797	0.624	
Dsechellia1 X	Dprocardinoides	0.764	0.584	
Dsechellia1 X	Dneocardini	0.794	0.645	
Dsechellia1 X	Dpolymorpha	0.766	0.627	
Dsechellia1 X	Dcardinoides	0.786	0.610	
sech1 X	Derecta1	0.195	0.257	
sech1 X	Dsimulans1	0.016	0.080	**
sech1 X	Dsimulans2	0.016	0.080	**
sech1 X	Dsimulans3	0.002	0.080	***
sech1 X	sim1	0.209	0.080	

sech1 X	Dsimulans5	0.014	0.080	**
sech1 X	Dsimulans6	0.027	0.080	*
sech1 X	Dsimulans7	0.027	0.080	*
sech1 X	Dsimulans8	0.017	0.080	**
sech1 X	Dsimulans9	0.220	0.080	
sech1 X	Dm11	0.021	0.160	***
sech1 X	Dmelanogaster1	0.008	0.160	***
sech1 X	mel1	0.016	0.160	***
sech1 X	mel2	0.008	0.160	***
sech1 X	Dmelanogaster5	0.002	0.160	***
sech1 X	Dmelanogaster6	0.015	0.160	***
sech1 X	Dmelanogaster9	0.203	0.160	
sech1 X	yak1	0.224	0.280	
sech1 X	Dwillistoni1	0.767	0.798	
sech1 X	willi1	0.871	0.798	
sech1 X	dhMiF2	0.762	0.775	
sech1 X	Dmercatorum	0.760	0.647	
sech1 X	Dbuzzatii	0.779	0.742	
sech1 X	Dzottii	0.772	0.697	
sech1 X	Dacutilabella	0.775	0.592	
sech1 X	Dcardini1	0.797	0.639	
sech1 X	Dcardini2	0.799	0.639	
sech1 X	Dcardini3	0.799	0.639	
sech1 X	Dneomorpha	0.787	0.616	
sech1 X	parth1	0.776	0.624	
sech1 X	Dparthenogenetica4	0.797	0.624	
sech1 X	Dparthenogenetica5	0.792	0.624	
sech1 X	Dparthenogenetica6	0.798	0.624	
sech1 X	Dprocardinoides	0.766	0.584	
sech1 X	Dneocardini	0.795	0.645	
sech1 X	Dpolymorpha	0.768	0.627	
sech1 X	Dcardinoides	0.787	0.610	
Dsechellia5 X	Derecta1	0.195	0.257	
Dsechellia5 X	Dsimulans1	0.018	0.080	**
Dsechellia5 X	Dsimulans2	0.018	0.080	**
Dsechellia5 X	Dsimulans3	0.000	0.080	***
Dsechellia5 X	sim1	0.210	0.080	
Dsechellia5 X	Dsimulans5	0.012	0.080	**
Dsechellia5 X	Dsimulans6	0.024	0.080	*
Dsechellia5 X	Dsimulans7	0.026	0.080	*
Dsechellia5 X	Dsimulans8	0.016	0.080	**
Dsechellia5 X	Dsimulans9	0.219	0.080	
Dsechellia5 X	Dm11	0.019	0.160	***
Dsechellia5 X	Dmelanogaster1	0.006	0.160	***
Dsechellia5 X	mel1	0.015	0.160	***
Dsechellia5 X	mel2	0.006	0.160	***
Dsechellia5 X	Dmelanogaster5	0.000	0.160	***
Dsechellia5 X	Dmelanogaster6	0.013	0.160	***
Dsechellia5 X	Dmelanogaster9	0.204	0.160	
Dsechellia5 X	yak1	0.223	0.280	
Dsechellia5 X	Dwillistoni1	0.766	0.798	
Dsechellia5 X	willi1	0.868	0.798	

Dsechellia5 X	dhMiF2	0.761	0.775	
Dsechellia5 X	Dmercatorum	0.758	0.647	
Dsechellia5 X	Dbuzzatii	0.777	0.742	
Dsechellia5 X	Dzottii	0.772	0.697	
Dsechellia5 X	Dacutilabella	0.776	0.592	
Dsechellia5 X	Dcardini1	0.796	0.639	
Dsechellia5 X	Dcardini2	0.799	0.639	
Dsechellia5 X	Dcardini3	0.799	0.639	
Dsechellia5 X	Dneomorpha	0.786	0.616	
Dsechellia5 X	parth1	0.774	0.624	
Dsechellia5 X	Dparthenogenetica4	0.796	0.624	
Dsechellia5 X	Dparthenogenetica5	0.791	0.624	
Dsechellia5 X	Dparthenogenetica6	0.797	0.624	
Dsechellia5 X	Dprocardinoides	0.764	0.584	
Dsechellia5 X	Dneocardini	0.794	0.645	
Dsechellia5 X	Dpolymorpha	0.766	0.627	
Dsechellia5 X	Dcardinoides	0.786	0.610	
Dsechellia7 X	Derecta1	0.195	0.257	
Dsechellia7 X	Dsimulans1	0.000	0.080	***
Dsechellia7 X	Dsimulans2	0.000	0.080	***
Dsechellia7 X	Dsimulans3	0.018	0.080	**
Dsechellia7 X	sim1	0.200	0.080	
Dsechellia7 X	Dsimulans5	0.030	0.080	*
Dsechellia7 X	Dsimulans6	0.030	0.080	*
Dsechellia7 X	Dsimulans7	0.026	0.080	*
Dsechellia7 X	Dsimulans8	0.016	0.080	**
Dsechellia7 X	Dsimulans9	0.225	0.080	
Dsechellia7 X	Dm11	0.025	0.160	***
Dsechellia7 X	Dmelanogaster1	0.012	0.160	***
Dsechellia7 X	mel1	0.014	0.160	***
Dsechellia7 X	mel2	0.012	0.160	***
Dsechellia7 X	Dmelanogaster5	0.012	0.160	***
Dsechellia7 X	Dmelanogaster6	0.019	0.160	***
Dsechellia7 X	Dmelanogaster9	0.204	0.160	
Dsechellia7 X	yak1	0.227	0.280	
Dsechellia7 X	Dwillistoni1	0.772	0.798	
Dsechellia7 X	willi1	0.868	0.798	
Dsechellia7 X	dhMiF2	0.767	0.775	
Dsechellia7 X	Dmercatorum	0.765	0.647	
Dsechellia7 X	Dbuzzatii	0.784	0.742	
Dsechellia7 X	Dzottii	0.786	0.697	
Dsechellia7 X	Dacutilabella	0.782	0.592	
Dsechellia7 X	Dcardini1	0.802	0.639	
Dsechellia7 X	Dcardini2	0.799	0.639	
Dsechellia7 X	Dcardini3	0.799	0.639	
Dsechellia7 X	Dneomorpha	0.792	0.616	
Dsechellia7 X	parth1	0.776	0.624	
Dsechellia7 X	Dparthenogenetica4	0.802	0.624	
Dsechellia7 X	Dparthenogenetica5	0.797	0.624	
Dsechellia7 X	Dparthenogenetica6	0.803	0.624	
Dsechellia7 X	Dprocardinoides	0.773	0.584	
Dsechellia7 X	Dneocardini	0.800	0.645	

Dsechellia7 X	Dpolymorpha	0.774	0.627	
Dsechellia7 X	Dcardinoides	0.793	0.610	
Dsechellia8 X	Derecta1	0.195	0.257	
Dsechellia8 X	Dsimulans1	0.000	0.080	***
Dsechellia8 X	Dsimulans2	0.000	0.080	***
Dsechellia8 X	Dsimulans3	0.018	0.080	**
Dsechellia8 X	sim1	0.200	0.080	
Dsechellia8 X	Dsimulans5	0.030	0.080	*
Dsechellia8 X	Dsimulans6	0.030	0.080	*
Dsechellia8 X	Dsimulans7	0.026	0.080	*
Dsechellia8 X	Dsimulans8	0.016	0.080	**
Dsechellia8 X	Dsimulans9	0.225	0.080	
Dsechellia8 X	Dm11	0.025	0.160	***
Dsechellia8 X	Dmelanogaster1	0.012	0.160	***
Dsechellia8 X	mel1	0.014	0.160	***
Dsechellia8 X	mel2	0.012	0.160	***
Dsechellia8 X	Dmelanogaster5	0.012	0.160	***
Dsechellia8 X	Dmelanogaster6	0.019	0.160	***
Dsechellia8 X	Dmelanogaster9	0.204	0.160	
Dsechellia8 X	yak1	0.227	0.280	
Dsechellia8 X	Dwillistoni1	0.772	0.798	
Dsechellia8 X	willi1	0.868	0.798	
Dsechellia8 X	dhMiF2	0.767	0.775	
Dsechellia8 X	Dmercatorum	0.765	0.647	
Dsechellia8 X	Dbuzzatii	0.784	0.742	
Dsechellia8 X	Dzottii	0.786	0.697	
Dsechellia8 X	Dacutilabella	0.782	0.592	
Dsechellia8 X	Dcardini1	0.802	0.639	
Dsechellia8 X	Dcardini2	0.799	0.639	
Dsechellia8 X	Dcardini3	0.799	0.639	
Dsechellia8 X	Dneomorpha	0.792	0.616	
Dsechellia8 X	parth1	0.776	0.624	
Dsechellia8 X	Dparthenogenetica4	0.802	0.624	
Dsechellia8 X	Dparthenogenetica5	0.797	0.624	
Dsechellia8 X	Dparthenogenetica6	0.803	0.624	
Dsechellia8 X	Dprocardinoides	0.773	0.584	
Dsechellia8 X	Dneocardini	0.800	0.645	
Dsechellia8 X	Dpolymorpha	0.774	0.627	
Dsechellia8 X	Dcardinoides	0.793	0.610	
Dsechellia9 X	Derecta1	0.180	0.257	*
Dsechellia9 X	Dsimulans1	0.000	0.080	***
Dsechellia9 X	Dsimulans2	0.000	0.080	***
Dsechellia9 X	Dsimulans3	0.018	0.080	**
Dsechellia9 X	sim1	0.188	0.080	
Dsechellia9 X	Dsimulans5	0.031	0.080	*
Dsechellia9 X	Dsimulans6	0.031	0.080	*
Dsechellia9 X	Dsimulans7	0.027	0.080	*
Dsechellia9 X	Dsimulans8	0.017	0.080	*
Dsechellia9 X	Dsimulans9	0.218	0.080	
Dsechellia9 X	Dm11	0.026	0.160	***
Dsechellia9 X	Dmelanogaster1	0.012	0.160	***
Dsechellia9 X	mel1	0.014	0.160	***

Dsechellia9 X	mel2	0.012	0.160	***
Dsechellia9 X	Dmelanogaster5	0.013	0.160	***
Dsechellia9 X	Dmelanogaster6	0.019	0.160	***
Dsechellia9 X	Dmelanogaster9	0.204	0.160	
Dsechellia9 X	yak1	0.224	0.280	
Dsechellia9 X	Dwillistoni1	0.762	0.798	
Dsechellia9 X	willi1	0.860	0.798	
Dsechellia9 X	dhMiF2	0.757	0.775	
Dsechellia9 X	Dmercatorum	0.758	0.647	
Dsechellia9 X	Dbuzzatii	0.774	0.742	
Dsechellia9 X	Dzottii	0.785	0.697	
Dsechellia9 X	Dacutilabella	0.773	0.592	
Dsechellia9 X	Dcardini1	0.793	0.639	
Dsechellia9 X	Dcardini2	0.775	0.639	
Dsechellia9 X	Dcardini3	0.775	0.639	
Dsechellia9 X	Dneomorpha	0.783	0.616	
Dsechellia9 X	parth1	0.768	0.624	
Dsechellia9 X	Dparthenogenetica4	0.793	0.624	
Dsechellia9 X	Dparthenogenetica5	0.787	0.624	
Dsechellia9 X	Dparthenogenetica6	0.794	0.624	
Dsechellia9 X	Dprocardinoides	0.763	0.584	
Dsechellia9 X	Dneocardini	0.796	0.645	
Dsechellia9 X	Dpolymorpha	0.769	0.627	
Dsechellia9 X	Dcardinoides	0.789	0.610	
sech2 X	Derecta1	0.196	0.257	
sech2 X	Dsimulans1	0.017	0.080	**
sech2 X	Dsimulans2	0.017	0.080	**
sech2 X	Dsimulans3	0.002	0.080	***
sech2 X	sim1	0.211	0.080	
sech2 X	Dsimulans5	0.014	0.080	**
sech2 X	Dsimulans6	0.026	0.080	*
sech2 X	Dsimulans7	0.025	0.080	*
sech2 X	Dsimulans8	0.016	0.080	**
sech2 X	Dsimulans9	0.220	0.080	
sech2 X	Dm11	0.021	0.160	***
sech2 X	Dmelanogaster1	0.008	0.160	***
sech2 X	mel1	0.016	0.160	***
sech2 X	mel2	0.008	0.160	***
sech2 X	Dmelanogaster5	0.002	0.160	***
sech2 X	Dmelanogaster6	0.015	0.160	***
sech2 X	Dmelanogaster9	0.205	0.160	
sech2 X	yak1	0.225	0.280	
sech2 X	Dwillistoni1	0.763	0.798	
sech2 X	willi1	0.864	0.798	
sech2 X	dhMiF2	0.758	0.775	
sech2 X	Dmercatorum	0.755	0.647	
sech2 X	Dbuzzatii	0.775	0.742	
sech2 X	Dzottii	0.772	0.697	
sech2 X	Dacutilabella	0.775	0.592	
sech2 X	Dcardini1	0.794	0.639	
sech2 X	Dcardini2	0.799	0.639	
sech2 X	Dcardini3	0.799	0.639	

sech2 X	Dneomorpha	0.783	0.616	
sech2 X	parth1	0.772	0.624	
sech2 X	Dparthenogenetica4	0.794	0.624	
sech2 X	Dparthenogenetica5	0.788	0.624	
sech2 X	Dparthenogenetica6	0.794	0.624	
sech2 X	Dprocardinoides	0.761	0.584	
sech2 X	Dneocardini	0.791	0.645	
sech2 X	Dpolymorpha	0.766	0.627	
sech2 X	Dcardinoides	0.784	0.610	
Dsechellia13 X	Derecta1	0.195	0.257	
Dsechellia13 X	Dsimulans1	0.000	0.080	***
Dsechellia13 X	Dsimulans2	0.000	0.080	***
Dsechellia13 X	Dsimulans3	0.018	0.080	**
Dsechellia13 X	sim1	0.200	0.080	
Dsechellia13 X	Dsimulans5	0.030	0.080	*
Dsechellia13 X	Dsimulans6	0.030	0.080	*
Dsechellia13 X	Dsimulans7	0.026	0.080	*
Dsechellia13 X	Dsimulans8	0.016	0.080	**
Dsechellia13 X	Dsimulans9	0.225	0.080	***
Dsechellia13 X	Dm11	0.025	0.160	***
Dsechellia13 X	Dmelanogaster1	0.012	0.160	***
Dsechellia13 X	mel1	0.014	0.160	***
Dsechellia13 X	mel2	0.012	0.160	***
Dsechellia13 X	Dmelanogaster5	0.012	0.160	***
Dsechellia13 X	Dmelanogaster6	0.019	0.160	***
Dsechellia13 X	Dmelanogaster9	0.204	0.160	
Dsechellia13 X	yak1	0.227	0.280	
Dsechellia13 X	Dwillistoni1	0.772	0.798	
Dsechellia13 X	willi1	0.868	0.798	
Dsechellia13 X	dhMiF2	0.767	0.775	
Dsechellia13 X	Dmercatorum	0.765	0.647	
Dsechellia13 X	Dbuzzatii	0.784	0.742	
Dsechellia13 X	Dzottii	0.786	0.697	
Dsechellia13 X	Dacutilabella	0.782	0.592	
Dsechellia13 X	Dcardini1	0.802	0.639	
Dsechellia13 X	Dcardini2	0.799	0.639	
Dsechellia13 X	Dcardini3	0.799	0.639	
Dsechellia13 X	Dneomorpha	0.792	0.616	
Dsechellia13 X	parth1	0.776	0.624	
Dsechellia13 X	Dparthenogenetica4	0.802	0.624	
Dsechellia13 X	Dparthenogenetica5	0.797	0.624	
Dsechellia13 X	Dparthenogenetica6	0.803	0.624	
Dsechellia13 X	Dprocardinoides	0.773	0.584	
Dsechellia13 X	Dneocardini	0.800	0.645	
Dsechellia13 X	Dpolymorpha	0.774	0.627	
Dsechellia13 X	Dcardinoides	0.793	0.610	
Dsechellia15 X	Derecta1	0.195	0.257	
Dsechellia15 X	Dsimulans1	0.000	0.080	***
Dsechellia15 X	Dsimulans2	0.000	0.080	***
Dsechellia15 X	Dsimulans3	0.018	0.080	**
Dsechellia15 X	sim1	0.200	0.080	
Dsechellia15 X	Dsimulans5	0.030	0.080	*

Dsechellia15 X	Dsimulans6	0.030	0.080	*
Dsechellia15 X	Dsimulans7	0.026	0.080	*
Dsechellia15 X	Dsimulans8	0.016	0.080	**
Dsechellia15 X	Dsimulans9	0.225	0.080	
Dsechellia15 X	Dm11	0.025	0.160	***
Dsechellia15 X	Dmelanogaster1	0.012	0.160	***
Dsechellia15 X	mel1	0.014	0.160	***
Dsechellia15 X	mel2	0.012	0.160	***
Dsechellia15 X	Dmelanogaster5	0.012	0.160	***
Dsechellia15 X	Dmelanogaster6	0.019	0.160	***
Dsechellia15 X	Dmelanogaster9	0.204	0.160	
Dsechellia15 X	yak1	0.227	0.280	
Dsechellia15 X	Dwillistoni1	0.772	0.798	
Dsechellia15 X	willi1	0.868	0.798	
Dsechellia15 X	dhMiF2	0.767	0.775	
Dsechellia15 X	Dmercatorum	0.765	0.647	
Dsechellia15 X	Dbuzzatii	0.784	0.742	
Dsechellia15 X	Dzottii	0.786	0.697	
Dsechellia15 X	Dacutilabella	0.782	0.592	
Dsechellia15 X	Dcardini1	0.802	0.639	
Dsechellia15 X	Dcardini2	0.799	0.639	
Dsechellia15 X	Dcardini3	0.799	0.639	
Dsechellia15 X	Dneomorpha	0.792	0.616	
Dsechellia15 X	parth1	0.776	0.624	
Dsechellia15 X	Dparthenogenetica4	0.802	0.624	
Dsechellia15 X	Dparthenogenetica5	0.797	0.624	
Dsechellia15 X	Dparthenogenetica6	0.803	0.624	
Dsechellia15 X	Dprocardinoides	0.773	0.584	
Dsechellia15 X	Dneocardini	0.800	0.645	
Dsechellia15 X	Dpolymorpha	0.774	0.627	
Dsechellia15 X	Dcardinoides	0.793	0.610	
Dsechellia18 X	Derecta1	0.195	0.257	
Dsechellia18 X	Dsimulans1	0.000	0.080	***
Dsechellia18 X	Dsimulans2	0.000	0.080	***
Dsechellia18 X	Dsimulans3	0.018	0.080	**
Dsechellia18 X	sim1	0.200	0.080	
Dsechellia18 X	Dsimulans5	0.030	0.080	*
Dsechellia18 X	Dsimulans6	0.030	0.080	*
Dsechellia18 X	Dsimulans7	0.026	0.080	*
Dsechellia18 X	Dsimulans8	0.016	0.080	**
Dsechellia18 X	Dsimulans9	0.225	0.080	
Dsechellia18 X	Dm11	0.025	0.160	***
Dsechellia18 X	Dmelanogaster1	0.012	0.160	***
Dsechellia18 X	mel1	0.014	0.160	***
Dsechellia18 X	mel2	0.012	0.160	***
Dsechellia18 X	Dmelanogaster5	0.012	0.160	***
Dsechellia18 X	Dmelanogaster6	0.019	0.160	***
Dsechellia18 X	Dmelanogaster9	0.204	0.160	
Dsechellia18 X	yak1	0.227	0.280	
Dsechellia18 X	Dwillistoni1	0.772	0.798	
Dsechellia18 X	willi1	0.868	0.798	
Dsechellia18 X	dhMiF2	0.767	0.775	

Dsechellia18 X	Dmercatorum	0.765	0.647	
Dsechellia18 X	Dbuzzatii	0.784	0.742	
Dsechellia18 X	Dzottii	0.786	0.697	
Dsechellia18 X	Dacutilabella	0.782	0.592	
Dsechellia18 X	Dcardini1	0.802	0.639	
Dsechellia18 X	Dcardini2	0.799	0.639	
Dsechellia18 X	Dcardini3	0.799	0.639	
Dsechellia18 X	Dneomorpha	0.792	0.616	
Dsechellia18 X	parth1	0.776	0.624	
Dsechellia18 X	Dparthenogenetica4	0.802	0.624	
Dsechellia18 X	Dparthenogenetica5	0.797	0.624	
Dsechellia18 X	Dparthenogenetica6	0.803	0.624	
Dsechellia18 X	Dprocardinoides	0.773	0.584	
Dsechellia18 X	Dneocardini	0.800	0.645	
Dsechellia18 X	Dpolymorpha	0.774	0.627	
Dsechellia18 X	Dcardinoides	0.793	0.610	
sech3 X	Derecta1	0.195	0.257	
sech3 X	Dsimulans1	0.035	0.080	*
sech3 X	Dsimulans2	0.035	0.080	*
sech3 X	Dsimulans3	0.016	0.080	**
sech3 X	sim1	0.209	0.080	
sech3 X	Dsimulans5	0.029	0.080	*
sech3 X	Dsimulans6	0.041	0.080	
sech3 X	Dsimulans7	0.042	0.080	
sech3 X	Dsimulans8	0.032	0.080	
sech3 X	Dsimulans9	0.224	0.080	
sech3 X	Dm11	0.037	0.160	***
sech3 X	Dmelanogaster1	0.023	0.160	***
sech3 X	mel1	0.033	0.160	***
sech3 X	mel2	0.023	0.160	***
sech3 X	Dmelanogaster5	0.017	0.160	***
sech3 X	Dmelanogaster6	0.025	0.160	***
sech3 X	Dmelanogaster9	0.206	0.160	
sech3 X	yak1	0.227	0.280	
sech3 X	Dwillistoni1	0.761	0.798	
sech3 X	willi1	0.866	0.798	
sech3 X	dhMiF2	0.756	0.775	
sech3 X	Dmercatorum	0.755	0.647	
sech3 X	Dbuzzatii	0.773	0.742	
sech3 X	Dzottii	0.773	0.697	
sech3 X	Dacutilabella	0.770	0.592	
sech3 X	Dcardini1	0.793	0.639	
sech3 X	Dcardini2	0.784	0.639	
sech3 X	Dcardini3	0.784	0.639	
sech3 X	Dneomorpha	0.781	0.616	
sech3 X	parth1	0.770	0.624	
sech3 X	Dparthenogenetica4	0.793	0.624	
sech3 X	Dparthenogenetica5	0.789	0.624	
sech3 X	Dparthenogenetica6	0.794	0.624	
sech3 X	Dprocardinoides	0.756	0.584	
sech3 X	Dneocardini	0.795	0.645	
sech3 X	Dpolymorpha	0.764	0.627	

sech3 X	Dcardinoides	0.780	0.610	
Dsechellia22 X	Derecta1	0.202	0.257	
Dsechellia22 X	Dsimulans1	0.006	0.080	***
Dsechellia22 X	Dsimulans2	0.006	0.080	***
Dsechellia22 X	Dsimulans3	0.024	0.080	*
Dsechellia22 X	sim1	0.211	0.080	
Dsechellia22 X	Dsimulans5	0.036	0.080	
Dsechellia22 X	Dsimulans6	0.036	0.080	
Dsechellia22 X	Dsimulans7	0.033	0.080	*
Dsechellia22 X	Dsimulans8	0.024	0.080	*
Dsechellia22 X	Dsimulans9	0.231	0.080	
Dsechellia22 X	Dm11	0.031	0.160	***
Dsechellia22 X	Dmelanogaster1	0.018	0.160	***
Dsechellia22 X	mel1	0.021	0.160	***
Dsechellia22 X	mel2	0.018	0.160	***
Dsechellia22 X	Dmelanogaster5	0.018	0.160	***
Dsechellia22 X	Dmelanogaster6	0.026	0.160	***
Dsechellia22 X	Dmelanogaster9	0.204	0.160	
Dsechellia22 X	yak1	0.233	0.280	
Dsechellia22 X	Dwillistoni1	0.779	0.798	
Dsechellia22 X	willi1	0.864	0.798	
Dsechellia22 X	dhMiF2	0.774	0.775	
Dsechellia22 X	Dmercatorum	0.772	0.647	
Dsechellia22 X	Dbuzzatii	0.791	0.742	
Dsechellia22 X	Dzottii	0.790	0.697	
Dsechellia22 X	Dacutilabella	0.790	0.592	
Dsechellia22 X	Dcardini1	0.810	0.639	
Dsechellia22 X	Dcardini2	0.817	0.639	
Dsechellia22 X	Dcardini3	0.817	0.639	
Dsechellia22 X	Dneomorpha	0.800	0.616	
Dsechellia22 X	parth1	0.784	0.624	
Dsechellia22 X	Dparthenogenetica4	0.810	0.624	
Dsechellia22 X	Dparthenogenetica5	0.804	0.624	
Dsechellia22 X	Dparthenogenetica6	0.810	0.624	
Dsechellia22 X	Dprocardinoides	0.780	0.584	
Dsechellia22 X	Dneocardini	0.807	0.645	
Dsechellia22 X	Dpolymorpha	0.783	0.627	
Dsechellia22 X	Dcardinoides	0.801	0.610	
Dsechellia26 X	Derecta1	0.196	0.257	
Dsechellia26 X	Dsimulans1	0.026	0.080	*
Dsechellia26 X	Dsimulans2	0.026	0.080	*
Dsechellia26 X	Dsimulans3	0.006	0.080	***
Dsechellia26 X	sim1	0.210	0.080	
Dsechellia26 X	Dsimulans5	0.019	0.080	**
Dsechellia26 X	Dsimulans6	0.026	0.080	*
Dsechellia26 X	Dsimulans7	0.021	0.080	*
Dsechellia26 X	Dsimulans8	0.025	0.080	*
Dsechellia26 X	Dsimulans9	0.218	0.080	
Dsechellia26 X	Dm11	0.020	0.160	***
Dsechellia26 X	Dmelanogaster1	0.013	0.160	***
Dsechellia26 X	mel1	0.019	0.160	***
Dsechellia26 X	mel2	0.013	0.160	***

Dsechellia26 X	Dmelanogaster5	0.007	0.160	***
Dsechellia26 X	Dmelanogaster6	0.021	0.160	***
Dsechellia26 X	Dmelanogaster9	0.205	0.160	
Dsechellia26 X	yak1	0.222	0.280	
Dsechellia26 X	Dwillistoni1	0.767	0.798	
Dsechellia26 X	willi1	0.843	0.798	
Dsechellia26 X	dhMiF2	0.754	0.775	
Dsechellia26 X	Dmercatorum	0.759	0.647	
Dsechellia26 X	Dbuzzatii	0.783	0.742	
Dsechellia26 X	Dzottii	0.773	0.697	
Dsechellia26 X	Dacutilabella	0.778	0.592	
Dsechellia26 X	Dcardini1	0.800	0.639	
Dsechellia26 X	Dcardini2	0.800	0.639	
Dsechellia26 X	Dcardini3	0.800	0.639	
Dsechellia26 X	Dneomorpha	0.787	0.616	
Dsechellia26 X	parth1	0.776	0.624	
Dsechellia26 X	Dparthenogenetica4	0.800	0.624	
Dsechellia26 X	Dparthenogenetica5	0.794	0.624	
Dsechellia26 X	Dparthenogenetica6	0.800	0.624	
Dsechellia26 X	Dprocardinoides	0.762	0.584	
Dsechellia26 X	Dneocardini	0.797	0.645	
Dsechellia26 X	Dpolymorpha	0.787	0.627	
Dsechellia26 X	Dcardinoides	0.789	0.610	
Dsechellia27 X	Derecta1	0.165	0.257	*
Dsechellia27 X	Dsimulans1	0.222	0.080	
Dsechellia27 X	Dsimulans2	0.222	0.080	
Dsechellia27 X	Dsimulans3	0.216	0.080	
Dsechellia27 X	sim1	0.102	0.080	
Dsechellia27 X	Dsimulans5	0.227	0.080	
Dsechellia27 X	Dsimulans6	0.230	0.080	
Dsechellia27 X	Dsimulans7	0.224	0.080	
Dsechellia27 X	Dsimulans8	0.241	0.080	
Dsechellia27 X	Dsimulans9	0.037	0.080	
Dsechellia27 X	Dm11	0.236	0.160	
Dsechellia27 X	Dmelanogaster1	0.222	0.160	
Dsechellia27 X	mel1	0.230	0.160	
Dsechellia27 X	mel2	0.214	0.160	
Dsechellia27 X	Dmelanogaster5	0.221	0.160	
Dsechellia27 X	Dmelanogaster6	0.220	0.160	
Dsechellia27 X	Dmelanogaster9	0.089	0.160	
Dsechellia27 X	yak1	0.134	0.280	***
Dsechellia27 X	Dwillistoni1	0.771	0.798	
Dsechellia27 X	willi1	0.841	0.798	
Dsechellia27 X	dhMiF2	0.764	0.775	
Dsechellia27 X	Dmercatorum	0.772	0.647	
Dsechellia27 X	Dbuzzatii	0.785	0.742	
Dsechellia27 X	Dzottii	0.807	0.697	
Dsechellia27 X	Dacutilabella	0.768	0.592	
Dsechellia27 X	Dcardini1	0.795	0.639	
Dsechellia27 X	Dcardini2	0.772	0.639	
Dsechellia27 X	Dcardini3	0.772	0.639	
Dsechellia27 X	Dneomorpha	0.787	0.616	

Dsechellia27 X	parth1	0.798	0.624	
Dsechellia27 X	Dparthenogenetica4	0.795	0.624	
Dsechellia27 X	Dparthenogenetica5	0.784	0.624	
Dsechellia27 X	Dparthenogenetica6	0.790	0.624	
Dsechellia27 X	Dprocardinoides	0.789	0.584	
Dsechellia27 X	Dneocardini	0.805	0.645	
Dsechellia27 X	Dpolymorpha	0.764	0.627	
Dsechellia27 X	Dcardinoides	0.804	0.610	
Dsechellia28 X	Derecta1	0.218	0.257	
Dsechellia28 X	Dsimulans1	0.048	0.080	
Dsechellia28 X	Dsimulans2	0.048	0.080	
Dsechellia28 X	Dsimulans3	0.039	0.080	
Dsechellia28 X	sim1	0.196	0.080	
Dsechellia28 X	Dsimulans5	0.058	0.080	
Dsechellia28 X	Dsimulans6	0.059	0.080	
Dsechellia28 X	Dsimulans7	0.048	0.080	
Dsechellia28 X	Dsimulans8	0.058	0.080	
Dsechellia28 X	Dsimulans9	0.197	0.080	
Dsechellia28 X	Dm11	0.052	0.160	**
Dsechellia28 X	Dmelanogaster1	0.048	0.160	**
Dsechellia28 X	mel1	0.058	0.160	**
Dsechellia28 X	mel2	0.048	0.160	**
Dsechellia28 X	Dmelanogaster5	0.040	0.160	**
Dsechellia28 X	Dmelanogaster6	0.050	0.160	**
Dsechellia28 X	Dmelanogaster9	0.114	0.160	
Dsechellia28 X	yak1	0.213	0.280	
Dsechellia28 X	Dwillistoni1	0.755	0.798	
Dsechellia28 X	willi1	0.817	0.798	
Dsechellia28 X	dhMiF2	0.741	0.775	
Dsechellia28 X	Dmercatorum	0.752	0.647	
Dsechellia28 X	Dbuzzatii	0.773	0.742	
Dsechellia28 X	Dzottii	0.781	0.697	
Dsechellia28 X	Dacutilabella	0.772	0.592	
Dsechellia28 X	Dcardini1	0.800	0.639	
Dsechellia28 X	Dcardini2	0.784	0.639	
Dsechellia28 X	Dcardini3	0.784	0.639	
Dsechellia28 X	Dneomorpha	0.787	0.616	
Dsechellia28 X	parth1	0.774	0.624	
Dsechellia28 X	Dparthenogenetica4	0.800	0.624	
Dsechellia28 X	Dparthenogenetica5	0.791	0.624	
Dsechellia28 X	Dparthenogenetica6	0.800	0.624	
Dsechellia28 X	Dprocardinoides	0.758	0.584	
Dsechellia28 X	Dneocardini	0.788	0.645	
Dsechellia28 X	Dpolymorpha	0.788	0.627	
Dsechellia28 X	Dcardinoides	0.779	0.610	
Dsechellia29 X	Derecta1	0.157	0.257	*
Dsechellia29 X	Dsimulans1	0.222	0.080	
Dsechellia29 X	Dsimulans2	0.222	0.080	
Dsechellia29 X	Dsimulans3	0.217	0.080	
Dsechellia29 X	sim1	0.095	0.080	
Dsechellia29 X	Dsimulans5	0.228	0.080	
Dsechellia29 X	Dsimulans6	0.231	0.080	

Dsechellia29 X	Dsimulans7	0.225	0.080	
Dsechellia29 X	Dsimulans8	0.234	0.080	
Dsechellia29 X	Dsimulans9	0.073	0.080	
Dsechellia29 X	Dm11	0.238	0.160	
Dsechellia29 X	Dmelanogaster1	0.222	0.160	
Dsechellia29 X	mel1	0.231	0.160	
Dsechellia29 X	mel2	0.214	0.160	
Dsechellia29 X	Dmelanogaster5	0.218	0.160	
Dsechellia29 X	Dmelanogaster6	0.221	0.160	
Dsechellia29 X	Dmelanogaster9	0.076	0.160	*
Dsechellia29 X	yak1	0.139	0.280	**
Dsechellia29 X	Dwillistoni1	0.742	0.798	
Dsechellia29 X	willi1	0.871	0.798	
Dsechellia29 X	dhMiF2	0.744	0.775	
Dsechellia29 X	Dmercatorum	0.752	0.647	
Dsechellia29 X	Dbuzzatii	0.757	0.742	
Dsechellia29 X	Dzottii	0.791	0.697	
Dsechellia29 X	Dacutilabella	0.764	0.592	
Dsechellia29 X	Dcardini1	0.786	0.639	
Dsechellia29 X	Dcardini2	0.779	0.639	
Dsechellia29 X	Dcardini3	0.779	0.639	
Dsechellia29 X	Dneomorpha	0.777	0.616	
Dsechellia29 X	parth1	0.775	0.624	
Dsechellia29 X	Dparthenogenetica4	0.786	0.624	
Dsechellia29 X	Dparthenogenetica5	0.773	0.624	
Dsechellia29 X	Dparthenogenetica6	0.786	0.624	
Dsechellia29 X	Dprocardinoides	0.777	0.584	
Dsechellia29 X	Dneocardini	0.796	0.645	
Dsechellia29 X	Dpolymorpha	0.757	0.627	
Dsechellia29 X	Dcardinoides	0.794	0.610	
Derecta1 X	Dsimulans1	0.195	0.223	
Derecta1 X	Dsimulans2	0.195	0.223	
Derecta1 X	Dsimulans3	0.196	0.223	
Derecta1 X	sim1	0.153	0.223	
Derecta1 X	Dsimulans5	0.207	0.223	
Derecta1 X	Dsimulans6	0.211	0.223	
Derecta1 X	Dsimulans7	0.204	0.223	
Derecta1 X	Dsimulans8	0.223	0.223	
Derecta1 X	Dsimulans9	0.162	0.223	
Derecta1 X	Dm11	0.225	0.271	
Derecta1 X	Dmelanogaster1	0.201	0.271	
Derecta1 X	mel1	0.214	0.271	
Derecta1 X	mel2	0.194	0.271	
Derecta1 X	Dmelanogaster5	0.194	0.271	*
Derecta1 X	Dmelanogaster6	0.190	0.271	*
Derecta1 X	Dmelanogaster9	0.117	0.271	***
Derecta1 X	yak1	0.178	0.155	
Derecta1 X	Dwillistoni1	0.724	0.780	
Derecta1 X	willi1	0.873	0.780	
Derecta1 X	dhMiF2	0.726	0.777	
Derecta1 X	Dmercatorum	0.726	0.620	
Derecta1 X	Dbuzzatii	0.741	0.745	

Derecta1 X	Dzottii	0.751	0.711	
Derecta1 X	Dacutilabella	0.743	0.610	
Derecta1 X	Dcardini1	0.764	0.687	
Derecta1 X	Dcardini2	0.770	0.687	
Derecta1 X	Dcardini3	0.770	0.687	
Derecta1 X	Dneomorpha	0.763	0.652	
Derecta1 X	parth1	0.748	0.656	
Derecta1 X	Dparthenogenetica4	0.764	0.656	
Derecta1 X	Dparthenogenetica5	0.761	0.656	
Derecta1 X	Dparthenogenetica6	0.773	0.656	
Derecta1 X	Dprocardinoides	0.745	0.607	
Derecta1 X	Dneocardini	0.781	0.682	
Derecta1 X	Dpolymorpha	0.775	0.672	
Derecta1 X	Dcardinoides	0.775	0.637	
Dsimulans1 X	Dm11	0.025	0.146	***
Dsimulans1 X	Dmelanogaster1	0.012	0.146	***
Dsimulans1 X	mel1	0.014	0.146	***
Dsimulans1 X	mel2	0.012	0.146	***
Dsimulans1 X	Dmelanogaster5	0.012	0.146	***
Dsimulans1 X	Dmelanogaster6	0.019	0.146	***
Dsimulans1 X	Dmelanogaster9	0.204	0.146	
Dsimulans1 X	yak1	0.227	0.259	
Dsimulans1 X	Dwillistoni1	0.772	0.771	
Dsimulans1 X	willi1	0.868	0.771	
Dsimulans1 X	dhMiF2	0.767	0.749	
Dsimulans1 X	Dmercatorum	0.765	0.619	
Dsimulans1 X	Dbuzzatii	0.784	0.738	
Dsimulans1 X	Dzottii	0.786	0.684	
Dsimulans1 X	Dacutilabella	0.782	0.586	
Dsimulans1 X	Dcardini1	0.802	0.649	
Dsimulans1 X	Dcardini2	0.799	0.649	
Dsimulans1 X	Dcardini3	0.799	0.649	
Dsimulans1 X	Dneomorpha	0.792	0.610	
Dsimulans1 X	parth1	0.776	0.632	
Dsimulans1 X	Dparthenogenetica4	0.802	0.632	
Dsimulans1 X	Dparthenogenetica5	0.797	0.632	
Dsimulans1 X	Dparthenogenetica6	0.803	0.632	
Dsimulans1 X	Dprocardinoides	0.773	0.588	
Dsimulans1 X	Dneocardini	0.800	0.640	
Dsimulans1 X	Dpolymorpha	0.774	0.626	
Dsimulans1 X	Dcardinoides	0.793	0.609	
Dsimulans2 X	Dm11	0.025	0.146	***
Dsimulans2 X	Dmelanogaster1	0.012	0.146	***
Dsimulans2 X	mel1	0.014	0.146	***
Dsimulans2 X	mel2	0.012	0.146	***
Dsimulans2 X	Dmelanogaster5	0.012	0.146	***
Dsimulans2 X	Dmelanogaster6	0.019	0.146	***
Dsimulans2 X	Dmelanogaster9	0.204	0.146	
Dsimulans2 X	yak1	0.227	0.259	
Dsimulans2 X	Dwillistoni1	0.772	0.771	
Dsimulans2 X	willi1	0.868	0.771	
Dsimulans2 X	dhMiF2	0.767	0.749	

Dsimulans2 X	Dmercatorum	0.765	0.619	
Dsimulans2 X	Dbuzzatii	0.784	0.738	
Dsimulans2 X	Dzottii	0.786	0.684	
Dsimulans2 X	Dacutilabella	0.782	0.586	
Dsimulans2 X	Dcardini1	0.802	0.649	
Dsimulans2 X	Dcardini2	0.799	0.649	
Dsimulans2 X	Dcardini3	0.799	0.649	
Dsimulans2 X	Dneomorpha	0.792	0.610	
Dsimulans2 X	parth1	0.776	0.632	
Dsimulans2 X	Dparthenogenetica4	0.802	0.632	
Dsimulans2 X	Dparthenogenetica5	0.797	0.632	
Dsimulans2 X	Dparthenogenetica6	0.803	0.632	
Dsimulans2 X	Dprocardinoides	0.773	0.588	
Dsimulans2 X	Dneocardini	0.800	0.640	
Dsimulans2 X	Dpolymorpha	0.774	0.626	
Dsimulans2 X	Dcardinoides	0.793	0.609	
Dsimulans3 X	Dm11	0.019	0.146	***
Dsimulans3 X	Dmelanogaster1	0.006	0.146	***
Dsimulans3 X	mel1	0.015	0.146	***
Dsimulans3 X	mel2	0.006	0.146	***
Dsimulans3 X	Dmelanogaster5	0.000	0.146	***
Dsimulans3 X	Dmelanogaster6	0.013	0.146	***
Dsimulans3 X	Dmelanogaster9	0.205	0.146	
Dsimulans3 X	yak1	0.223	0.259	
Dsimulans3 X	Dwillistoni1	0.768	0.771	
Dsimulans3 X	willi1	0.871	0.771	
Dsimulans3 X	dhMiF2	0.763	0.749	
Dsimulans3 X	Dmercatorum	0.758	0.619	
Dsimulans3 X	Dbuzzatii	0.780	0.738	
Dsimulans3 X	Dzottii	0.775	0.684	
Dsimulans3 X	Dacutilabella	0.778	0.586	
Dsimulans3 X	Dcardini1	0.799	0.649	
Dsimulans3 X	Dcardini2	0.799	0.649	
Dsimulans3 X	Dcardini3	0.799	0.649	
Dsimulans3 X	Dneomorpha	0.788	0.610	
Dsimulans3 X	parth1	0.777	0.632	
Dsimulans3 X	Dparthenogenetica4	0.799	0.632	
Dsimulans3 X	Dparthenogenetica5	0.793	0.632	
Dsimulans3 X	Dparthenogenetica6	0.799	0.632	
Dsimulans3 X	Dprocardinoides	0.766	0.588	
Dsimulans3 X	Dneocardini	0.796	0.640	
Dsimulans3 X	Dpolymorpha	0.769	0.626	
Dsimulans3 X	Dcardinoides	0.789	0.609	
sim1 X	Dm11	0.243	0.146	
sim1 X	Dmelanogaster1	0.210	0.146	
sim1 X	mel1	0.209	0.146	
sim1 X	mel2	0.197	0.146	
sim1 X	Dmelanogaster5	0.214	0.146	
sim1 X	Dmelanogaster6	0.199	0.146	
sim1 X	Dmelanogaster9	0.043	0.146	*
sim1 X	yak1	0.111	0.259	**
sim1 X	Dwillistoni1	0.683	0.771	

sim1 X	willi1	0.785	0.771	
sim1 X	dhMiF2	0.683	0.749	
sim1 X	Dmercatorum	0.696	0.619	
sim1 X	Dbuzzatii	0.726	0.738	
sim1 X	Dzottii	0.745	0.684	
sim1 X	Dacutilabella	0.725	0.586	
sim1 X	Dcardini1	0.757	0.649	
sim1 X	Dcardini2	0.768	0.649	
sim1 X	Dcardini3	0.768	0.649	
sim1 X	Dneomorpha	0.747	0.610	
sim1 X	parth1	0.730	0.632	
sim1 X	Dparthenogenetica4	0.757	0.632	
sim1 X	Dparthenogenetica5	0.735	0.632	
sim1 X	Dparthenogenetica6	0.760	0.632	
sim1 X	Dprocardinoides	0.733	0.588	
sim1 X	Dneocardini	0.762	0.640	
sim1 X	Dpolymorpha	0.777	0.626	
sim1 X	Dcardinoides	0.766	0.609	
Dsimulans5 X	Dm11	0.031	0.146	***
Dsimulans5 X	Dmelanogaster1	0.018	0.146	***
Dsimulans5 X	mel1	0.030	0.146	***
Dsimulans5 X	mel2	0.018	0.146	***
Dsimulans5 X	Dmelanogaster5	0.012	0.146	***
Dsimulans5 X	Dmelanogaster6	0.026	0.146	***
Dsimulans5 X	Dmelanogaster9	0.221	0.146	
Dsimulans5 X	yak1	0.234	0.259	
Dsimulans5 X	Dwillistoni1	0.764	0.771	
Dsimulans5 X	willi1	0.869	0.771	
Dsimulans5 X	dhMiF2	0.759	0.749	
Dsimulans5 X	Dmercatorum	0.757	0.619	
Dsimulans5 X	Dbuzzatii	0.776	0.738	
Dsimulans5 X	Dzottii	0.766	0.684	
Dsimulans5 X	Dacutilabella	0.774	0.586	
Dsimulans5 X	Dcardini1	0.795	0.649	
Dsimulans5 X	Dcardini2	0.813	0.649	
Dsimulans5 X	Dcardini3	0.813	0.649	
Dsimulans5 X	Dneomorpha	0.784	0.610	
Dsimulans5 X	parth1	0.773	0.632	
Dsimulans5 X	Dparthenogenetica4	0.795	0.632	
Dsimulans5 X	Dparthenogenetica5	0.789	0.632	
Dsimulans5 X	Dparthenogenetica6	0.795	0.632	
Dsimulans5 X	Dprocardinoides	0.762	0.588	
Dsimulans5 X	Dneocardini	0.792	0.640	
Dsimulans5 X	Dpolymorpha	0.764	0.626	
Dsimulans5 X	Dcardinoides	0.785	0.609	
Dsimulans6 X	Dm11	0.025	0.146	***
Dsimulans6 X	Dmelanogaster1	0.018	0.146	***
Dsimulans6 X	mel1	0.022	0.146	***
Dsimulans6 X	mel2	0.018	0.146	***
Dsimulans6 X	Dmelanogaster5	0.019	0.146	***
Dsimulans6 X	Dmelanogaster6	0.020	0.146	***
Dsimulans6 X	Dmelanogaster9	0.225	0.146	

Dsimulans6 X	yak1	0.238	0.259	
Dsimulans6 X	Dwillistoni1	0.768	0.771	
Dsimulans6 X	willi1	0.854	0.771	
Dsimulans6 X	dhMiF2	0.762	0.749	
Dsimulans6 X	Dmercatorum	0.760	0.619	
Dsimulans6 X	Dbuzzatii	0.778	0.738	
Dsimulans6 X	Dzottii	0.778	0.684	
Dsimulans6 X	Dacutilabella	0.778	0.586	
Dsimulans6 X	Dcardini1	0.798	0.649	
Dsimulans6 X	Dcardini2	0.788	0.649	
Dsimulans6 X	Dcardini3	0.788	0.649	
Dsimulans6 X	Dneomorpha	0.788	0.610	
Dsimulans6 X	parth1	0.772	0.632	
Dsimulans6 X	Dparthenogenetica4	0.798	0.632	
Dsimulans6 X	Dparthenogenetica5	0.792	0.632	
Dsimulans6 X	Dparthenogenetica6	0.799	0.632	
Dsimulans6 X	Dprocardinoides	0.765	0.588	
Dsimulans6 X	Dneocardini	0.796	0.640	
Dsimulans6 X	Dpolymorpha	0.777	0.626	
Dsimulans6 X	Dcardinoides	0.788	0.609	
Dsimulans7 X	Dm11	0.035	0.146	***
Dsimulans7 X	Dmelanogaster1	0.020	0.146	***
Dsimulans7 X	mel1	0.018	0.146	***
Dsimulans7 X	mel2	0.020	0.146	***
Dsimulans7 X	Dmelanogaster5	0.020	0.146	***
Dsimulans7 X	Dmelanogaster6	0.027	0.146	***
Dsimulans7 X	Dmelanogaster9	0.200	0.146	
Dsimulans7 X	yak1	0.226	0.259	
Dsimulans7 X	Dwillistoni1	0.773	0.771	
Dsimulans7 X	willi1	0.864	0.771	
Dsimulans7 X	dhMiF2	0.771	0.749	
Dsimulans7 X	Dmercatorum	0.767	0.619	
Dsimulans7 X	Dbuzzatii	0.785	0.738	
Dsimulans7 X	Dzottii	0.775	0.684	
Dsimulans7 X	Dacutilabella	0.784	0.586	
Dsimulans7 X	Dcardini1	0.807	0.649	
Dsimulans7 X	Dcardini2	0.799	0.649	
Dsimulans7 X	Dcardini3	0.799	0.649	
Dsimulans7 X	Dneomorpha	0.795	0.610	
Dsimulans7 X	parth1	0.779	0.632	
Dsimulans7 X	Dparthenogenetica4	0.807	0.632	
Dsimulans7 X	Dparthenogenetica5	0.801	0.632	
Dsimulans7 X	Dparthenogenetica6	0.807	0.632	
Dsimulans7 X	Dprocardinoides	0.777	0.588	
Dsimulans7 X	Dneocardini	0.804	0.640	
Dsimulans7 X	Dpolymorpha	0.789	0.626	
Dsimulans7 X	Dcardinoides	0.793	0.609	
Dsimulans8 X	Dm11	0.026	0.146	***
Dsimulans8 X	Dmelanogaster1	0.008	0.146	***
Dsimulans8 X	mel1	0.017	0.146	***
Dsimulans8 X	mel2	0.008	0.146	***
Dsimulans8 X	Dmelanogaster5	0.008	0.146	***

Dsimulans8 X	Dmelanogaster6	0.008	0.146	***
Dsimulans8 X	Dmelanogaster9	0.164	0.146	
Dsimulans8 X	yak1	0.232	0.259	
Dsimulans8 X	Dwillistoni1	0.738	0.771	
Dsimulans8 X	willi1	0.817	0.771	
Dsimulans8 X	dhMiF2	0.753	0.749	
Dsimulans8 X	Dmercatorum	0.753	0.619	
Dsimulans8 X	Dbuzzatii	0.763	0.738	
Dsimulans8 X	Dzottii	0.751	0.684	
Dsimulans8 X	Dacutilabella	0.776	0.586	
Dsimulans8 X	Dcardini1	0.802	0.649	
Dsimulans8 X	Dcardini2	0.813	0.649	
Dsimulans8 X	Dcardini3	0.813	0.649	
Dsimulans8 X	Dneomorpha	0.793	0.610	
Dsimulans8 X	parth1	0.771	0.632	
Dsimulans8 X	Dparthenogenetica4	0.802	0.632	
Dsimulans8 X	Dparthenogenetica5	0.795	0.632	
Dsimulans8 X	Dparthenogenetica6	0.803	0.632	
Dsimulans8 X	Dprocardinoides	0.776	0.588	
Dsimulans8 X	Dneocardini	0.793	0.640	
Dsimulans8 X	Dpolymorpha	0.804	0.626	
Dsimulans8 X	Dcardinoides	0.785	0.609	
Dsimulans9 X	Dm11	0.243	0.146	
Dsimulans9 X	Dmelanogaster1	0.225	0.146	
Dsimulans9 X	mel1	0.220	0.146	
Dsimulans9 X	mel2	0.217	0.146	
Dsimulans9 X	Dmelanogaster5	0.224	0.146	
Dsimulans9 X	Dmelanogaster6	0.217	0.146	
Dsimulans9 X	Dmelanogaster9	0.096	0.146	
Dsimulans9 X	yak1	0.132	0.259	**
Dsimulans9 X	Dwillistoni1	0.772	0.771	
Dsimulans9 X	willi1	0.832	0.771	
Dsimulans9 X	dhMiF2	0.765	0.749	
Dsimulans9 X	Dmercatorum	0.772	0.619	
Dsimulans9 X	Dbuzzatii	0.783	0.738	
Dsimulans9 X	Dzottii	0.802	0.684	
Dsimulans9 X	Dacutilabella	0.782	0.586	
Dsimulans9 X	Dcardini1	0.797	0.649	
Dsimulans9 X	Dcardini2	0.777	0.649	
Dsimulans9 X	Dcardini3	0.777	0.649	
Dsimulans9 X	Dneomorpha	0.789	0.610	
Dsimulans9 X	parth1	0.796	0.632	
Dsimulans9 X	Dparthenogenetica4	0.797	0.632	
Dsimulans9 X	Dparthenogenetica5	0.785	0.632	
Dsimulans9 X	Dparthenogenetica6	0.791	0.632	
Dsimulans9 X	Dprocardinoides	0.788	0.588	
Dsimulans9 X	Dneocardini	0.807	0.640	
Dsimulans9 X	Dpolymorpha	0.778	0.626	
Dsimulans9 X	Dcardinoides	0.806	0.609	
Dm11 X	yak1	0.252	0.285	
Dm11 X	Dwillistoni1	0.767	0.770	
Dm11 X	willi1	0.861	0.770	

Dm11 X	dhMiF2	0.762	0.778
Dm11 X	Dmercatorum	0.760	0.666
Dm11 X	Dbuzzatii	0.780	0.764
Dm11 X	Dzottii	0.783	0.727
Dm11 X	Dacutilabella	0.780	0.631
Dm11 X	Dcardini1	0.801	0.657
Dm11 X	Dcardini2	0.756	0.657
Dm11 X	Dcardini3	0.756	0.657
Dm11 X	Dneomorpha	0.791	0.611
Dm11 X	parth1	0.772	0.642
Dm11 X	Dparthenogenetica4	0.799	0.642
Dm11 X	Dparthenogenetica5	0.793	0.642
Dm11 X	Dparthenogenetica6	0.794	0.642
Dm11 X	Dprocardinoides	0.765	0.624
Dm11 X	Dneocardini	0.797	0.636
Dm11 X	Dpolymorpha	0.764	0.625
Dm11 X	Dcardinoides	0.791	0.636
Dmelanogaster1 X	yak1	0.229	0.285
Dmelanogaster1 X	Dwillistoni1	0.772	0.770
Dmelanogaster1 X	willi1	0.862	0.770
Dmelanogaster1 X	dhMiF2	0.767	0.778
Dmelanogaster1 X	Dmercatorum	0.765	0.666
Dmelanogaster1 X	Dbuzzatii	0.784	0.764
Dmelanogaster1 X	Dzottii	0.780	0.727
Dmelanogaster1 X	Dacutilabella	0.782	0.631
Dmelanogaster1 X	Dcardini1	0.802	0.657
Dmelanogaster1 X	Dcardini2	0.799	0.657
Dmelanogaster1 X	Dcardini3	0.799	0.657
Dmelanogaster1 X	Dneomorpha	0.792	0.611
Dmelanogaster1 X	parth1	0.776	0.642
Dmelanogaster1 X	Dparthenogenetica4	0.802	0.642
Dmelanogaster1 X	Dparthenogenetica5	0.797	0.642
Dmelanogaster1 X	Dparthenogenetica6	0.803	0.642
Dmelanogaster1 X	Dprocardinoides	0.770	0.624
Dmelanogaster1 X	Dneocardini	0.800	0.636
Dmelanogaster1 X	Dpolymorpha	0.774	0.625
Dmelanogaster1 X	Dcardinoides	0.793	0.636
mel1 X	yak1	0.226	0.285
mel1 X	Dwillistoni1	0.762	0.770
mel1 X	willi1	0.839	0.770
mel1 X	dhMiF2	0.760	0.778
mel1 X	Dmercatorum	0.753	0.666
mel1 X	Dbuzzatii	0.779	0.764
mel1 X	Dzottii	0.803	0.727
mel1 X	Dacutilabella	0.788	0.631
mel1 X	Dcardini1	0.799	0.657
mel1 X	Dcardini2	0.798	0.657
mel1 X	Dcardini3	0.798	0.657
mel1 X	Dneomorpha	0.787	0.611
mel1 X	parth1	0.767	0.642
mel1 X	Dparthenogenetica4	0.799	0.642
mel1 X	Dparthenogenetica5	0.791	0.642

mel1 X	Dparthenogenetica6	0.799	0.642
mel1 X	Dprocardinoides	0.762	0.624
mel1 X	Dneocardini	0.792	0.636
mel1 X	Dpolymorpha	0.778	0.625
mel1 X	Dcardinoides	0.784	0.636
mel2 X	yak1	0.221	0.285
mel2 X	Dwillistoni1	0.776	0.770
mel2 X	willi1	0.867	0.770
mel2 X	dhMiF2	0.764	0.778
mel2 X	Dmercatorum	0.769	0.666
mel2 X	Dbuzzatii	0.788	0.764
mel2 X	Dzottii	0.786	0.727
mel2 X	Dacutilabella	0.786	0.631
mel2 X	Dcardini1	0.806	0.657
mel2 X	Dcardini2	0.799	0.657
mel2 X	Dcardini3	0.799	0.657
mel2 X	Dneomorpha	0.796	0.611
mel2 X	parth1	0.780	0.642
mel2 X	Dparthenogenetica4	0.806	0.642
mel2 X	Dparthenogenetica5	0.800	0.642
mel2 X	Dparthenogenetica6	0.807	0.642
mel2 X	Dprocardinoides	0.773	0.624
mel2 X	Dneocardini	0.804	0.636
mel2 X	Dpolymorpha	0.779	0.625
mel2 X	Dcardinoides	0.797	0.636
Dmelanogaster5 X	yak1	0.228	0.285
Dmelanogaster5 X	Dwillistoni1	0.759	0.770
Dmelanogaster5 X	willi1	0.864	0.770
Dmelanogaster5 X	dhMiF2	0.754	0.778
Dmelanogaster5 X	Dmercatorum	0.754	0.666
Dmelanogaster5 X	Dbuzzatii	0.772	0.764
Dmelanogaster5 X	Dzottii	0.763	0.727
Dmelanogaster5 X	Dacutilabella	0.770	0.631
Dmelanogaster5 X	Dcardini1	0.790	0.657
Dmelanogaster5 X	Dcardini2	0.799	0.657
Dmelanogaster5 X	Dcardini3	0.799	0.657
Dmelanogaster5 X	Dneomorpha	0.782	0.611
Dmelanogaster5 X	parth1	0.766	0.642
Dmelanogaster5 X	Dparthenogenetica4	0.790	0.642
Dmelanogaster5 X	Dparthenogenetica5	0.785	0.642
Dmelanogaster5 X	Dparthenogenetica6	0.791	0.642
Dmelanogaster5 X	Dprocardinoides	0.757	0.624
Dmelanogaster5 X	Dneocardini	0.788	0.636
Dmelanogaster5 X	Dpolymorpha	0.758	0.625
Dmelanogaster5 X	Dcardinoides	0.780	0.636
Dmelanogaster6 X	yak1	0.222	0.285
Dmelanogaster6 X	Dwillistoni1	0.772	0.770
Dmelanogaster6 X	willi1	0.851	0.770
Dmelanogaster6 X	dhMiF2	0.769	0.778
Dmelanogaster6 X	Dmercatorum	0.768	0.666
Dmelanogaster6 X	Dbuzzatii	0.784	0.764
Dmelanogaster6 X	Dzottii	0.782	0.727

Dmelanogaster6 X	Dacutilabella	0.783	0.631	
Dmelanogaster6 X	Dcardini1	0.805	0.657	
Dmelanogaster6 X	Dcardini2	0.785	0.657	
Dmelanogaster6 X	Dcardini3	0.785	0.657	
Dmelanogaster6 X	Dneomorpha	0.794	0.611	
Dmelanogaster6 X	parth1	0.777	0.642	
Dmelanogaster6 X	Dparthenogenetica4	0.805	0.642	
Dmelanogaster6 X	Dparthenogenetica5	0.799	0.642	
Dmelanogaster6 X	Dparthenogenetica6	0.805	0.642	
Dmelanogaster6 X	Dprocardinoides	0.770	0.624	
Dmelanogaster6 X	Dneocardini	0.808	0.636	
Dmelanogaster6 X	Dpolymorpha	0.783	0.625	
Dmelanogaster6 X	Dcardinoides	0.797	0.636	
Dmelanogaster9 X	yak1	0.125	0.285	***
Dmelanogaster9 X	Dwillistoni1	0.751	0.770	
Dmelanogaster9 X	willi1	0.903	0.770	
Dmelanogaster9 X	dhMiF2	0.763	0.778	
Dmelanogaster9 X	Dmercatorum	0.760	0.666	
Dmelanogaster9 X	Dbuzzatii	0.766	0.764	
Dmelanogaster9 X	Dzottii	0.783	0.727	
Dmelanogaster9 X	Dacutilabella	0.784	0.631	
Dmelanogaster9 X	Dcardini1	0.793	0.657	
Dmelanogaster9 X	Dcardini2	0.796	0.657	
Dmelanogaster9 X	Dcardini3	0.796	0.657	
Dmelanogaster9 X	Dneomorpha	0.773	0.611	
Dmelanogaster9 X	parth1	0.769	0.642	
Dmelanogaster9 X	Dparthenogenetica4	0.790	0.642	
Dmelanogaster9 X	Dparthenogenetica5	0.791	0.642	
Dmelanogaster9 X	Dparthenogenetica6	0.796	0.642	
Dmelanogaster9 X	Dprocardinoides	0.788	0.624	
Dmelanogaster9 X	Dneocardini	0.800	0.636	
Dmelanogaster9 X	Dpolymorpha	0.745	0.625	
Dmelanogaster9 X	Dcardinoides	0.806	0.636	
yak1 X	Dwillistoni1	0.751	0.801	
yak1 X	willi1	0.853	0.801	
yak1 X	dhMiF2	0.759	0.787	
yak1 X	Dmercatorum	0.758	0.656	
yak1 X	Dbuzzatii	0.774	0.783	
yak1 X	Dzottii	0.784	0.734	
yak1 X	Dacutilabella	0.776	0.606	
yak1 X	Dcardini1	0.785	0.687	
yak1 X	Dcardini2	0.783	0.687	
yak1 X	Dcardini3	0.783	0.687	
yak1 X	Dneomorpha	0.779	0.652	
yak1 X	parth1	0.776	0.661	
yak1 X	Dparthenogenetica4	0.785	0.661	
yak1 X	Dparthenogenetica5	0.773	0.661	
yak1 X	Dparthenogenetica6	0.790	0.661	
yak1 X	Dprocardinoides	0.771	0.616	
yak1 X	Dneocardini	0.793	0.662	
yak1 X	Dpolymorpha	0.758	0.654	
yak1 X	Dcardinoides	0.799	0.637	

Dwillistoni1 X	dhMiF2	0.081	0.801	***
Dwillistoni1 X	Dmercatorum	0.066	0.772	***
Dwillistoni1 X	Dbuzzatii	0.069	0.747	***
Dwillistoni1 X	Dzottii	0.133	0.791	***
Dwillistoni1 X	Dacutilabella	0.126	0.687	***
Dwillistoni1 X	Dcardini1	0.089	0.698	***
Dwillistoni1 X	Dcardini2	0.148	0.698	***
Dwillistoni1 X	Dcardini3	0.148	0.698	***
Dwillistoni1 X	Dneomorpha	0.091	0.708	***
Dwillistoni1 X	parth1	0.091	0.728	***
Dwillistoni1 X	Dparthenogenetica4	0.095	0.728	***
Dwillistoni1 X	Dparthenogenetica5	0.099	0.728	***
Dwillistoni1 X	Dparthenogenetica6	0.095	0.728	***
Dwillistoni1 X	Dprocardinoides	0.089	0.684	***
Dwillistoni1 X	Dneocardini	0.084	0.663	***
Dwillistoni1 X	Dpolymorpha	0.214	0.691	***
Dwillistoni1 X	Dcardinoides	0.101	0.677	***
willi1 X	dhMiF2	0.601	0.801	***
willi1 X	Dmercatorum	0.614	0.772	***
willi1 X	Dbuzzatii	0.580	0.747	***
willi1 X	Dzottii	0.575	0.791	***
willi1 X	Dacutilabella	0.613	0.687	
willi1 X	Dcardini1	0.604	0.698	*
willi1 X	Dcardini2	0.641	0.698	
willi1 X	Dcardini3	0.641	0.698	
willi1 X	Dneomorpha	0.601	0.708	*
willi1 X	parth1	0.611	0.728	*
willi1 X	Dparthenogenetica4	0.604	0.728	**
willi1 X	Dparthenogenetica5	0.618	0.728	*
willi1 X	Dparthenogenetica6	0.607	0.728	**
willi1 X	Dprocardinoides	0.607	0.684	
willi1 X	Dneocardini	0.600	0.663	
willi1 X	Dpolymorpha	0.644	0.691	
willi1 X	Dcardinoides	0.613	0.677	
dhMiF2 X	Dmercatorum	0.048	0.341	***
dhMiF2 X	Dbuzzatii	0.053	0.320	***
dhMiF2 X	Dzottii	0.120	0.302	***
dhMiF2 X	Dacutilabella	0.096	0.752	***
dhMiF2 X	Dcardini1	0.063	0.726	***
dhMiF2 X	Dcardini2	0.103	0.726	***
dhMiF2 X	Dcardini3	0.103	0.726	***
dhMiF2 X	Dneomorpha	0.061	0.732	***
dhMiF2 X	parth1	0.065	0.730	***
dhMiF2 X	Dparthenogenetica4	0.069	0.730	***
dhMiF2 X	Dparthenogenetica5	0.066	0.730	***
dhMiF2 X	Dparthenogenetica6	0.069	0.730	***
dhMiF2 X	Dprocardinoides	0.063	0.743	***
dhMiF2 X	Dneocardini	0.058	0.708	***
dhMiF2 X	Dpolymorpha	0.171	0.751	***
dhMiF2 X	Dcardinoides	0.073	0.783	***
Dmercatorum X	Dbuzzatii	0.048	0.379	***
Dmercatorum X	Dzottii	0.126	0.306	***

Dmercatorum X	Dacutilabella	0.106	0.592	***
Dmercatorum X	Dcardini1	0.057	0.686	***
Dmercatorum X	Dcardini2	0.086	0.686	***
Dmercatorum X	Dcardini3	0.086	0.686	***
Dmercatorum X	Dneomorpha	0.056	0.628	***
Dmercatorum X	parth1	0.055	0.627	***
Dmercatorum X	Dparthenogenetica4	0.062	0.627	***
Dmercatorum X	Dparthenogenetica5	0.065	0.627	***
Dmercatorum X	Dparthenogenetica6	0.063	0.627	***
Dmercatorum X	Dprocardinoides	0.057	0.610	***
Dmercatorum X	Dneocardini	0.052	0.602	***
Dmercatorum X	Dpolymorpha	0.186	0.617	***
Dmercatorum X	Dcardinoides	0.066	0.657	***
Dbuzzatii X	Dzottii	0.130	0.319	***
Dbuzzatii X	Dacutilabella	0.110	0.732	***
Dbuzzatii X	Dcardini1	0.064	0.723	***
Dbuzzatii X	Dcardini2	0.095	0.723	***
Dbuzzatii X	Dcardini3	0.095	0.723	***
Dbuzzatii X	Dneomorpha	0.065	0.739	***
Dbuzzatii X	parth1	0.064	0.751	***
Dbuzzatii X	Dparthenogenetica4	0.071	0.751	***
Dbuzzatii X	Dparthenogenetica5	0.074	0.751	***
Dbuzzatii X	Dparthenogenetica6	0.069	0.751	***
Dbuzzatii X	Dprocardinoides	0.063	0.738	***
Dbuzzatii X	Dneocardini	0.059	0.738	***
Dbuzzatii X	Dpolymorpha	0.172	0.727	***
Dbuzzatii X	Dcardinoides	0.072	0.763	***
Dzottii X	Dacutilabella	0.102	0.671	***
Dzottii X	Dcardini1	0.120	0.694	***
Dzottii X	Dcardini2	0.151	0.694	***
Dzottii X	Dcardini3	0.151	0.694	***
Dzottii X	Dneomorpha	0.128	0.686	***
Dzottii X	parth1	0.131	0.715	***
Dzottii X	Dparthenogenetica4	0.129	0.715	***
Dzottii X	Dparthenogenetica5	0.129	0.715	***
Dzottii X	Dparthenogenetica6	0.135	0.715	***
Dzottii X	Dprocardinoides	0.134	0.681	***
Dzottii X	Dneocardini	0.114	0.678	***
Dzottii X	Dpolymorpha	0.321	0.690	***
Dzottii X	Dcardinoides	0.127	0.732	***
Dacutilabella X	Dcardini1	0.114	0.264	***
Dacutilabella X	Dcardini2	0.149	0.264	
Dacutilabella X	Dcardini3	0.149	0.264	
Dacutilabella X	Dneomorpha	0.118	0.238	**
Dacutilabella X	parth1	0.123	0.248	**
Dacutilabella X	Dparthenogenetica4	0.121	0.248	**
Dacutilabella X	Dparthenogenetica5	0.120	0.248	**
Dacutilabella X	Dparthenogenetica6	0.121	0.248	**
Dacutilabella X	Dprocardinoides	0.103	0.219	**
Dacutilabella X	Dneocardini	0.109	0.220	**
Dacutilabella X	Dpolymorpha	0.242	0.176	
Dacutilabella X	Dcardinoides	0.119	0.224	**

Dcardini1 X	Dneomorpha	0.036	0.325	***
Dcardini1 X	parth1	0.053	0.298	***
Dcardini1 X	Dparthenogenetica4	0.006	0.298	***
Dcardini1 X	Dparthenogenetica5	0.009	0.298	***
Dcardini1 X	Dparthenogenetica6	0.030	0.298	***
Dcardini1 X	Dprocardinoides	0.060	0.294	***
Dcardini1 X	Dneocardini	0.018	0.363	***
Dcardini1 X	Dpolymorpha	0.140	0.278	**
Dcardini1 X	Dcardinoides	0.006	0.271	***
Dcardini2 X	Dneomorpha	0.045	0.325	***
Dcardini2 X	parth1	0.105	0.298	***
Dcardini2 X	Dparthenogenetica4	0.045	0.298	***
Dcardini2 X	Dparthenogenetica5	0.060	0.298	***
Dcardini2 X	Dparthenogenetica6	0.000	0.298	***
Dcardini2 X	Dprocardinoides	0.089	0.294	***
Dcardini2 X	Dneocardini	0.015	0.363	***
Dcardini2 X	Dpolymorpha	0.031	0.278	***
Dcardini2 X	Dcardinoides	0.052	0.271	***
Dcardini3 X	Dneomorpha	0.045	0.325	***
Dcardini3 X	parth1	0.105	0.298	***
Dcardini3 X	Dparthenogenetica4	0.045	0.298	***
Dcardini3 X	Dparthenogenetica5	0.060	0.298	***
Dcardini3 X	Dparthenogenetica6	0.000	0.298	***
Dcardini3 X	Dprocardinoides	0.089	0.294	***
Dcardini3 X	Dneocardini	0.015	0.363	***
Dcardini3 X	Dpolymorpha	0.031	0.278	***
Dcardini3 X	Dcardinoides	0.052	0.271	***
Dneomorpha X	parth1	0.062	0.310	***
Dneomorpha X	Dparthenogenetica4	0.042	0.310	***
Dneomorpha X	Dparthenogenetica5	0.046	0.310	***
Dneomorpha X	Dparthenogenetica6	0.042	0.310	***
Dneomorpha X	Dprocardinoides	0.061	0.275	***
Dneomorpha X	Dneocardini	0.029	0.322	***
Dneomorpha X	Dpolymorpha	0.144	0.195	
Dneomorpha X	Dcardinoides	0.039	0.303	***
parth1 X	Dprocardinoides	0.059	0.199	***
parth1 X	Dneocardini	0.056	0.372	***
parth1 X	Dpolymorpha	0.184	0.261	
parth1 X	Dcardinoides	0.063	0.159	**
Dparthenogenetica4 X	Dprocardinoides	0.066	0.199	***
Dparthenogenetica4 X	Dneocardini	0.024	0.372	***
Dparthenogenetica4 X	Dpolymorpha	0.147	0.261	**
Dparthenogenetica4 X	Dcardinoides	0.013	0.159	***
Dparthenogenetica5 X	Dparthenogenetica6	0.030	0.000	
Dparthenogenetica5 X	Dprocardinoides	0.069	0.199	***
Dparthenogenetica5 X	Dneocardini	0.027	0.372	***
Dparthenogenetica5 X	Dpolymorpha	0.153	0.261	*
Dparthenogenetica5 X	Dcardinoides	0.010	0.159	***
Dparthenogenetica6 X	Dprocardinoides	0.066	0.199	***
Dparthenogenetica6 X	Dneocardini	0.024	0.372	***
Dparthenogenetica6 X	Dpolymorpha	0.148	0.261	**
Dparthenogenetica6 X	Dcardinoides	0.038	0.159	***

Dprocardinoides X	Dneocardini	0.054	0.353	***
Dprocardinoides X	Dpolymorpha	0.170	0.248	
Dprocardinoides X	Dcardinoides	0.070	0.173	**
Dneocardini X	Dpolymorpha	0.124	0.247	**
Dneocardini X	Dcardinoides	0.025	0.345	***
Dpolymorpha X	Dcardinoides	0.155	0.230	

sech1 groups the sequences: Dsechellia2, Dsechellia3, Dsechellia4, Dsechellia6, Dsechellia10. Dsechellia11, Dsechellia14 and Dsechellia19; **sech2** groups sequences: Dsechellia12, Dsechellia16, Dsechellia17 and Dsechellia23; **sech3** groups: Dsechellia20. Dsechellia21, Dsechellia24 and Dsechellia25; **mel1**: Dmelanogaster2 and Dmelanogaster7; **mel2**: Dmelanogaster3 and Dmelanogaster4; **sim1**: Dsimulans4 and Dsimulans10; **yak1** groups all sequences from *D. yakuba* (Dyakuba1, Dyakuba2, Dyakuba3, Dyakuba4, Dyakuba5, Dyakuba6, Dyakuba7, Dyakuba8, Dyakuba9, Dyakuba10. Dyakuba11 and Dyakuba12); **Dbuzzatii** groups all sequences from *D. buzzatii* (Dbuzzatii1, Dbuzzatii2, Dbuzzatii3, Dbuzzatii4 and Dbuzzatii5); **Dmercatorum** groups all sequences from Dmercatorum (Dmercatorum1, Dmercatorum2, Dmercatorum3, Dmercatorum4, Dmercatorum5 and Dmercatorum6); **Dneomorpha** groups all sequences from *D. neomorpha* (Dneomorpha1, Dneomorpha2 and Dneomorpha3); **Parth1** groups sequences from *D. parthenogenetica* (Dparthenogenetica1, Dparthenogenetica2 and Dparthenogenetica3) with exception from sequences Dparthenogenetica4 and Dparthenogenetica5; **Dprocardinoides** groups all sequences from *D. procardinoides* (Dprocardinoides1, Dprocardinoides2 and Dprocardinoides3); **Dacutilabella** groups all sequences from *D. acutilabella* (Dacutilabella1, Dacutilabella2, Dacutilabella3, Dacutilabella4 and Dacutilabella5); and **Dzottii** groups the two sequences from *D. zottii* (Dzottii1 and Dzottii2).

CAPÍTULO IV

Trabalho em preparação para ser submetido à revista científica Hereditas

Brief report: Chromosomal localization of the retrotransposable element *micropia* in the *cardini* group of the *Drosophila* genus (Diptera: Drosophilidae)

Juliana Cordeiro[1]; Gisele de Souza da Silva[2]; Vera Lucia Silva Valente[1, 3]

[1] Programa de Pós Graduação em Genética e Biologia Molecular; Universidade Federal do Rio Grande do Sul; Porto Alegre; Rio Grande do Sul; Brazil

[2] Curso de Ciências Biológicas; Universidade LaSalle; Canoas; Rio Grande do Sul; Brazil

[3] Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul; Porto Alegre; Rio Grande do Sul; Brazil

Corresponding author: Juliana Cordeiro, Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, CEP 91501-970, Porto Alegre, Rio Grande do Sul, Brazil. Phone: +55-51-33166729, Fax: +55-51-33167311. E-mail: jlncdr@gmail.com

Abstract

This work analyzes the possible *micropia* related sequences insertion sites in the polytene chromosomes of six species of the *cardini* group from *Drosophila* genus: *D. cardini*, *D. cardinoides*, *D. neocardini*, *D. neomorpha*, *D. parthenogenetica* and *D. polymorpha*. This analysis showed an interesting distribution pattern, where the *micropia* element is present from five to 18 copies number among the species, despite its recent invasion in their genomes. Considering all species together, the IV chromosome, in *D. cardini*, or the IIIL chromosome, in the other species, presented the majority of the insertion sites (19). It was also observed that *micropia* is present near or in the break point of some inversions in *D. cardinoides*, *D. polymorpha* and *D. parthenogenetica*. The significance of these results is discussed.

Running title: *micropia* chromosomal localization in *cardini* group

Keywords: *micropia* – *cardini* group – *in situ* hybridization – chromosomal localization

Introduction

Due to the transposable elements (TEs) dynamics in the host genome they have been indicated as responsible for causing gene mutation and chromosomal rearrangements, leading to genome evolution (Capy et al. 1998; Miller and Capy 2006). Studies on *Drosophila* species were the first to suggest (Lyttle and Haymer 1992; Regner et al. 1996; Ledevèze et al. 1998) and show evidences (Cáceres et al. 1999, 2001; Puig et al. 2004) of the TEs involvement in chromosomal rearrangement, indicating their presence in the chromosomal inversion breakpoints. These last works suggest that TEs can promote these variations due to ectopic recombination (Marzo et al. 2008).

More than a half of the *Drosophila* species are polymorphic for paracentric inversions (Krimbas and Powell 1992) and some species of the *cardini* group follow this pattern (Heed and Russel 1971). This group consists of 16 species with Neotropical distribution which ranges from southern USA to southern Brazil and northern Argentina and Chile, including the Caribbean islands on Central America (Heed and Rusell 1971; Vilela et al. 2002). The group is divided in two subgroups, the *dunni* subgroup and the *cardini* subgroup (Heed 1962) which together compose one of the intriguing groups of the *Drosophila* genus in terms of distribution, abdominal pigmentation and evolution pattern (Heed and Russel 1971; Hollocher et al. 2000; Brisson et al. 2006). The phylogenetic relationships within the *cardini* group are not totally elucidated (Brisson et al. 2006) however it is well accepted that *D. cardini* is the basal species of the group with a diploid number of chromosomes 2n=12 with all acrocentric chromosomes, while the other species have 2n=8 with acrocentric XY chromosome pair and metacentric autosomes (Heed 1962). In this work we analyzed six species of the *cardini* subgroup (*D. cardini*, *D. cardinoides*,

D. neocardini, *D. neomorpha*, *D. parthenogenetica* and *D. polymorpha*) for the location of the *micropia* element.

Micropia was discovered in the lampbrush loops of *D. hydei* (Huijser et al. 1988), and is classified as a retrotransposable element with long terminal repeats (LTRs) belonging to the *Ty*3 superfamily (Capy et al. 1998). However it seems that at least some copy of a *micropia* related sequence might be putatively active in the species of the *Drosophila* subgenus (*cardini* and *repleta* group) and *Sophophora* subgenus (*willistoni* and *melanogaster* group) (Cordeiro et al. unpublished data, chapter III of this Thesis). Based on these findings the present study offers a contribution to the growing body of evidence that has shown that TEs may act as a rearrangement promoter and/or only as a secondary invader of these regions (Cassals et al. 2006; Marzo et al. 2008).

Materials and Method

Samples

One strain of each of the following species was studied: *D. cardini* (strain from Serra do Japi, Sao Paulo State, Brazil - 23º14'S 46º58'W), *D. cardinoides*, *D. neocardini*, *D. polymorpha* (strains from Florianópolis, Santa Catarina State, Brazil - 27º35'S 48º22'W), *D. neomorpha* (strain from Joinville, Santa Catarina State, Brazil - 26º17'S 49º00'W) and *D. parthenogenetica* (strain from Puebla, Mexico, Tuckson Stock Center number 15181-2221.01). These strains have been maintained in the laboratory, reared in corn flour culture medium in a controlled chamber (17°C ± 1°C, 60% r.h.) and carefully fed with yeast solution once a day. To ensure the identification of the species, we reanalyzed the stocks using the literature available about external morphologies and genitalia characters (Val 1982; Vilela and Bächli 1990; De Toni et al. 2005).

In situ hybridization on polytene chromosomes

To analyze the insertion sites of the *micropia* retroelement, species-specific clones were used as probes for *in situ* hybridization on polytene chromosomes. Polymerase chain reaction (PCR) assays were performed with the primers Mic1777 (5' CTCCCCTTTTGCCAGTCCT 3') and Mic2570 (5' TTGAGCTAGCGTCGGTGTG 3') (Cordeiro et al. 2008). The following conditions for a 25μl PCR reaction were used: 25ng/μl template DNA, 20pMol of each primer, 0.2mM of each nucleotide, 1.5mM $MgCl_2$ and 1 unit Taq DNA polymerase in 1x Polymerase Buffer (all from Invitrogen). Amplification parameters for *D. cardini*, *D. neomorpha* and *D. parthenogenetica* were 95°C for 2min, 35 cycles at 95°C for 30s, 57°C for 30s and 72°C for 1min, followed by an extension cycle at 72°C for 10min. The annealing temperature used to amplify *micropia* in *D. cardinoides*, *D. neocardini* and *D. polymorpha* was 64°C. The amplicons size was approximately 850bp; exception for *D. polymorpha* with 700bp. The fragments were inserted into pGEM®-T Easy Vector (Promega), for *D. parthenogenetica*, and into TOPO-TA Cloning® Vector (Invitrogen), for the others species. Polytene chromosome squashing, hybridization, and detection were carried out as described by Montgomery et al. (1987), with small modifications. All probes were labelled with biotin-14-dATP (Invitrogen) according to the nick translation technique; the detection was carried out using the ABC®-Elite kit (Vector Laboratories) and the hybridizations were performed at 37°C. Hybridization signals were located on the polytene chromosomes using the improved cytological maps of each species (Cordeiro et al. unpublished data, chapter V of this Thesis) and their identification was done comparing them among the nucleus analyzed for each species, where three larvae per species were used and at least three nuclei per larvae were photomicrographed.

Results and Discussion

Considering that *D. cardini* has 2n=12 and the other species of the group *cardini* have 2n=8, here we designate the chromosomal complements as, for example, II/IIL chromosome; where the first is the *D. cardini* chromosome and then the others species chromosomal arm. The Figs. 1, 2 and 3 show the hybridization signals of the *micropia* retroelement detected in the chromosomes of the *cardini* group species. Fig. 4 and Table 1 summarize the putative insertion sites of *micropia* retroelement detected by *in situ* hybridization in each chromosomal arm of the species.

Analyzing the Figs. and observing the Table 1, it is possible to identify five hybridizations signals in *D. neomorpha*, six in *D. cardini* and in *D. parthenogenetica*, 12 in *D. neocardini*, 14 in *D. cardinoides*, and 18 in *D. polymorpha*. Considering the total insertions per chromosome in Table 1, the IV/IIIL chromosomes have the majority of the *micropia* insertions (19), followed by the III/IIR chromosomes (12), by II/IIL chromosomes (11), by X chromosome (10) and by V/IIIR chromosomes (nine). Indeed the IIIL chromosomal arm is the most polymorphic for paracentric inversion for *D. cardinoides* and *D. neocardini* (Rohde and Valente 1996; De Toni et al. 2001). *Drosophila polymorpha* is the most polymorphic species of the group with inversions described for all its chromosomal arms (Cordeiro et al. unpublished data, chapter V of this thesis), and for this species we localized *micropia* hybridization signals in all chromosomal arms, totalizing 18 insertion sites (Fig 3; Table 1).

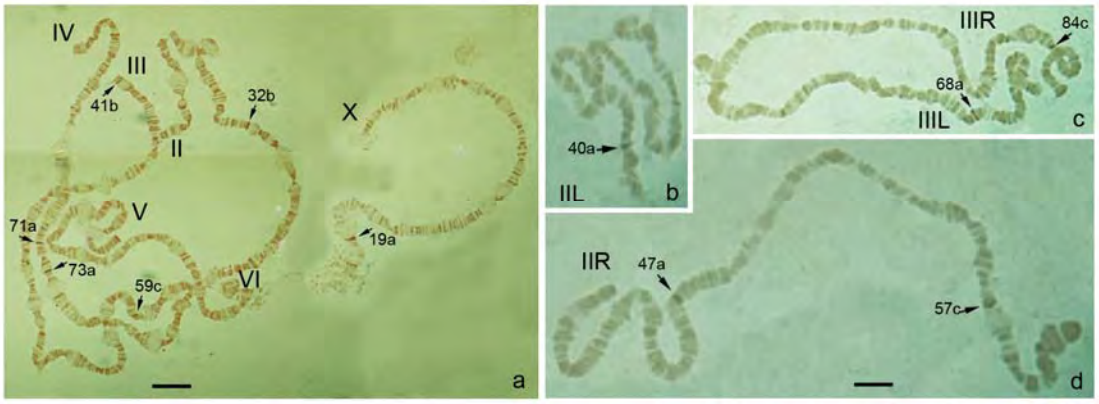


Fig. 1. *Drosophila cardini* (a) and *D. neomorpha* (b, c, d) polytene chromosomes hybridized with the *micropia* probe. The arrows indicate the hybridization signals; beside it, there is the chromosomal section of the reference photomaps. The identification of the chromosome and the chromosomal arms is indicated near the tip of them. Bar: 10μm

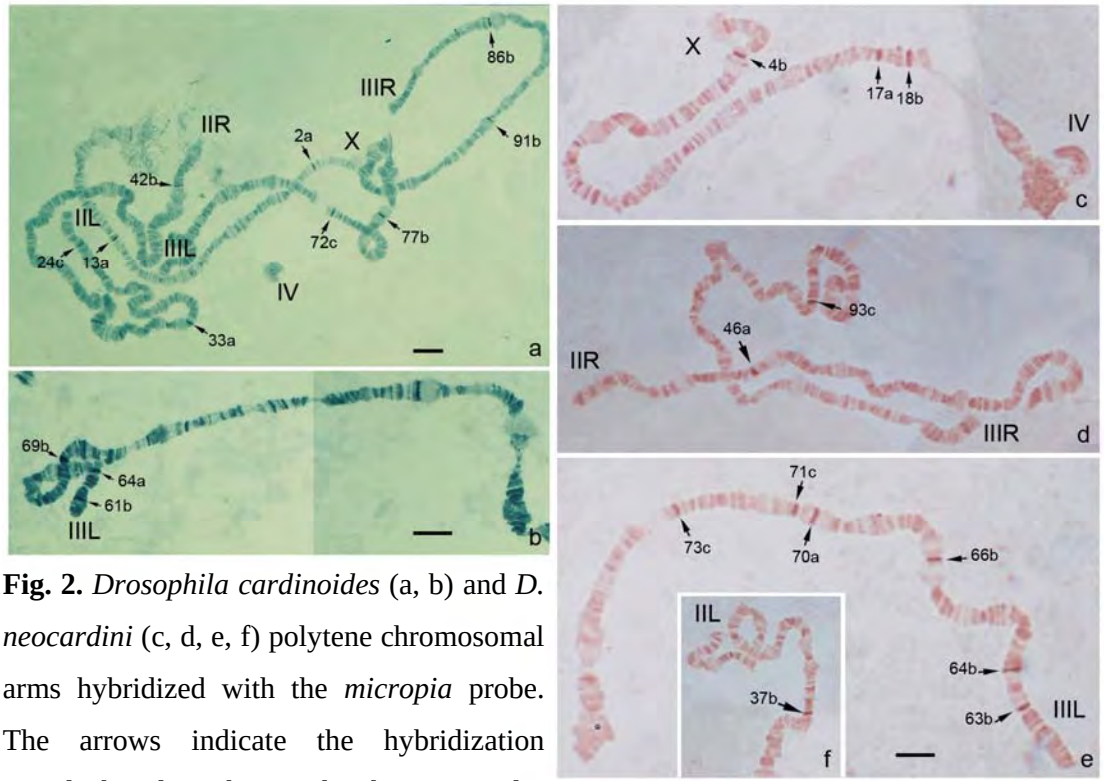


Fig. 2. *Drosophila cardinoides* (a, b) and *D. neocardini* (c, d, e, f) polytene chromosomal arms hybridized with the *micropia* probe. The arrows indicate the hybridization signals; beside it, there is the chromosomal sections of the reference photomaps. The identification of the chromosomal arms is indicated near the tip of them. Bar: 10μm

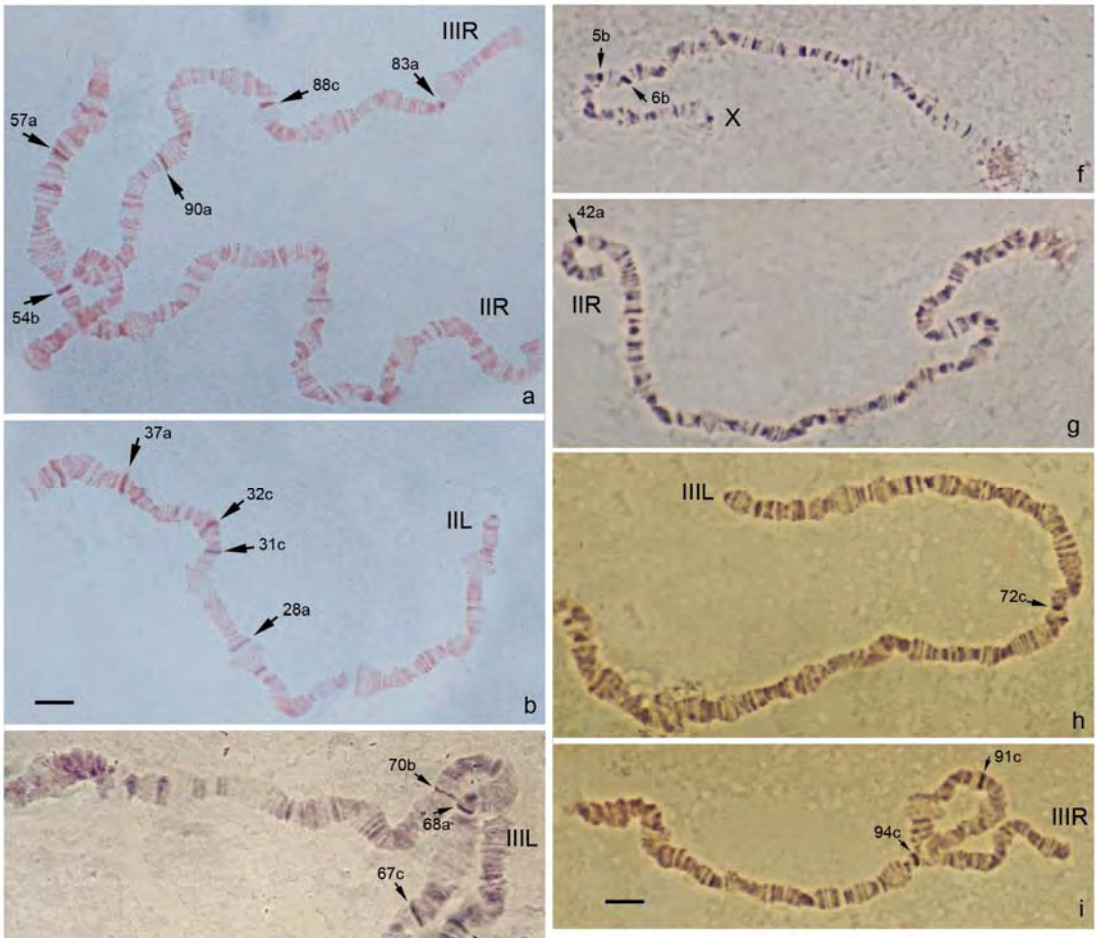


Fig. 3. *Drosophila polymorpha* (a, b, c, d, e) and *D. parthenogenetica* (f, g, h, i) polytene chromosomes hybridized with the *micropia* probe. The arrows indicate the hybridization signals; beside it, there is the chromosomal section of the reference photomaps. The identification of the chromosomal arms is indicated near the tip of them. Bar: 10µm

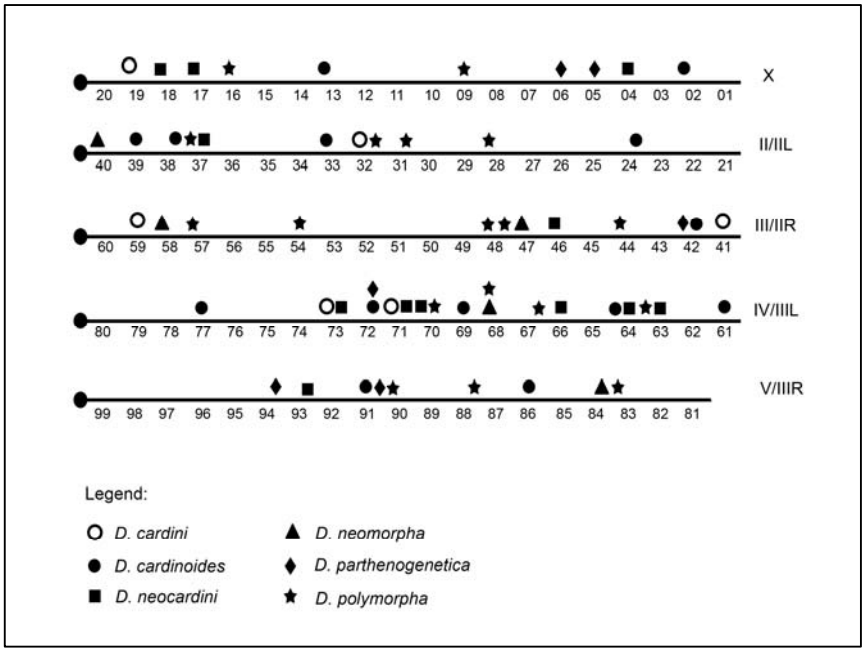


Fig. 4. Scheme showing a general pattern of the *micropia* insertion sites on the polytene chromosomes of *D. cardini*, *D. cardinoides*, *D. neocardini*, *D. neomorpha*, *D. parthenogenetica* and *D. polymorpha* (see legend in the Fig.). The chromosomes and chromosomal arms are identified on the right. The chromocenter is located on the left represented by a filled circle; and the sections of the polytene chromosomes are indicated under the lines that represent the chromosomes.

Table 1: Putative insertion sites of *micropia* retroelement detected by *in situ* hybridization on the chromosomes of the six *cardini* species group analyzed in this study.

Chromosome/ Chromosomal arms	***D. cardini***	***D. cardinoides***	***D. neocardini***	***D. neomorpha***	***D. parthenogenetica***	***D. polymorpha***	**Total per chromosome**
X	19a	2a, 13a	4b, 17a, 18a	--	5b, 6b	9b, 16a	10
II/IIL	32b	24c, 33a, 38c, 39b	37b	40a	--	28a, 31c, 32c, 37a	11
III/IIR	41b, 59c	42b	46a	47a, 57c	42a	44a**, 48a**,48c*, 54b, 57a	12
IV/IIIL	71a, 73a	61b, 64a*, 69b, 72c*, 77b	63b, 64b, 66b, 70a, 71c, 73c	68a	72c	63a, 67c, 68a, 70b	19
V/IIIR	--	86b, 91b	93c	84c	91c*, 94c	83a*, 88c, 90a	09
Total per species	06	14	12	05	06	18	

* insertion site next to chromosomal inversion break point.
** insertion site in the chromosomal inversion break point.

However, the V chromosome of *D. cardini*, the X chromosome of *D. neomorpha* and the IIL chromosome of *D. parthenogenetica* did not present any hybridization signal. Moreover, these three species showed the lower number of hybridization signals here detected and, coincidently, they are the species with less heterozygous chromosomal inversion polymorphism in the *cardini* subgroup (De Toni et al. 2006; Cordeiro et al. unpublished data, chapter V of this Thesis). These three species have limited ecological characteristics, as verified by the very low frequency on field collections (De Toni et al. 2005; 2007) which, as a consequence, may reflect in the low polymorphism for heterozygous inversions in these species. It is interesting to observe that the species with a wide distribution and high frequency on field collections; *D. neocardini*, *D. cardinoides* and *D. polymorpha*; obtained the higher number of *micropia* insertions. Although it was suggested for *D. melanogaster* that the TEs acquisition and accumulation in the genome may parallel the species colonization process; for *D. simulans* it was not observed (Vieira et al. 1999). *Drosophila melanogaster* has a worldwide distribution that is thought to have occurred a long time ago; otherwise, the worldwide dispersal of its sibling species, *D. simulans*, seems to have occurred more recently (Nardon et al. 2005). Therefore, *D. simulans* populations still seem to be in the process of being invaded by TEs (Vieira et al. 1999). When populations are subjected to new environmental conditions may have some of their TEs mobilized (Capy et al. 2000) increasing their copy number in the genome. Therefore, this can explain the insertion sites number differences among the species of the *cardini* group; where *D. neocardini*, *D. polymorpha* and *D. neocardini* have a higher insertion site number then *D. cardini*, *D. neomorpha* and *D. parthenogenetica*.

The dot chromosomes (data not shown) and the centromeric heterochromatin (observed on the region base of the X chromosomes in Figs 1, 2 and 3) of all species do not

react with the probes used for *in situ* hybridizations, indicating that the *cardini* group species do not have this *micropia* sequences insertions in these structures. The presence of a TE in the heterochromatin is a strong inference that this copy might be not active (Capy et al. 1998); therefore, our result suggests an evidence of the activity of these *micropia* sequences in the genome of the *cardini* group species. Contrasting with these results, in *D. hydei micropia* is highly present in the heterochromatin of this species and it seems to be active (Hennig et al. 1983).

The representativeness of *micropia* related sequences copy number for the *repleta* group species was estimated between seven and 17 copies (Almeida and Carareto 2004), for *D. melanogaster* between 16 and 32 copies (Lankenau 1990), and for *D. cardinoides D. neocardini* and *D. polymorpha* between four and six copies (Cordeiro et al. 2008); all of these data was based on Southern Blot analysis. For *D. hydei*, Hennig et al. (1983) estimated 11 euchromatic copies of *micropia* using *in situ* hybridization on polytene chromosomes. Regarding the data for the *cardini* group, here we improved our previous results using the accurate *in situ* hybridization technique for this purpose.

Despite being indicated as rearrangements promoters (Cáceres et al. 1999, 2001; Puig et al. 2004) the relationship between chromosomal inversions and the number of TEs in the genome is not explained simply by means of inversion generation. Such association may also be explained by the generation of hotspots for TEs insertions after the occurrence of the inversion, where these elements may act as secondary invaders of these regions (Casals et al. 2006; Ranz et al. 2007).

Until now there is no evidence that retrotransposable elements with LTR are the promoters of the chromosomal inversions as observed for the transposon *Galileo* (Cáceres et al. 1999); however, the elements with LTR are the most representative TEs in the

genome of *D. melanogaster*, *D. simulans* and *D. yakuba* (Bartolomé et al. 2009). Our results found some hybridization signals near or in the inversion break points for *D. cardinoides* (Fig. 2b), *D. polymorpha* (Figs. 3a and 3d) and for *D. parthenogenetica* (Fig 3i; see Table 1), as it follows: for *D. cardinoides*, sections 64a and 72c, near the distal break point of inversion IIILD and IIILC, respectively; for *D. polymorpha*, sections 44a and 48a, in the distal break point of inversion IIRD, IIRA, respectively, and sections 48c and 83a, near the distal break point of inversion IIRD and IIIRA, respectively; and for *D. parthenogenetica*, section 94c, near the proximal break point of inversion IIIRA. Considering these results, the precise inference of the action of this element in the generation of inversions is not possible to be suggested; so, in these cases, *micropia* might be acting as a secondary invader of these chromosomal sites.

The TEs seem to be maintained in populations as equilibrium between the transposing events, which can amplify their copy number (Charlesworth and Charlesworth 1983), and the action of opposing forces that reduce their copies (Charlesworth and Charlesworth 1983; Montgomery et al. 1987; Langley et al. 1988). Regarding mechanism that leads to copies reduction, it is inferred that selection can act either directly, against deleterious insertions of TEs, or indirectly, against deleterious chromosomal rearrangements produced by non-homologous ectopic recombination between the TEs (Biémont et al. 1997; Guerreiro et al. 2008). In the first model the action of selection against deleterious insertions would lead to a lower number of TEs copies in the X chromosome, when compared to autosomes, due to the possible deleterious effect of recessive insertional mutations in hemizygous males of *Drosophila* species (Montgomery et al. 1987; Hoogland and Biémont 1996). This hypothesis was tested for the *cardini* group

species here studied and the number of insertion sites in X chromosomes is not statistically different from that of autosomes ($\chi^2 = 0.593$, Gl = 0.05).

In the second model, where the selection acts against ectopic rearrangements of TEs, an accumulation of TEs in reduced crossing over regions is expected, such as the base and the tip of the chromosomes (Hoogland and Biémont 1996). However, in the *cardini* group species the *micropia* element was detected in all chromosome extensions with no statistical significance difference, and also, this retroelement is equally distributed on the X chromosomes and autosomes (Fig. 4). Otherwise, for *D. melanogaster* no global tendency was observed in the relationship between the TEs chromosomal distribution and reduced crossing over region for the elements *P*, *I*, *copia*, *mdg1*, *mdg3*, *412*, *297* and *roo*, but for *hobo* element it was observed a preferential insertion in the tip and base of the chromosomes (Hoogland and Biémont 1996). Moreover, Vieira and Biémont (1996) observed a strong tendency toward a lower copy number of the *copia* and *412* elements on the X chromosome of *D. melanogaster* populations, demonstrating that the TEs dynamics may be controlled by different factors, depending on the interaction between the host genome and the TE itself.

The *cardini* group consists of 16 described species inhabiting different areas of Neotropical America subdivided in two subgroups: the *dunni* subgroup, with eight species restricted to the Caribbean islands, and the *cardini* subgroup composed of the other eight species that are present in the mainland, comprising an area that starts in USA and stretches to south Brazil, covering also the north of Argentina and Chile (Heed and Rusell 1971; Vilela et al. 2002). We have shown that all species from the *cardini* subgroup have sequences related to *micropia* in their genomes (Cordeiro et al. unpublished data, chapter III of this Thesis). However, the species from the *dunni* subgroup seems to have *micropia*

copies with a very divergent nucleotide sequence. To explain these findings, we hypothesized that those species might have acquired the *micropia* copies recently (3.7MYA) by an horizontal transfer event after the species divergence, that might have been occurred about (10.7MYA) (Cordeiro et al. unpublished data, chapter III of this Thesis). Moreover, based on the analysis of approximately 700bp of reverse transcriptase sequence, it seems that *micropia* might be active in the *cardini* subgroup species with exception of *D. polymorpha*.

Some studies have shown that this TE expression gives rise to a complex set of sense and antisense RNAs (Lankenau et al. 1994; Almeida and Carareto 2004). In this model, the antisense transcript, produced by the copies present in the Y chromosome, is potentially driven against the entire reverse transcriptase domain, thus contributing to the *micropia* RNA degradation (Lankenau 1993). In the *cardini* group the *micropia* control seems to have a different mechanism, since it is not possible do detect any hybridization signal that could be related to the Y chromosome, i.e., in the heterochromatin.

Acknowledgements – The authors thank BSc. Carolina Flores Garcia and Dr Marindia Depra for valuable help. This work was supported by grants from CAPES, CNPq, and FAPERGS research funding.

References

Almeida, L. M. and Carareto, C. M. A. 2004. Identification of two subfamilies of *micropia* transposable element in species of the *repleta* group of *Drosophila*. - Genetica 121:155-164.

Bartolomé, C., Bello, X. and Maside, X. 2009. Widespread evidence for horizontal transfer of transposable elements across *Drosophila* genomes. - Genome Biol 10(2):R22.

Biémont, C., Vieira, C., Hoogland, C. et al. 1997. Maintenance of transposable element copy number in natural populations of *Drosophila melanogaster* and *Drosophila simulans*. - Genetica 100:161-166.

Brisson, J. A., Wilder, J. and Hollocher, H. 2006. Phylogenetic analysis of the *cardini* group of *Drosophila* with respect to changes in pigmentation. - Evolution 60:1228-1241.

Cáceres, M., Puig, M. and Ruiz, A. 2001. Molecular characterization of two natural hotspots in the *Drosophila buzzatii* genome induced by transposon insertions. – Genome Res 11:1353-1364.

Cáceres, M., Ranz, J. M., Barbadilla, A. et al. 1999. Generation of a widespread *Drosophila* inversion by a transposable element. - Science 285:415-418.

Capy, P., Bazin, C., Higuet, D. et al. 1998. Dynamics and evolution of transposable elements. Texas: Landes Bioscience 197pp.

Casals, F., González, J. and Ruiz, A. 2006. Abundance and chromosomal distribution of six *Drosophila buzzatii* transposons: *BuT1*, *BuT2*, *BuT3*, *BuT4*, *BuT5*, and *BuT6*. - Chromosoma 115:403-12.

Charlesworth, B. and Charlesworth, D. 1983. The population dynamics of transposable elements. – Genet. Res. 42: 1-27.

Cordeiro, J., Robe, L. J., Loreto, E. L. et al. 2008. The LTR retrotransposon *micropia* in the *cardini* group of *Drosophila* (Diptera: Drosophilidae): a possible case of horizontal transfer. - Genetica 134:335-344.

De Toni, D. C., Brisson, J. A., Hofmann, P. R. P. et al. 2005. Frist record of *Drosophila parthenogenetica* and *D. neomorpha*, *cardini* group, Heed, 1962 (*Drosophila*, Drosophilidae), in Brazil. - Drosoph. Inf. Serv. 88:33-38.

De Toni, D. C., Gottschalk, M. S., Cordeiro, J. et al. 2007. Study of the Drosophilidae (Diptera) communities on Atlantic Forest Islands of Santa Catarina State, Brazil. - Neotropical Ent. 36:356-375.

De Toni, D. C., Herédia, F. O. and Valente, V. L. S. 2001. Chromosomal variability of *Drosophila polymorpha* populations from Atlantic Forest remnants of continental and insular environments in the State of Santa Catarina, Brazil. - Caryologia 4 (54):329-337.

De Toni, D. C., Loureiro, M. A., Hofmann, P. R. P. et al. 2006. Reference photomap of the salivary gland polytene chromosomes of *Drosophila neomorpha* (Streisinger, 1946) - Drosoph. Inf. Serv. 89: 73-77.

Guerreiro, M. P. G., Chávez-Sandoval, B. E., Balanyà, J., et al. 2008. Distribution of the transposable elements *bilbo* and *gypsy* in original and colonizing populations of *Drosophila subobscura.* - BMC Evol. Biol. 8:234.

Heed, W. B. 1962. Genetic characteristics of island populations. - Univ. Texas. Publ. Stud. Genet. 6205:173–206.

Heed, W. B. and Russell, J. S. 1971. Phylogeny and population structure in island and continental species of the *cardini* group of *Drosophila* studied by inversion analysis. – Univ. Texas. Publ. Stud. Genet. 7103:91-130.

Hennig, W., Huijser, P., Vogt, P., Jackle, H. and Edstrom, J-E. 1983. Molecular cloning of microdissected lampbrush loop DNA sequences of *Drosophila hydei*. - EMBO J 2 (10): 1746-1983.

Hollocher, H., Hatcher, J. L. and Dyreson, E. G. 2000. Evolution of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. - Evolution 54:2046-2056.

Hoogland, C. and Biémont, C. 1996. Chromosomal distribution of transposable elements in *Drosophila melanogaster*: test of the ectopic recombination model for maintenance of insertion site number. – Genetics 144:197-204.

Huijser, P., Kirchhoff, C., Lankenau, D. H. et al. 1988. Retrotransposon-like sequences are expressed in Y chromosomal lampbrush loops of *Drosophila hydei*. – J. Mol. Biol. 203:689-697.

Krimbas, C. D. and Powell, J. R. 1992. *Drosophila* inversion polymorphism. Boca Raton, Florida: CRC Press 560pp.

Ladevèze, V., Aulard, S., Chaminade, N. et al. 1998. *Hobo* transposons causing chromosomal breakpoints. - Proc. Biol. Sci. 265:1157-1159.

Langley, C. H., Montgomery, E. A., Hudson, R., Kaplan, N. and Charlesworth, B. 1988. On the role of unequal exchange in the containment of transposable element copy number. – Genet. Res. 52:223-235.

Lankenau, *D.* H. 1990. Molecular structure and evolution of a retrotranspon family in *Drosophila*. Ph.D. Thesis, university of Nijmegen, The Netherlands.

Lankenau, D. H. 1993. The retrotransposon family *micropia* in *Drosophila* species. - In: McDonald, J. (ed) Transposable Elements and Evolution. Amsterdam: Kluwer Publishers. pp 232-241.

Lankenau, S., Corces, G. V. and Lankenau, D. H. 1994. The *Drosophila micropia* retrotransposon encodes a testis-specific antisense RNA complementary to reverse transcriptase. – Mol. Biol. Evol. 17:1542-1557.

Lyttle, T. W. and Haymer, D. S. 1992. The role of the transposable element *hobo* in the origin of the endemic inversions in wild populations of *Drosophila melanogaster*. - Genetical Res. 86:113-126.

Marzo, M., Puig, M. and Ruiz, A. 2008. The Foldback-like element *Galileo* belongs to the *P* superfamily of DNA transposons and is widespread within the *Drosophila* genus. – Proc. Natl. Acad. Sci. USA. 105:2957-2962.

Miller, W. J. and Capy, P. 2006. Applying mobile genetic elements for genome analysis and evolution. – Mol. Biotechnol. 33:161-174.

Montgomery, E. A., Charlesworth, B. and Langley, C. H. 1987. A test for the role of a natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. – Genet. Res. 49:31-41.

Puig, M., Cáceres, M. and Ruiz, A. 2004. Silencing of a gene adjacent to the breakpoint of a widespread *Drosophila* inversion by a transposon-induced antisense RNA. – Proc. Natl. Acad. Sci. USA. 101: 9013-9018.

Ranz, J. M., Maurin, D., Chan, Y. S., et al. 2007. Principles of genome evolution in the *Drosophila melanogaster* species group. - PLoS Biol 6:1366-1381.

Regner, L. P., Pereira, M. S. O., Alonso, C. E. V., et al. 1996. Genomic distribution of *P* elements in *Drosophila willistoni* and a search for their relationship with chromosomal inversions. - J. Hered. 87(3):191-198.

Rohde, C. and Valente, V. L. S. 1996. Cytological maps and chromosomal polymorphism of *Drosophila polymorpha* and *Drosophila cardinoides*. – Braz. J. Genet. 19:27-32.

Val, F. C. 1982. The male genitalia of some Neotropical *Drosophila*: Notes and illustrations. – Pap. Avulsos Zool. 34: 309-347.

Vieira, C. and Biémont, C. 1996. Selection against transposable elements in *D. simulans* and *D. melanogaster*. – Genet. Res. 68(1):9-15.

Vieira, C., Lepetit, D., Dumont, S. and Biémont C. 1999. Wake up of transposable elements following *Drosophila simulans* worldwide colonization. – Mol. Biol. Evol. 16(9):1251-1255.

Vilela, C. R. and Bächli, G. 1990. Taxonomic studies on Neotropical species of seven genera of Drosophilidae (Diptera). – Mitt. Schweiz Ent. Ges. 63: 1-332.

Vilela, C. R., da Silva, A. F. G. and Sene, F. M. 2002. Preliminary data on the geographical distribution of *Drosophila* species within morphoclimatic domains of Brazil. – Revta. Bras. Ent. 2:139-148.

CAPÍTULO V

Trabalho em preparação a ser submetido à Cytogenetic and Genome Research

Chromosomal evolution in the *cardini* group of *Drosophila* (Diptera: Drosophilidae): Evolutionary analysis of six species of the *cardini* subgroup, their inversions and new photomaps

Juliana Cordeiro[1*]; Daniela Cristina De Toni[2*]; Gisele de Souza da Silva[3]; Vera Lucia da Silva Valente[1, 4]

[1] Programa de Pós Graduação em Genética e Biologia Molecular; Universidade Federal do Rio Grande do Sul; Porto Alegre; Rio Grande do Sul; Brazil

[2] Departamento de Biologia Celular, Embriologia e Genética; Universidade Federal de Santa Catarina; Florianópolis; Santa Catarina; Brazil

[3] Curso de Ciências Biológicas; Universidade LaSalle; Canoas; Rio Grande do Sul; Brazil

[4] Departamento de Genética; Instituto de Biociências; Universidade Federal do Rio Grande do Sul; Porto Alegre; Rio Grande do Sul; Brazil

* These authors contributed equally to this work

Corresponding author: Vera Lucia da Silva Valente

Departamento de Genetica, Instituto de Biociencias, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, Postal Code 91501-970, Porto Alegre, Rio Grande do Sul, Brazil. Phone: +55-51-33166729, Fax: +55-51-33167311. E-mail: vera.gaiesky@ufrgs.br

Abstract

In this study we present the first reference photomaps for *Drosophila cardini* and *D. parthenogenetica* and their first inversions found: 3A for *D. cardini*, 2RA and 3RA for *D. parthenogenetica*. Furthermore, we improved the *D. cardinoides*, the *D. neocardini* and the *D. polymorpha* reference photomaps previously published. Beyond the *D. cardini* and *D. parthenogenetica* populations, we also analyzed the inversion polymorphism for heterozygous rearrangements for different populations of other four species of the *cardini* subgroup, *D. cardinoides*, *D. neocardini*, *D. neomorpha*, and *D. polymorpha*. In these analyses, for *D. polymorpha,* we found the inversions 2RA, 2RC, 2RD and the rearrangement 2RA+C previously described. We also described new inversions in these species, the inversion 3RA for *D. neocardini*, 3LE and 3LF for *D. cardinoides*, 3LC and the rearrangement 2RA+D for *D. polymorpha*. *Drosophila neomorpha* presented a homokaryotypic banding pattern for the unique population here analyzed. After all, we can conclude that for the species populations distributed in south Brazil, *D. polymorpha* is by far the most polymorphic species of *cardini* subgroup; and perhaps, the analysis of different other populations of these species would provide different rearrangements which could contributed to the evolutionary studies of the *cardini* group based on the polytene chromosomes.

Running title: Evolutionary studies in *cardini* group

Key words: *Drosophila cardini* group - reference photomap - chromosomal evolution

Introduction

The publication of the 12 *Drosophila* genomes and the comparative analysis of these sequences with the reference photomap of polytene chromosomes of each species (Schaeffer *et al.,* 2008), gives rise to a new era in the evolutionary studies of the genus (Bhutkar *et al.,* 2008). Several authors used the method of the triads of overlapping inversions (Sturtevant and Dobzhansky, 1936; Dobzhansky and Sturtevant, 1938) to infer phylogenetic relationships of species for some groups of flies: *D. pseudoobscura* populations (Sturtevant and Dobzhansky, 1936), *virilis* group (Throckmorton, 1982), *melanogaster* group (Lemeunier and Ashburner, 1976), *repleta* group (Wasserman, 1982; 1992), the Hawaiian drosophilids (Clayton *et al.,* 1972; Kaneshiro *et al.,* 1995), and *willistoni* subgroup (Rohde *et al.,* 2006). These studies show that the inversions found in polytene chromosomes in *Drosophila* species consists of good tools for the phylogenetic studies with valuable evolutionary information.

The species from the *cardini* group have been studied since 40's with works about their taxonomy and distribution (Sturtevant, 1942; Dobzhansky and Pavan, 1943; Stalker, 1953; Heed and Russell, 1971), crossing patterns (Heed, 1962; 1963), abdominal color polymorphisms (Da Cunha *et al.,* 1953; Heed and Khrishinamurthy, 1959; Hollocher *et al.,* 2000ab; Brisson *et al.,* 2005), chromosomal inversions, and evolutionary relationships (Martinez and Cordeiro, 1970; Heed and Russell, 1971; Brisson *et al.,* 2006). Moreover, recent collections on Neotropical Atlantic Forest have shown that the *cardini* group species have a very expressive presence in the Drosophilidae community (Vilela *et al.,* 2002; De Toni *et al.,* 2007; Chaves and Tidon, 2008). The *cardini* group consists of 16 described species inhabiting different areas of Neotropical America (Heed and Rusell, 1971) and is characterized by highly polymorphic pigmentation appearing (Heed and Krishnamurthy,

1959). Seven of the species belongs to *dunni* subgroup restrictedly distributed in the Caribbean islands; and the other nine species assemble the *cardini* subgroup present from USA to southern Brazil. This group has radiated equally on the Caribbean islands and the neighboring mainland (Heed, 1962) providing valuable tools to study the evolution in islands and continent simultaneously. Therefore, the *cardini* subgroup is a paraphyletic group that diverged 6.6 million years ago (Brisson *et al.,* 2006); and the paraphyletic situation was a consequence of the radiation of the *dunni* subgroup from the mainland to the islands.

The taxonomic studies of the *cardini* group report several difficulties in establishing criteria to separate the species within the group using morphological character only (Streisinger, 1946; Stalker, 1953; Heed and Wheeler, 1957; Vilela *et al.,* 2002). Concerning the molecular data, this group is phylogeneticaly related to the *calloptera*, *guarani* and *tripunctata* groups (Throckmorton, 1975; Grimaldi, 1990; Robe *et al.,* 2005) and the limits among them are not always clear-cut. Some phylogenetic relationships hypotheses within the *cardini* group had previously been proposed by Heed (1962) using male genitalia morphology data, by Heed and Russel (1971) with chromosomal inversions and intercrossing, by Napp and Cordeiro (1981) using electrophoresis pattern of isoenzymes, and by Robe *et al.* (2005) and Brisson *et al.* (2006) using nuclear and mitochondrial genes as molecular markers. However, when these data are compared some inconsistencies are observed specially regarding the *D. neocardini* phylogenetic relationships. According to the hypothesis presented by Robe *et al.*, (2005) and Brisson *et al.,* (2006), *D. neocardini* is phylogeneticaly closer to *D. cardinoides* than to *D. polymorpha*, whereas according to Heed (1962), Heed and Russel (1971) and to Napp and Cordeiro (1981), *D. neocardini* and *D. polymorpha* bear closer phylogenetic affinities.

These differences may be due to different selective constrains suffered by these different markers.

In this study, we analyzed the polytene chromosomes of six species from the *cardini* subgroup; *D. cardini*, *D. cardinoides*, *D. neocardini*, *D. neomorpha*, *D. parthenogenetica* and *D. polymorpha*; and we present improved reference photomaps of *D. cardinoides*, *D. neocardini* and *D. polymorpha*. Also, we analyzed new heterozygous inversions data found in these species, and we present a first step to establish a clear chromosomal evolution pattern in these species. Differently from Heed and Russel (1971), here we show the banding similarities between the polytene chromosomes, which are an essential data to comprehend the chromosomal evolutionary pattern.

Materials and Methods

Fly stocks

Table 1 lists the information about the species and populations strains studied in this work; all of them have been maintained in the laboratory and reared in corn flour culture medium (Marques *et al.*, 1966) in a controlled chamber (17°C ± 1°C, 60% rh). The strains were established using three fertile females collected from nature and, for the strains maintained as a pool in the laboratory, we performed at least five isofemale strains. To ensure the identification of the species kindly sent by our collectors' collaborators and by the National *Drosophila* Species Research Center (NDSRC), we reanalyzed the stock using the available literature about external morphologies and genitalia characters.

Table 1: Information about the species and populations analyzed in this study

Species	Locality	N. of isostrains	Coordinate	Year of collection/Collector
D. cardini	Itaqui – RS[a]	05	29º 08'S 56º 33'W	2004/MD Fantinel
	Porto Alegre – RS[a]	05	30º 07'S 51º 09'W	2007/HJ Schmitz
	Paulo Faria – SP	07	20º 01'S 49º 24'W	2006/L Maddi-Ravazzi
	Serra do Japi – SP	06	23º 14'S 46º 58'W	2006/C Rohde
	Serra do Cipo – MG	06	19º 12'S 43º 42'W	2005/C Rohde
	Tangara da Serra – MT	05	14º 37'S 57º 29'W	2007/MS Gottschalk
	Cabralia – BA	03	12º 36'S 41º 24'W	2006/C Rhode
D. cardinoides	Florianopolis – SC[a]	05	27º 35'S 48º 22'W	2001/MS Gottschalk
	Porto Alegre – RS[a]	05	30º 03'S 51º 07'W	2002/MS Gottschalk
	São Sebastiao – SP	05	23º 47'S 45º 33'W	2005/C Rohde
	Itaunas – ES	02	18º 24'S 39º 41'W	2005/C Rohde
	Grao Mogol – MG	04	16º 33'S 42º 53'W	2005/L Maddi-Ravazzi
	Serra do Cipo – MG	04	19º 12'S 43º 42'W	2006/C Rohde
	Mucuri – BA	03	18º 05'S 39º 35'W	2005/C Rohde
D. neocardini	Ilha do Arvoredo – SC[a]	05	27º 16'S 48º 21'W	1999/ DC De Toni
	Ilha do Campeche – SC[a]	05	27º 41'S 48º 27'W	2006/ DC De Toni
	Morro do Bau – SC	02	26º 46'S 49º 02'W	2006/ DC De Toni
	Florianopolis – SC	07	27º 35'S 48º 22'W	2002/ MS Gottschalk
	Joinville – SC	04	26º 17'S 49º 00'W	2004/J Döge
	Mucuri – BA	02	18º 05'S 39º 35'W	2005/C Rohde
D. neomorpha	Joinville – SC	05	26º 17'S 49º 00'W	2007/DC De Toni
D. parthenogenetica	Ilha do Campeche – SC	05	27º 41'S 48º 27'W	2006/DC De Toni
	Sinaloa – Mexico[a]	05	15181-2221.00[b]	1971/W Heed
	Los Hornos – Mexico[a, c]	05	--	1971/W Heed
D. polymorpha	Ilha de Ratones Grande – SC	03	27º 28'S 48º 33'W	2005/DC De Toni
	Ilha do Campeche – SC	07	27º 41'S 48º 27'W	2005/DC De Toni
	Florianopolis – SC[a]	05	27º 35'S 48º 22'W	2002/MS Gottschalk
	Joinville – SC	07	26º 17'S 49º 00'W	2005/J Döge
	Porto Alegre – RS[a]	05	30º 03'S 51º 07'W	2002/MS Gottchalk
	Turvo – RS[a]	05	27º 13'S 53º 51'W	1999/L Basso
	Santa Maria – RS	05	29º 42'S 53º 47'W	2003/J Cordeiro
	Sao Sebastiao – SP	02	23º 47'S 45º 33'W	2005/C Rohde
	Serra do Cipo – MG	02	19º 12'S 43º 42'W	2005/C Rohde
	Cabralia – BA	01	12º 36'S 41º 24'W	2005/C Rohde
	Itauna – ES	01	18º 24'S 39º 41'W	2005/C Rohde

[a]strains maintained as pool in our laboratory.

[b]National *Drosophila* Species Research Center (NDSRC) stock number

[c]strain provided by Dr Hope Hollocher from University of Notre Dame – Indiana/USA.

Cytological preparations

The salivary glands of third instar female larvae were prepared according to Ashburner (1989, p30) using Ringer's solution for dissection. Among three and five female larvae were analyzed for each isofemale strain. These strategies were used in an effort to sample all inversions present in each species. The brain ganglia metaphase chromosomes of the species were prepared according to Santos-Colares *et al.,* (2002).

Polytene chromosome reference photomaps

The slides were examined using a Zeiss photomicroscope with phase contrast at 1,000x magnification and images were taken on Kodak Asa 100 color print film. After comparing the banding pattern of each chromosome within each species strains, we chose one homokaryotipic strain from Serra do Cipo/Brazil from *D. cardini* and the Sinaloa/Mexico strain from *D. parthenogenetica* to produce their first polytene reference photomap. The identification of the chromosomal complement of these species were done comparing the banding pattern of each polytene chromosome between the species that already have their reference photomaps (*D. cardinoides*, *D. polymorpha* – Rohde and Valente, 1996 – *D. neocardini* and *D. neomorpha* – De Toni *et al.,* 2001a; 2006). Aiming the improvement of the *D. cardinoides*, *D. neocardini* and *D polymorpha* photomaps; we used different photomicrographs, from different strains, that were representative of the banding pattern of each polytene chromosome; following the banding pattern of previous published photomaps of each species (Rohde and Valente, 1996; De Toni *et al.,* 2001a). In the comparative analyses of the polytene chromosomes we used the photomaps here produced and improved, and the *D. neomorpha* photomap published by De Toni *et al.,* (2006). The best images were captured using a film scanner and processed with the

Adobe® Photoshop® program (Adobe® Photoshop® CS3 Extended, v 10.0.1, San Jose, California, Adobe Systems Inc.).

Results

Metaphase chromosomes characterization

Figure 1 depicts the metaphase chromosomes of *D. cardini* (A and B) and *D. parthenogenetica* (C and D) where it is possible to observe that the first species has 2n=12, being all acrocentric chromosomes; and the second species has 2n=8, with metacentric autosomes and acrocentric sexual chromosomes, as the other species of the group. The recognition of the XY pair in the metaphase preparations was facilitated by the analysis of two characteristics: the heteropicnotic state of the Y chromosome, and the tendency of the homologous chromosomes to be next to each other. Regarding the species from the *cardini* subgroup, only *D. parthenogenetica* shows a visible secondary constriction in the X chromosomal pair (Figure 1D), as also observed by Stalker (1953). However, it is known that this kind of constriction is associated with rRNA synthesis (Sumner, 2003); otherwise it can also be used as a structural chromosomal marker to differentiate this species from its related species. As *D. cardini* is believed to be the basal species of the *cardini* group (Heed and Russell, 1971; Brisson *et al.*, 2006), the chromosomal composition of the other species might have been formed through fusions of the ancestral species chromosomes (Heed, 1962; Heed and Russell, 1971).

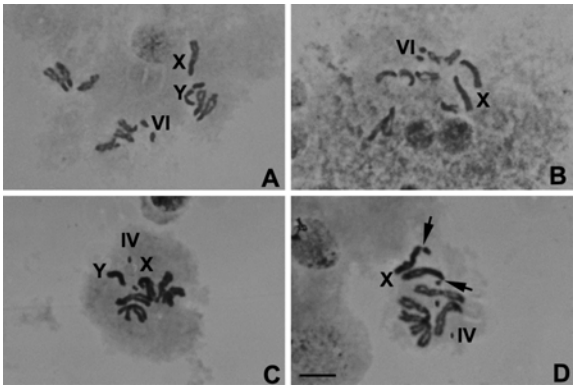


Figure 1: Metaphase chromosomes of *D. cardini* male (A) and female (B), and *D. parthenogenetica* male (C) and female (D). Arrows indicate the secondary constriction on *D. parthenogenetica* X chromosomes. Bar = 10µm.

Reference photomaps of polytene chromosomes

The assembly of the photomaps followed a rigorous comparison of each chromosome within each species strains making possible to identify the exact chromosomal banding pattern of each chromosome. Also we conducted an analysis of each chromosome between the species aiming the correct identification of each chromosome. In order to promote the precise localization of the banding pattern and of the inversions break points, the polytene chromosomes of all species were divided into 20 sections each, always beginning from the tip of the chromosomes, and subdivided into "a" and "b" or "a", "b" and "c" subsections, based on banding similarities between the species and based on the division of the previous photomaps performed by Rohde and Valente (1996) and by De Toni *et al.* (2001a). The V chromosome of *D. cardini*, the IIIR chromosomal arms of the other species and the dot chromosomes of all of them are the exceptions, where the first are divided into 19 sections and the latter is the 100th section for all species. In this way, the improved reference photomaps of *D. cardinoides*, *D. polymorpha* (Rohde and Valente, 1996) and *D. neocardini* (De Toni *et al.*, 2001a) are shown in Supplementary Material Figure 1S. Figure 2 shows the first reference photomaps of *D. cardini* (Figure 2A) and *D. parthenogenetica* (Figure 2B).

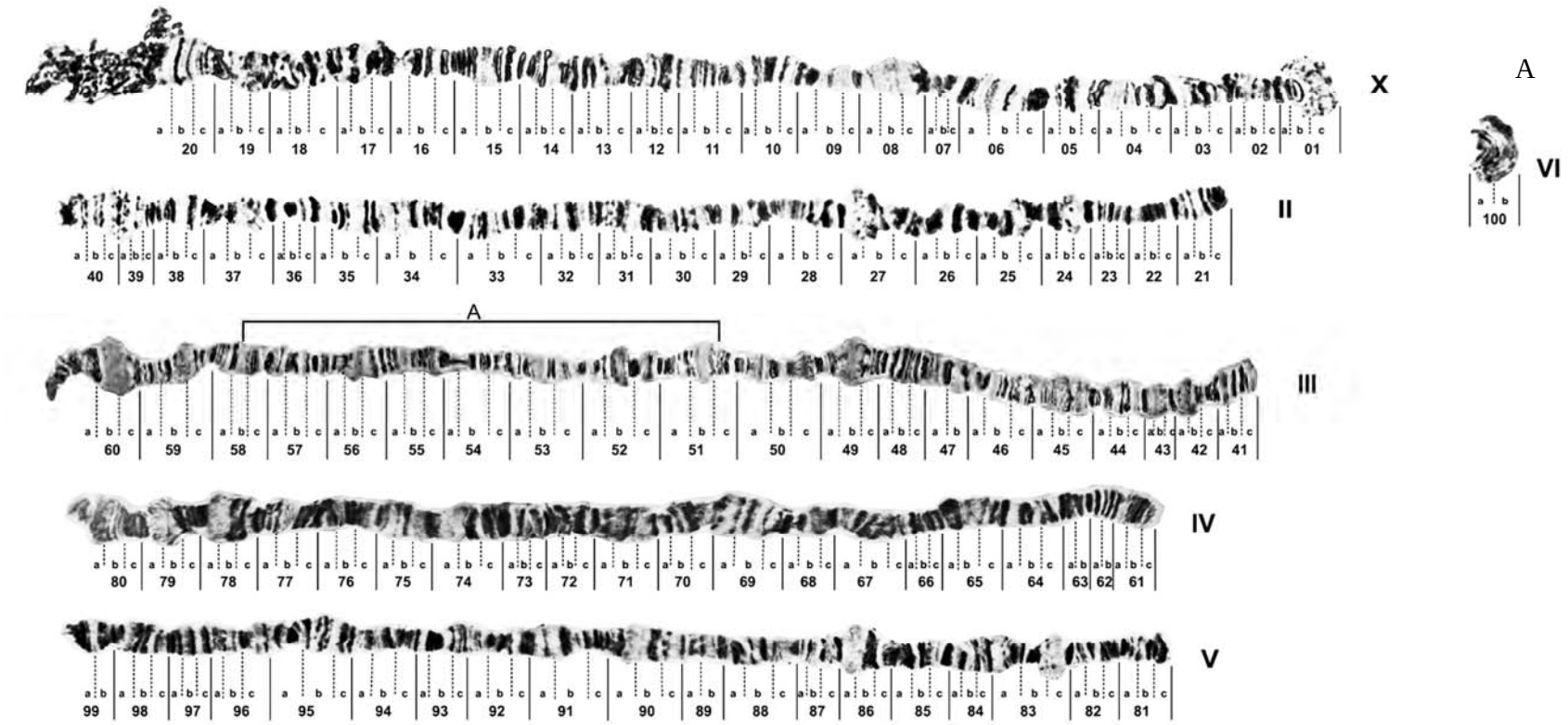


Figure 2. Reference photomap of *D. cardini* (A) and *D. parthenogenetica* (B). The distal regions of the chromosomes (base) are in the right side of the figure. The lines indicate the break points of the inversion IIIA of *D. cardini* and inversions IIRA and IIIRA of *D. parthenogenetica*.

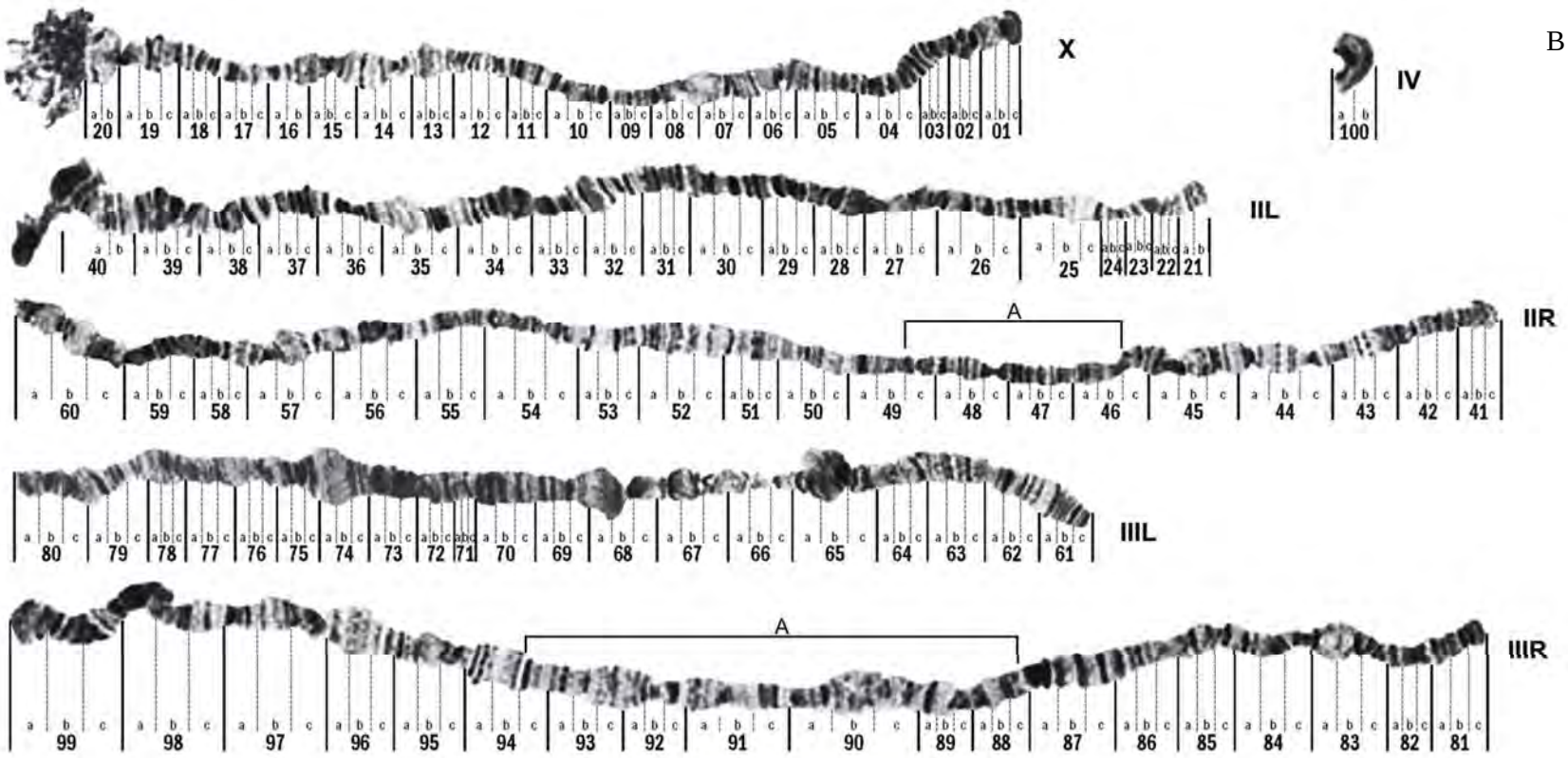


Figure 2B: Reference photomap of *D. parthenogenetica*. See legend in the previous page.

Inversion polymorphism

Although Heed and Russel (1971) have studied several inversions in the *cardini* group, it is not possible to compare their data with the data obtained in this study, as there is no demonstration of where the inversions analyzed by that authors are located in the polytene chromosomes of the *cardini* group species. Considering this, here we are describing the first inversion for *D. cardini* and two first inversions for *D. parthenogenetica* (Table 2 and Figure 3). Here, we also are describing two new inversions for *D. cardinoides*, one for *D. neocardini* and one new inversion rearrangement for *D. polymorpha* (Table 2 and Figure 3 and 4). Table 2 summarizes the information about the heterozygous inversions observed in this study with their respective chromosomal localization; Figure 3 shows these inversions found in the species *D. cardini*, *D. cardinoides*, *D. neocardini*, *D. parthenogenetica* and Figure 4 shows the inversions found in *D. polymorpha* here analyzed. Table 3 shows which populations of each species presented each of the inversions found in this study. For the sake of simplicity, the inversions nomenclature follows the Arabic numbers, instead of the Roman numbers used for the chromosomes.

Table 2: Heterozygous inversions observed in this study for the five species of the *cardini* group analyzed, with their chromosomal localization.

Species	Heterozygous inversion	Chromosome localization	Figure where it is shown
D. cardini	3A*	51c – 58b	3A
D. neocardini	3RA*	83c – 91a	3B
D. cardinoides	3LE*	62a – 68c	3C
	3LF*	69b – 77c	3D
D. parthenogenetica	2RA*	46b – 49b	3E
	3RA*	84b – 88b	3F
D. polymorpha	2RA[a]	48a – 53a	4A
	2RC[b]	47c – 48a	4C
	2RD[b]	45c – 47a	4B
	3LC*	63b – 68a	4F
	2RA + C[b]		4D
	2RA + D*		4E

*****new inversions found for these species in this study

[a] inversion first descried in Rohde and Valente (1996)

[b] inversions first described in De Toni *et al*., 2001a

[c] inversion first described in De Toni, 2002

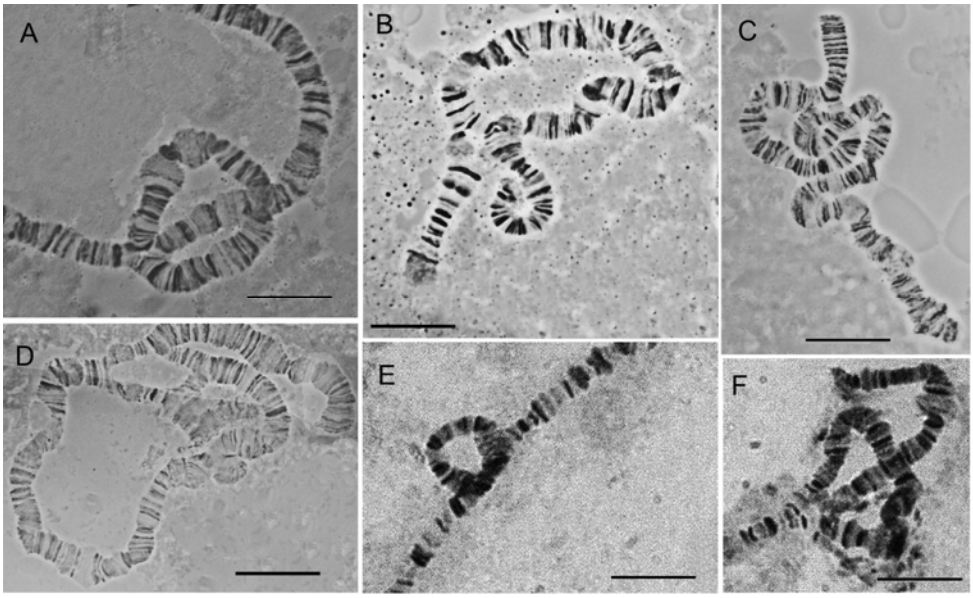


Figure 3: Heterozygous inversions found in the polytene chromosomes of populations of: *Drosophila cardini*, 3A (A); *D. neocardini*, 3RA (B), *D. cardinoides*, 3LE (C) and 3LF (C) and of *D. parthenogenetica*, 2RA (E) and 3RA (F). Bar = 10µm.

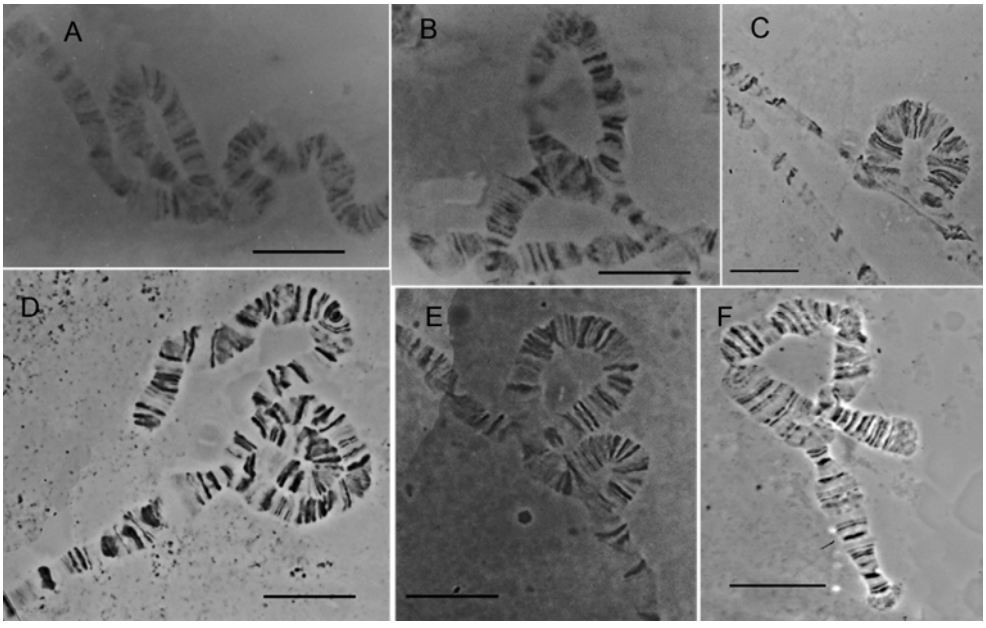


Figure 4: Heterozygous inversions found in the polytene chromosomes of *Drosophila polymorpha* populations: 2RA (A), 2RD (B), 2RC (C), 2RA + C (D), 2RA + D (E) and 3LC (F). Bar = 10µm.

Table 3: Populations of the *Drosophila cardini* species group where the inversions and arrangements were detected in the present study.

Species	Locality	Inversions
D. cardini	Itaqui – RS 29°08'S 56°33'W	3A
	Porto Alegre – RS 30º04'S 51º08W	3A
	Paulo Faria – SP 20°01'S 49°23'W	3A
	Serra do Japi – SP 23º14'S 46º58'W	3A
	Serra do Cipo – MG 18°59'S 43°38'W	3A
	Tangara da Serra – MT 14°37'S 57°29'W	3A
	Cabralia – BA 12°50'S 41°50'3W	3A
D. cardinoides	Florianópolis – SC 27º35'S 48º33'W	3LE
	Porto Alegre – RS 30º04'S 51º08W	3LE; 3LF
	Sao Sebastiao – SP 23°45'S 45°24'W	st
	Itaunas – ES 18°25'S 39°42'W	st
	Grao Mogol – MG 16°33'S 42°53'W	st
	Serra do Cipo – MG 18°59'S 43°38'W	st
	Mucuri – BA 18°03'S 39°32'W	st
D. neocardini	Ilha do Arvoredo – SC 27º17'S 48º21'W	st
	Ilha do Campeche – SC 27º35'S 48º22'W	st
	Morro do Bau – SC 26°46'S 48°54'W	3RA
	Florianópolis – SC 27º35'S 48º33'W	3RA
	Joinville – SC 26º17'S 49º00'W	3RA
	Mucuri – BA 18°03'S 39°32'W	st
D. neomorpha	Joinville – SC 26º17'S 49º00'W	st
D. parthenogenetica	Ilha do Campeche – SC 27º35'S 48º22'W	st
	Sinaloa – Mexico 28°46'N 110°25'W	st
	Los Hornos – Mexico 24°21'N 107°29'W	2RA; 3RA
D. polymorpha	Ilha de Ratones Grande – SC 27°28'S 48°33'W	2RA; 2RD
	Ilha do Campeche – SC 27º35'S 48º22'W	2RA
	Florianópolis – SC 27º35'S 48º33'W	2RA; 2RA+D
	Joinville – SC 26º17'S 49º00'W	2RA
	Porto Alegre – RS 30º04'S 51º08W	2RA; 2RC
	Turvo – RS 27º20'S 53º40'W	2RA; 2RA+C; 3LC
	Santa Maria – RS 29º32'S 53º42'W	2RA; 2RC; 3LC
	Sao Sebastiao – SP 23°45'S 45°24'W	2RA
	Serra do Cipo – MG 18°59'S 43°38'W	2RA
	Cabralia – BA 12°50'S 41°50'3W	2RA
	Itauna – ES 18°25'S 39°42'W	2RA

st: population with the standard band/interband rearrangement for all chromosomal arms

In the analysis of these data we can observe some characteristics of the inversion pattern shown by each population species. First, all *D. cardini* populations presented the inversion 3A, showing that this inversion seems to be well widespread. Second, only the *D. parthenogenetica* population from Los Hornos (Mexico) is polymorphic for inversions, although the other Mexican population, Sinaloa, is homokariotypic. Third, all populations of *D. polymorpha* have the 2RA rearrangement, from south (Rio Grande do Sul) to north (Bahia) of the distribution here studied. In this work it is the first time that the rearrangement 2RA+D is observed in this species, and it was found only in a population distributed in a restinga area of Florianopolis. Moreover, the rearrangement 2RA+C was found only in the *D. polymorpha* population from Turvo and the 2LC rearrangement was only in the Santa Maria population. Fourth, each of the new inversions here analyzed was found only in one of the species populations, in contrast to the inversions 3LE and 3LF of *D. cardinoides* that were found in the same population from Porto Alegre. Fifth, there were some species populations free of inversions, but *D. polymorpha* (see Table 3). We also could not observe any inversion in the only *D. neomorpha* population here analyzed from Joinville.

The IIIL chromosomal arm seems to be the most polymorphic for heterozygous inversions both for *D. cardinoides* and *D. neocardini*, with six and two inversions respectively; differently for *D. polymorpha* in which the IIIR chromosomal arm was the most polymorphic, with five inversions segregating. *Drosophila cardini*, *D. neomorpha* and *D. parthenogenetica* do not seem to be very polymorphic for heterozygous inversions, at least for the populations here analyzed.

Characterization and comparison of the polytene chromosomes

The comparisons of the chromosomes among the species were done using the new and improved photomaps, and the whole amount of photographs obtained in this study. For *D. neomorpha* we also used the published photomap (De Toni *et al.,* 2006). Intrinsic error may exist in the pairwise chromosomal analysis, one of them being the observational mistakes (Wasserman, 1992). To overcome this, we performed at least six times the same pairwise analysis but with different observer. Although *D. cardini* is the basal species of the *cardini* group (Brisson *et al.,* 2006); in general, the banding pattern of the polytene chromosomes of the six species here studied resembles the ones of *D. cardinoides*. In the comparisons involving *D. neomorpha* and *D. parthenogenetica* we have some difficulty to observe similarities between the banding patterns in the pairwise comparisons with the others species and between themselves. This could be a consequence of the proper difficulties of maintenance (rearing) of these species strains, where the attainment of good chromosomes becomes extremely laborious. So, for sake of simplicity, considering the metaphase chromosomes difference numbers we adopted the following designation for the chromosomes: chromosomal Roman number of *D. cardini* and chromosomal arm Roman number of the other species related chromosomal arm separated by a bar, example: II/IIL chromosomes.

Figure 5 shows the side by side positioning of each chromosome aiming the compared analysis of the banding pattern among the species, which amounted in 75 comparisons (Table 4; Supplementary Material Figure 2S). From these analyses some similar banding pattern emerged. In general the tip and the base (centromeric region) regions resemble each other, comparing the same chromosome of the different species;

however, throughout the chromosomes, small regions were detected that can be used as hallmarks, helping their recognition.

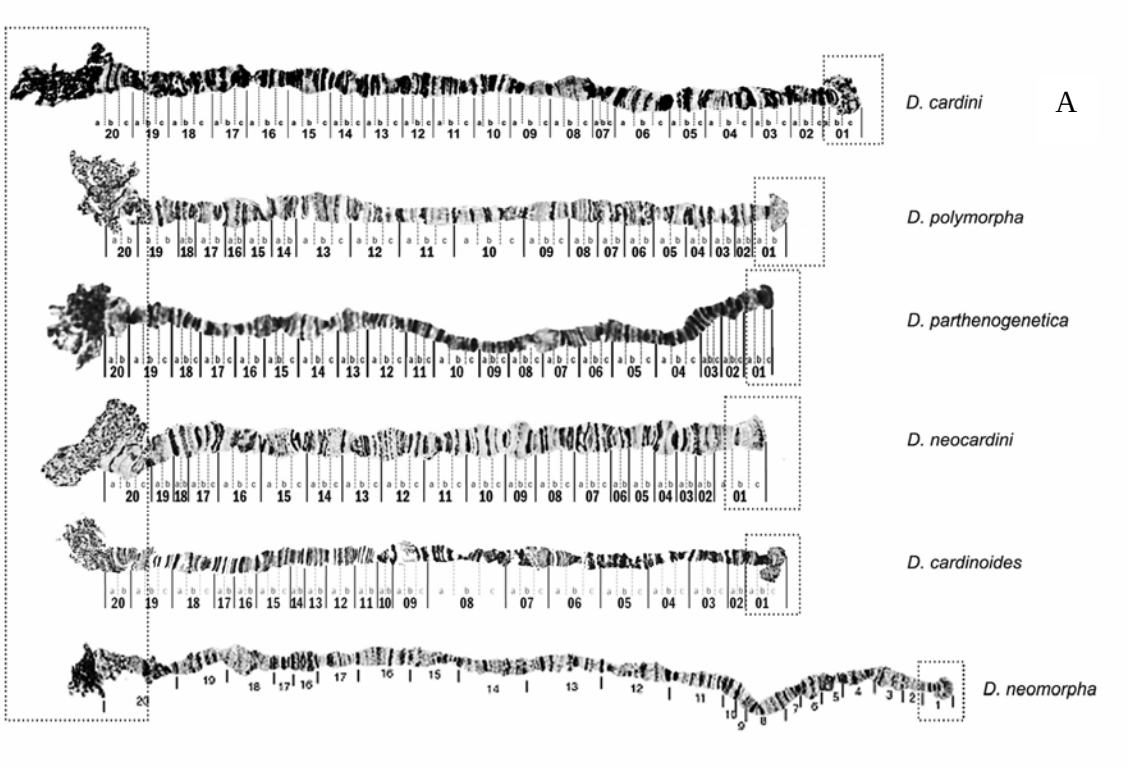


Figure 5. Comparisons among the chromosomes of the six *cardini* subgroup species. (A) X chromosomes. (B) Chromosome II of *Drosophila cardini* and IIL chromosomal arms of the other five species. (C) Chromosome III and IIR chromosomal arms. (D) Chromosome IV and IIIL chromosomal arms. (E) Chromosome V and IIIR chromosomal arms. (F) Chromosome VI and IV. The distal regions of the chromosomes (base) are in the right side. The dotted lines highlight the information about similarities discussed in the text. * In (B) and (C): regions supposed to be involved in an ancient pericentric inversion (see text).

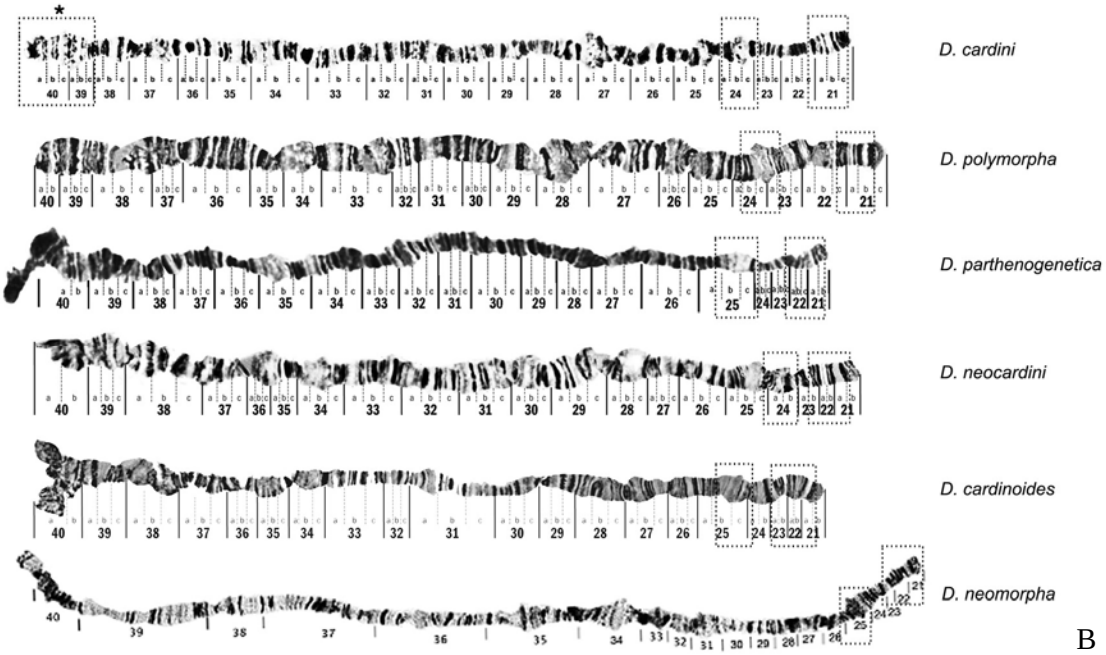


Figure 5B: Comparisons among the chromosomes of the six *cardini* subgroup species. Legend is in the previous page.

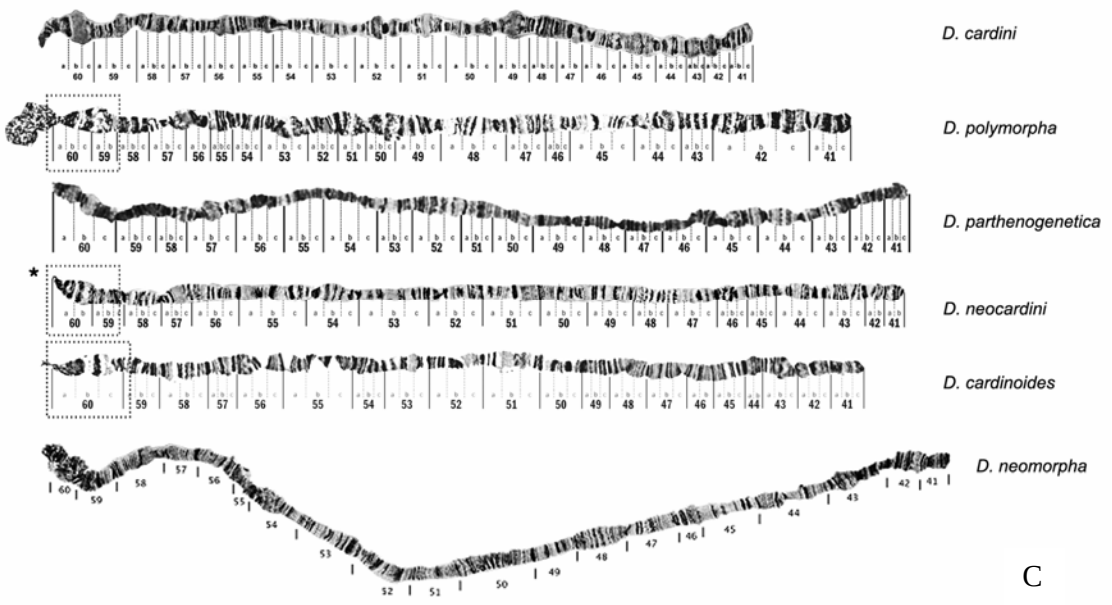


Figure 5C: Comparisons among the chromosomes of the six *cardini* subgroup species. Legend is in the previous page.

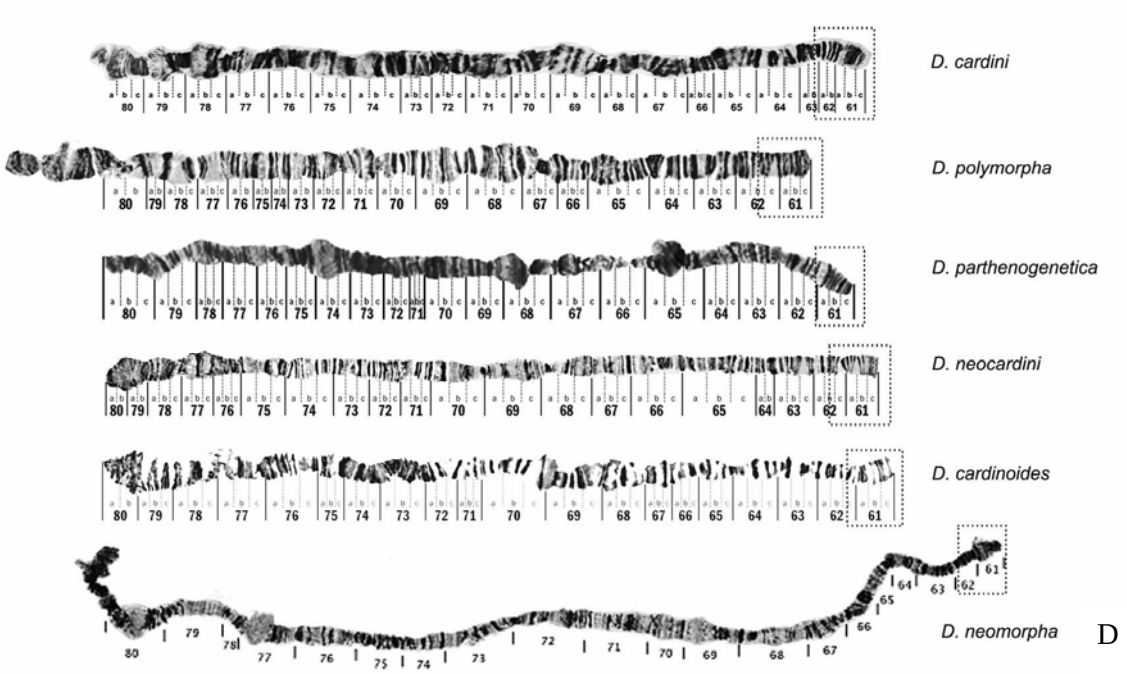


Figure 5D: Comparisons among the chromosomes of the six *cardini* subgroup species. Legend is in the previous page.

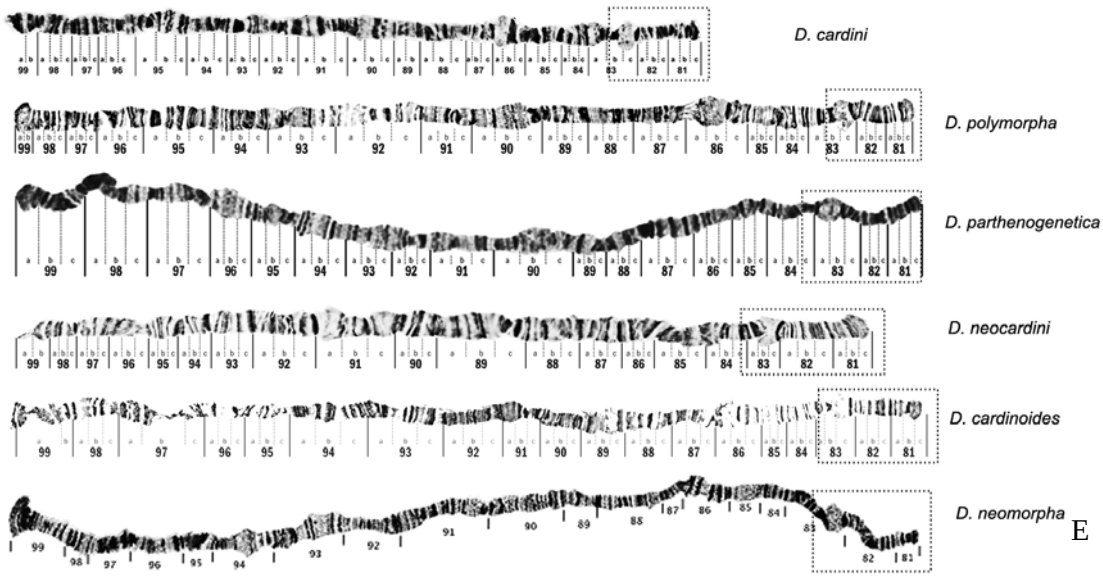


Figure 5E: Comparisons among the chromosomes of the six *cardini* subgroup species. Legend is in the previous page.

F

Figure 5F: Comparisons among the chromosomes of the six *cardini* subgroup species. Legend is in the previous page.

Table 4: Pairwise analysis of the chromosomes sections among the five species from *cardini* group analyzed

Chromosome arm	Pairwise comparison		Chromosome subsections		
	First species	Second species	tip	middle	base
X	*D. cardini*	*D. polymorpha*	1c-1a→1c-1a	--	20c-20a→20c-20a
		D. cardinoides	1c-1a→1c-1a	--	20c-20a→20c-20a
		D. neocardini	1c-5c→1c-6b	--	16a-20a→16c-20a
		D. neomorpha	1c-6a→1c-8b	--	19c-20a→19c-20a
		D. parthenogenetica	1c-2a→1c-1a	--	15c-20a→12a-20a
	D. polymorpha	*D. cardinoides*	1c-10b→1c-8b	--	20c-20a→20c-20a
		D. neocardini	1c-9c→1c-9c	--	18b-20a→17c-20a
		D. neomorpha	1c-6a→1c-6a	--	19a-20a→20c-20a
		D. parthenogenetica	1c-3b→1c-2a	12b-14a←5a-8b	18b-20a→18b-20a
	D. cardinoides	*D. neocardini*	--	--	--
		D. neomorpha	1c-3b→1c-4b	--	18a-20a→18a-20a
		D. parthenogenetica	1c-7a→1c-8a	--	17a-20a→18a-20a
	D. neocardini	*D. neomorpha*	1c-3c→1c-3c	--	20c-20a→20c-20a
		D. parthenogenetica	1c-4a→1c-4c	--	17a-20a→18a-20a
	D. neomorpha	*D. parthenogenetica*	1c-3c→1c-3c	16c-18a←7c-9a	19c-20a→20c-20a
II/IIL	*D. cardini*	*D. polymorpha*	21c-27a→21c-28a	28a-32b←31b-36a	37a-40a→37c-40a
		D. cardinoides	21c-26a→21c-27b	34c-34a←28c-29a	36b-40a→37a-40a
		D. neocardini	21c-32b→21c-32c	32b-37b←33c-37b	--
		D. neomorpha	21c-25b→21c-28b	--	37b-40a→39a-40a
		D. parthenogenetica	21c-27a→21c-28b	--	37c-40a→38c-40a
	D. polymorpha	*D. cardinoides*	21c-27b→21c-27a	33b-33a→32a-33b	39b-40a→39a-40a
		D. neocardini	21c-27c→21c-28c	33b-35a→33b-35c	38a-40a→38a-40a

Table 4: Continuation.

Chromosome arm	Pairwise comparison		Chromosome subsections		
	First species	Second species	tip	middle	base
II/IIL	*D. polymorpha*	*D. neomorpha*	21c-25c→21c-27c	33b-33a→31c-31b	35a-40a→37b-40a
		D. parthenogenetica	21c-25a→21c-26a	--	35a-40a→36a-40a
	D. cardinoides	*D. neocardini*	21c-30b→21c-30a	31c-37a←31c-37a	--
		D. neomorpha	21c-25b→21c-25a	30b-31a←35b-36b	39a-40a→40c-40a
		D. parthenogenetica	21c-28b→21c-29c	31b-32a→32b-33a	38c-40a→38a-40a
	D. neocardini	*D. neomorpha*	21c-24a→21c-25a	--	--
		D. parthenogenetica	21c-22a→21c-22a	--	38c-40a→38a-40a
	D. neomorpha	*D. parthenogenetica*	21c-25a→21c-25a	--	--
III/IIR	*D. cardini*	*D. polymorpha*	41c-43a→41c-42c	--	59c-60a→57a-60a
		D. cardinoides	41c-43a→41c-43c	53c-57a←50c-53a	60b-60a→60c-60a
		D. neocardini	--	52b-57c←51b-54a	--
		D. neomorpha	--	--	58b-60a→58b-60a
		D. parthenogenetica	41c-44a→41c-44c	46a-50a→55b-58b	60c-60a→59c-60a
				52c-55b→45a-48a	
	D. polymorpha	*D. cardinoides*	41c-44b→41c-44a	53a-55a←48a-50a	57a-60a→58a-60a
		D. neocardini	41c-44a→41c-47c	--	57c-60a→57b-60a
		D. neomorpha	41c-42a→41c-43a	--	59b-60a→59c-60a
		D. parthenogenetica	41c-42a→41c-43a	56c-57a→56c-57a	58b-60a→59b-60a
	D. cardinoides	*D. neocardini*	41c-44a→41c-44c	--	56b-60a→56b-60a
		D. neomorpha	41c-43a→41c-44c	--	59c-60a→59c-60a
		D. parthenogenetica	41c-43a→41c-43b	--	59c-60a→59c-60a

Table 4: Continuation.

Chromosome arm	Pairwise comparison		Chromosome subsections		
	First species	Second species	tip	middle	base
III/IIR	*D. neocardini*	*D. neomorpha*	41c-43c→41c-42b	46a-49c→46c-48a 53c-53a→53c-53a	57c-60a→57c-60a
		D. parthenogenetica	41c-44c→41c-44c	54c-55b←47b-48a	57a-60a→58b-60a
	D. neomorpha	*D. parthenogenetica*	--	--	--
IV/IIIL	*D. cardini*	*D. polymorpha*	61c-64a→61c-63a	68b-70b→67b-68b	75a-80a→75a-80a
		D. cardinoides	61c-66a→61c-64a	--	--
		D. neocardini	61c-66a→61c-65c	--	77a-80a→76c-80a
		D. neomorpha	61c-65a→61c-66a	--	77b-80a→77c-80a
		D. parthenogenetica	61c-62a→61c-62a	--	78b-80a→78a-80a
	D. polymorpha	*D. cardinoides*	61c-69b→61c-69a	--	77b-80a→77b-80a
		D. neocardini	61c-67a→61c-67c	--	--
		D. neomorpha	61c-62a→61c-64a	--	80b-80a→78b-80a
		D. parthenogenetica	61c-62a→61c-62b	70b-75a←67c-69a	77c-80a→78b-80a
	D. cardinoides	*D. neocardini*	61c-64b→61c-64b	--	77b-80a→76c-80a
		D. neomorpha	--	--	--
		D. parthenogenetica	--	--	--
	D. neocardini	*D. neomorpha*	--	--	--
		D. parthenogenetica	61c-63b→61c-63a	--	--
	D. neomorpha	*D. parthenogenetica*	--	--	77b-80a→74a-80a
V/IIIR	*D. cardini*	*D. polymorpha*	81c-86a→81c-87c	--	92a-99a→93a-99a
		D. cardinoides	81c-86a→81c-86a	--	93b-99a→94a-99a

Table 4: Continuation.

Chromosome arm	Pairwise comparison		Chromosome subsections		
	First species	Second species	tip	middle	base
V/IIIR		*D. neocardini*	81c-84c→81c-84c	--	96c-99a→96c-99a
		D. neomorpha	--	--	--
		D. parthenogenetica	81c-83b→81c-84c	--	--
	D. polymorpha	*D. cardinoides*	81c-86a→81c-86a	90c-93a→89c-92a	--
		D. neocardini	81c-85a→81c-84b	90c-91c→89c-90b	95b-99a→94b-99a
		D. neomorpha	--	--	--
		D. parthenogenetica	--	86b-89c←93a-96a	97c-99a→98a-99a
	D. cardinoides	*D. neocardini*	81c-84b→81c-84b	87a-89a→90a-92c	97b-99a→96c-99a
		D. neomorpha	--	93b-94a→92a-93a	--
		D. parthenogenetica	81c-84c→81c-85c	91b-93b→92c-94a	--
	D. neocardini	*D. neomorpha*	81c-83b→81c-83b	--	--
		D. parthenogenetica	81c-85c→81c-85c	--	--
	D. neomorpha	*D. parthenogenetica*	81c-83a→81c-84c	85c-88b→86c-89b	--
				91b-93b→92b-95b	

→ not inverted section
← inverted section
-- region of difficult comparison analysis

The X chromosome of all species (Figure 5A) is easily recognized by the puffed band present in the first section, and by the 20th section, the last section of this chromosome, that remains attached to the chromocentric heterochromatin. In all pairwise comparisons for this chromosome, only in the analysis between *D. cardinoides* and *D. neocardini* we were unable to identify precise banding similarities. This situation also occurred in most comparisons in the middle part of this chromosome. Table 4 summarizes the sections with clear similarities found among the species for all chromosomes arms, indicating the observed inverted and the not inverted sections; and the Supplementary Material Figure 2S shows a schematic representation of the data in Table 4. For instance, in the analysis of the X chromosomes between the *D. polymorpha* and *D. parthenogenetica* we found one region with inverted similarity: sections from 12b to 14a that in *D. polymorpha* slightly matches with sections from 5a to 8b of *D. parthenogenetica* in an inverted direction. Moreover, the only species that presented inversions in this chromosome is *D. polymorpha* with the XA, XB and XC inversions.

The chromosome II/IIL (Figure 5B) was divided from section 21 to 40; and in general, the first section of this chromosome share the same pattern among the species, which includes the strong bands in the tip of the chromosome followed by a large interband, and three strong bands. Moreover, the IIL chromosome shares a “hexagonal” puffed band (hallmark) observed on section 23 of *D. polymorpha*, 24 of *D. cardini* and *D. neocardini*, 25 of *D. cardinoides*, *D. neomorpha* and *D. parthenogenetica*. As observed for all chromosomes, the base of this chromosome between the species remains similar, although the similarity of this region in the comparisons between *D. cardini* and *D. neocardini, D. cardinoides* and *D. neocardini, D. neocardini* and *D. neomorpha* and *D. neomorpha* and *D. parthenogenetica* was difficult to establish. After the analysis of the

photomicrographs of the base region in different strains of *D. cardini* and *D. neocardini*, we could visualize a major similarity of the base banding pattern of sections 39 and 40 of *D. cardini* chromosome II (Figure 5B) with the sections 59 and 60 of *D. neocardini* chromosome IIR (Figure 5B and 5C, indicated by an *); suggesting the occurrence of a pericentric inversion between chromosomes II/IIL and III/IIR in the ancestral of *D. neocardini*, considering that *D. cardini* is the basal species of the group. In the search for inverted gene sequences, we found some sequences with similar banding pattern, as was the case of the other chromosomes. See Table 4 and Figure 2S for all the inverted or not inverted regions. Indeed, *D. polymorpha* is the only species with inversions found in this chromosomal arm: 2LA and 2LB.

In the III/IIR chromosome (Figure 5C), which was divided from section 41 to 60, the banding regions with relevance for their identification among species seem to be absent. However, this chromosome follows the general pattern of similar tips and bases; and the species *D. polymorpha, D. neocardini* and *D. cardinoides* share great similarity in their base region (Figure 5C). Nevertheless, the comparisons between *D. cardini* and *D. neocardini*, *D. cardini* and *D. neomorpha*, and between *D. neomorpha* and *D. parthenogenetica* seems to be difficult, preventing the establishment of major similarities. In this chromosome, heterozygous inversions were observed for *D. cardini*, its unique widespread inversion analyzed in this study, 3A; *D. parthenogenetica*, one of the two inversions 2RA and *D. polymorpha*, the widespread inversion 2RA, the inversion 2RB that although is not present in our samples it was observed in other populations (De Toni *et al*., 2001b), the 2RC present in the populations from the southernmost distribution of this species, Porto Alegre, Turvo and Santa Maria; and 2RD found in the population from Ratones Grande Island as also observed by De Toni *et al*. (2001b) (see Table 2; Figure 4).

The IV/IIIL chromosomes were divided in sections 61 (tip) to 80 (base). This chromosome is identified by the several bands and interbands with approximately the same width in the first sections (Figure 5D). The base region among the species resembles each other with no hallmark more evident. In the pairwise comparisons between *D. cardini* and *D. cardinoides*, *D. polymorpha* and *D. neocardini*, *D. neocardini* and *D. parthenogenetica* it was difficult to establish similarities in some parts of the chromosome (Table 4). However, regarding the comparisons between *D. cardinoides* and *D. neomorpha*, *D. cardinoides* and *D. parthenogenetica*, *D. neocardini* and *D. neomorpha*, the whole chromosome seems to be reorganized in a way that becomes hard to identify the band/interband similarities. In general, neither similarities in the middle of the chromosomes (Table 4), nor in the not inverted or in the inverted direction were detected in the comparisons. Considering all species together, this chromosome is the most polymorphic in the *cardini* subgroup with 11 heterozygous inversions. *Drosophila cardinoides* has four inversions already described: 3LA, 3LB, 3LC, 3LD (Rohde and Valente, 1996), and two new inversions found in our study, 3LE and 3LF, for populations from Florianopolis and Porto Alegre. *Drosophila neocardini* has two inversions, 3LA and 3LB (De Toni *et al*., 2001a; De Toni, 2002); and *D. polymorpha* has three, 3LA, 3LB (De Toni, 2002) and 3LC the new one found in this study present in the southern populations of Turvo and Santa Maria.

In the V/IIIR chromosome, that was divided in sections 81 to 99, the tip of all chromosomes resembles each other in the sections 81 to 83 (Figure 5E). In the comparison between *D. cardini* and *D. parthenogenetica*, *D. neocardini* and *D. neomorpha*, *D. neocardini* and *D. parthenogenetica* only the tip region of the chromosomes showed identified banding similarities (Table 4, Figure 2S). However, in the comparisons between

D. cardini and *D. neomorpha*, and *D. polymorpha* and *D. neomorpha* we could not identify the band/interband similarities in the whole extension of the chromosome. Although the comparisons among the IV/IIIL chromosomes showed whole chromosomes without similarities, the V/IIIR chromosomes showed fewer similarities in the overall analysis. In this chromosome it was observed the second new heterozygous inversion for *D. parthenogenetica* in our analysis, and a new inversion for *D. neocardini*, the 3RA present in the continental populations of its southern distribution. *Drosophila polymorpha* has five inversions already detected for chromosome IIIR: 3RA (De Toni *et al*., 2001a), 3RB, 3RC, 3RD and 3RE (De Toni, 2002).

The VI/IV chromosome (Figure 5F) corresponds to the dot chromosome of all the *cardini* group species and it was divided in one section with two subsections (100a and 100b). This chromosome is generally attached to the X chromosome; however it differs in shape and number of bands among the species. The attempt to establish the band number for this chromosome is not always precise; consequently we do not perform the pairwise comparison among them.

Discussion

Regarding the *cardini* group, our study is the first to presents the species photomaps and the chromosomal analysis aiming to generate a formal comparative evolutionary study, for this group of species. For this purpose, the construction of precise photomaps that allow the exact localization of nucleotide sequences is extremely important, as recently performed by Schaeffer *et al.* (2008) with the sequences of the *Drosophila* 12 Genomes Project (2007). Some authors have successfully performed studies

on the chromosomal evolution of some group of species in the *Drosophila* genus (Sturtevant and Dobzhansky, 1936; Clayton *et al.,* 1972; Lemeunier and Ashburner, 1976; Throckmorton 1982; Wasserman 1982, 1992; Kaneshiro *et al.*, 1995; Rohde *et al.,* 2006). These works suggest that the use of band/interband pattern comparisons is valuable in the evolutionary assays, although several sources of error may exist in this kind of analysis (Wasserman, 1992).

For the *Drosophila* genus the genome restructuration seems to be the rule of the evolution; where the gene order inversions are mainly generated by chromosomal inversions; although, in general, the relationships of sinteny of the genes remain conserved on the same chromosome arm among species (Bhutkar *et al.*, 2008). *Drosophila cardini* is the basal species of the *cardini* group (Heed and Russel, 1971; Brisson *et al.,* 2006), with a diploid chromosome number 2n=12 while all the other species of the group have 2n=8. Considering this, along the evolutionary history of the *cardini* group species seems to have undergone to a series of chromosomal restructuration, including chromosome fusions, pericentric and paracentric inversions. Analyzing the heterozygous chromosomal inversions in the six species of this group, it seems that the common situation is the occurrence of larger chromosomal inversions than small ones. Another source for the scramble pattern of the bands/interbands sequences might be the introgression mechanism that was identified to have happen among the *cardini* group species (Heed, 1962), providing a higher mixture of the gene sequences among these species. Therefore, the difficult to find banding similarities in the middle of the chromosomes for these species is possible due to these chromosomal reorganizing mechanisms acting together. Differently of some species, as *Drosophila willistoni* (Rohde, 2000; Schaeffer *et al*., 2008) that have several chromosomal regions serving as hallmark, the *cardini* group polytene

chromosomes have very little chromosomal regions of easy recognition leading the whole analyzing process very laborious.

Considering the chromosomal studies in the *cardini* group species, it is possible that the inversions described by Rohde and Valente (1996), by De Toni *et al*. (2001a), by De Toni (2002) and by us in this study was already observed by Heed and Russel (1971). However, the absence of the break points characterization draws and photomicrographs in this last work becomes any comparison impracticable. Although the *cardini* species present a high number of polymorphic heterozygous inversions detected by Heed and Russel (1971), the populations analyzed in this study showed a low number of inversions. These low records of chromosomal variants in our samples may be explained as a consequence of founder effect or even by the high inbreeding of the cultures in the laboratory environment. Yet, another explanation lies in the fact that some of these samples may be descendants of marginal populations.

According to the review by Singh (2008), marginal populations show a lower degree of chromosomal polymorphism than those in the centre of a given geographical distribution. Considering that marginal populations are under a strong pressure selection and that chromosomal recombination is an adaptative factor, the low level of chromosomal inversion polymorphism in these populations gives a high potential for free chromosomal recombination from which adaptative novelties can be generated (Singh, 2008). Considering this, the geographical distribution of the *cardini* subgroup species are designated to an area that stretches through the south of USA to the south of Brazil and north of Argentina. The *cardini* group species putative center of dispersion is indicated as been in the Central America (Heed and Russel, 1971). Therefore, the populations here analyzed, mainly from south Brazil, are potentially marginal populations form the six

species. Indeed, *D. neomorpha* and *D. parthenogenetica* recently had their first identification on south rainforests of Brazil (De Toni *et al.*, 2005).

For the species populations distributed in south Brazil, *D. polymorpha* is without doubt the most polymorphic species of *cardini* subgroup. When more populations from different distribution could be analyzed for heterozygous inversions, however, the probabilities of to found different rearrangements will certainly increase. Perhaps, the analysis of population of these six species from different places from Brazil would improve our data aiming the evolutionary studies of the *cardini* group based on the polytene chromosomes.

Acknowledgements - We are grateful to all the fly collectors for their assistance, which made this work possible. Thanks are also due to Dr. Hope Hollocher for the *D. parthenogenetica* strains, to Dr. Claudia Rohde for the initial help in this work, and to Daniel Dombrosky and Eduardo Faria for help with computational image programs. This work was supported by Brazilian grants from CAPES, CNPq, and FAPERGS research funding and fellowships.

References

Ashburner M (1989) *Drosophila*: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cambridge, England

Bhutkar A, Schaeffer SW, Russo SM, Xu M, Smith TF, Gelbart WM (2008) Chromosomal rearrangement inferred from comparisons of 12 *Drosophila* genomes. Genetics 179:1657-1680

Brisson JA, De Toni DC, Duncan I and Templeton AR (2005) Abdominal pigmentation variation in *Drosophila polymorpha*: geographic in the trait, and underlying phylogeography. Evolution 59(5):104-1059

Brisson JA, Wilder J and Hollocher H (2006) Phylogenetic analysis of the *cardini* group of *Drosophila* with respect to changes in pigmentation. Evolution 60:1228-1241.

Chaves NB and Tidon R (2008) Biogeographical aspects of drosophilids (Diptera, Drosophilidae) of the Brazilian Savana. Revta. Bras Ent. 52(3):340-348.

Clayton FE, Carson HL and Sato JE (1972) Polytene chromosomal relationships in Hawaiian species of *Drosophila*. Univ. Texas Pulbs. 7(7213):163-177.

Da Cunha AB, Brncic D and Salzano FM (1953) A comparative study of chromosomal polymorphism in certain South American species of *Drosophila*. Heredity 2(7): 193–202.

De Toni, DC (2002) Estudo da variabilidade genética e ecológica de comunidades de *Drosophila* em regiões de Mata Atlântica de ilhas do continente de Santa Catarina. PhD Thesis, Universidade Federal do Rio Grande do Sul, Brazil. 152pp.

De Toni DC, Araujo CB, Morales NB and Valente VLS (2001a) Reference photomap of the salivary gland polytene chromosomes of *Drosophila neocardini* (Streisinger, 1946). Drosoph. Inf. Serv. 84:88-91.

De Toni DC, Heredia FO and Valente VLS (2001b) Chromosomal variability of *Drosophila polymorpha* populations from Atlantic Forest remnants of continental and insular environments in the State of Santa Catarina, Brazil. Caryologia. 54:329-337.

De Toni DC, Brisson JA, Hofmann PRP, Martins M and Hollocher H (2005) First record of *Drosophila parthenogenetica* and *D. neomorpha*, *cardini* group, Heed, 1962 (Diptera, Drosophilidae), in Brazil. Drosoph. Inf. Serv. 88:33-38

De Toni DC, Loureiro MA, Hofmann PRP and Valente VLS (2006) Reference photomap of the salivary gland polytene chromosomes of *Drosophila neomorpha* (Heed and Wheeler, 1957). Drosoph. Inf. Serv. 89:73-77.

De Toni DC, Gottschalk MS, Cordeiro J, Hofmann PRP and Valente VLS (2007) Study of the Drosophilidae (Diptera) communities on Atlantic Forest islands of Santa Catarina State, Brazil. Neotropical Ent. 36(3):356-375

Dobzhansky T and Pavan C (1943) Studies on Brazilian species of *Drosophila*. Bolm. Fac. Filos. Cienc. S. Paulo 36(4):7-72.

Dobzhansky T and Sturtevant AH (1938) Inversions in the chromosomes of *Drosophila pseudoobscura*. Genetics 23: 28-64.

Drosophila 12 Genomes Consortium, Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, Iyer VN *et al*. (2007) Evolution of genes and genomes on the Drosophila phylogeny. Nature 450(7167):203-18.

Grimaldi DA (1990) A phylogenetic, revised classification of the genera in the Drosophilidae (Diptera). Bull. Am. Mus. Natl. Hist. 197:1–139.

Heed WB (1962) Genetic characteristics of island populations. Univ. Texas Publ. Stud. Genet. 6205:173-206.

Heed WB (1963) Density and distribution of *Drosophila polymorpha* and its color alleles in South America. Evolution 17: 502–518.

Heed WB and Krishnamurthy NB (1959) Genetic studies on the *cardini* group of *Drosophila* in the West Indies. Univ. Texas Publs. 5914:155-179.

Heed WB and Russell JS (1971) Phylogeny and population structure in island and continental species of the *cardini* group of *Drosophila* studied by inversion analysis. Univ. Texas Publs. Stud. Genet. 6(7103):91-130.

Heed WB and Wheeler MR (1957) Thirteen new species in the genus *Drosophila* from the neotropical region. Univ. Texas Publ. Stud. Genet. 5721:17-38.

Hollocher H, Hatcher JL and Dyreson EG (2000a) Evolution of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. Evolution 54 (6): 2046-2056.

Hollocher H, Hatcher JL and Dyreson EG (2000b) Genetic and developmental analysis of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. Evolution 54 (6): 2057-2071.

Kaneshiro KY, Gillespie RL and Carson HL (1995) Chromosomes and male genitalia of Hawaiian Drosophila: tools for interpreting phylogeny and geography, pp. 57–71 In: Wagner WL and Funk AK (eds) Hawaiian Biogeography. Evolution on a hot spot Archipelago. Smithsonian Institution Press, Washington

Lemeunier F, and Ashburner MA (1976) Relationships within the *melanogaster* species subgroup of the genus *Drosophila* (Sophophora). II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences. Proc. R. Soc. Lond. B. Biol. Sci. 193(1112):275-294.

Marques EK, Napp M, Winge W and Cordeiro A (1966) A corn meal, soybean flour, wheat germ medium for *Drosophila*. Drosph. Inf. Serv. 41:147-1966.

Martinez MN and Cordeiro AR (1970) Modifiers of color pattern genes in *Drosophila polymorpha*. Genetics 64: 573–587.

Napp M and Cordeiro AR (1981) Interspecific relationship in the *cardini* group of *Drosophila* studied by electrophoresis. Revta Bras Genet 4: 537–547.

Robe LJ, Valente VLS, Budnik M and Loreto ELS (2005) Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: A nuclear versus mitochondrial gene approach. Mol Phylogenet Evol. 3:623-640.

Rohde, C (2000) Polimorfismo cromossômico e elementos de transponíveis AM *Drosophila willistoni*. PhD Thesis. Universidade Federal do Rio Grande do Sul, Brazil. 238pp.

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

Rohde C and Valente VLS (1996) Cytological maps and chromosomal polymorphism of *Drosophila polymorpha* and *Drosophila cardinoides*. Braz. J. Genet. 19:27-32.

Santos-Colares M, Goñi B and Valente VLS (2002) An improved technique for mitotic and meiotic chromosomes of Neotropical species of *Drosophila*. Drosoph. Inf. Serv. 85:133-136.

Schaeffer SW, Bhutkar A, McAllister BF, Matsuda M, Matzkin LM, O'Grady PM, Rohde C, Valente VL, Aguadé M, Anderson WW et al. (2008) Polytene chromosomal maps of 11 Drosophila species: the order of genomic scaffolds inferred from genetic and physical maps. Genetics 179:1601-1655.

Singh BN (2008) Chromosome inversions and linkage disequilibrium in *Drosophila*. Curr. Sci. 94(4):459-464.

Stalker HD (1953) Taxonomy and hybridization in the *cardini* group of *Drosophila*. Ann. Ent. Soc. Am. 46:343-358.

Streisinger G (1946) The *cardini* species group of the genus *Drosophila*. J. N. Y. Ent. Soc. 54:105-113.

Sturtevant AH (1942) The classification of the genus *Drosophila*, with descriptions of nine new species. Univ. Texas. Publ. 4213:5-51.

Sturtevant AH and Dobzhansky T (1936) Inversions in the third chromosome of wild races of *Drosophila pseudobscura*, and their use in the study of the history of the species. Proc. Natl. Acad. Sci. USA 22:448-450.

Sumner AT (2003) Chromosomes: organization and function. Blackwell Publishing, London, United Kingdom. 287pp.

Throckmorton LH (1982) The virilis species group, pp. 227– 296 In: Ashburner, M., Carson, H.L., Thompson Jr, J.N. (eds) The Genetics and Biology of *Drosophila* Academic Press, London.

Throckmorton LH (1975) The phylogeny, ecology and geography of *Drosophila*. In: King, R.C. (ed.), Handbook of Genetics. Plenum, New York, pp. 421–469.

Vilela CR, da Silva AFG and Sene FM (2002) Preliminary data on the geographical distribution of *Drosophila* species within morphoclimatic domains of Brazil. III. The *cardini* group. Revta. Bras. Ent. 2(46):139-148.

Wasserman M (1982) Evolution of the *repleta* group. In: Ashburner, M., Carson, H. L. Thompson, J. N. (eds) The Genetics and Biology of Drosophila, Vol. 3b, Academic Press, London, pp.61-139.

Wasserman M (1992) Cytological evolution of the *Drosophila* repleta Species group. In: Powell, J. R., Krimbas, C. B. (eds) Inversion Polymorphism in Drosophila, CRC Press, Inc., Boca Raton, Florida, pp.455-541.

Supplementary Material

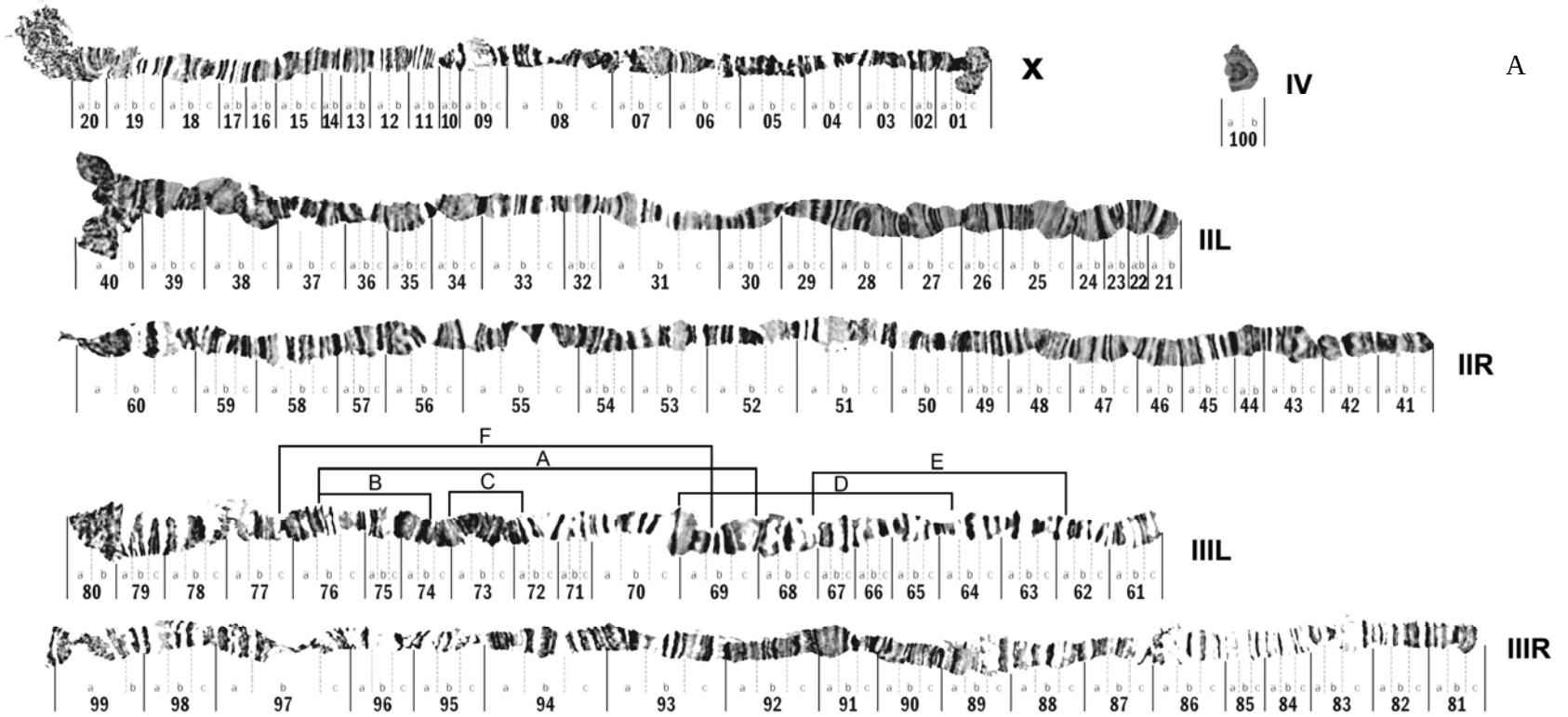


Figure 1SA. Improved reference photomaps of the polytene chromosomes of *Drosophila cardinoides* (A), *D. neocardini* (B) and *D. polymorpha* (C). The distal regions of the chromosomes (base) are in the right side of the figure. Full lines correspond to the break points of inversions described in this study and elsewhere (see text). Above the lines there is a letter that identifies the inversion in each chromosomal arm. * Indicates the new inversions found in this study.

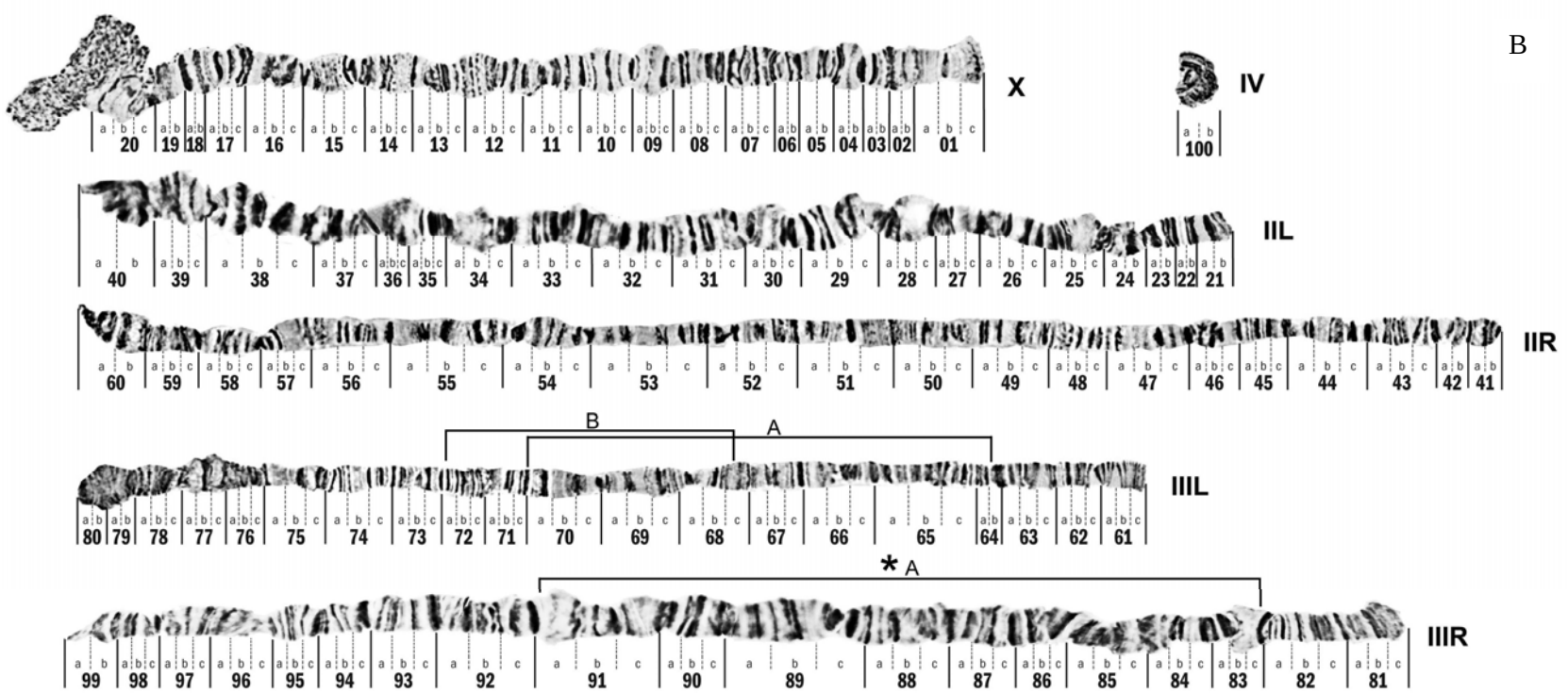


Figure 1SB: Improved reference photomap of *Drosophila neocardini.* Legend is in the previous page.

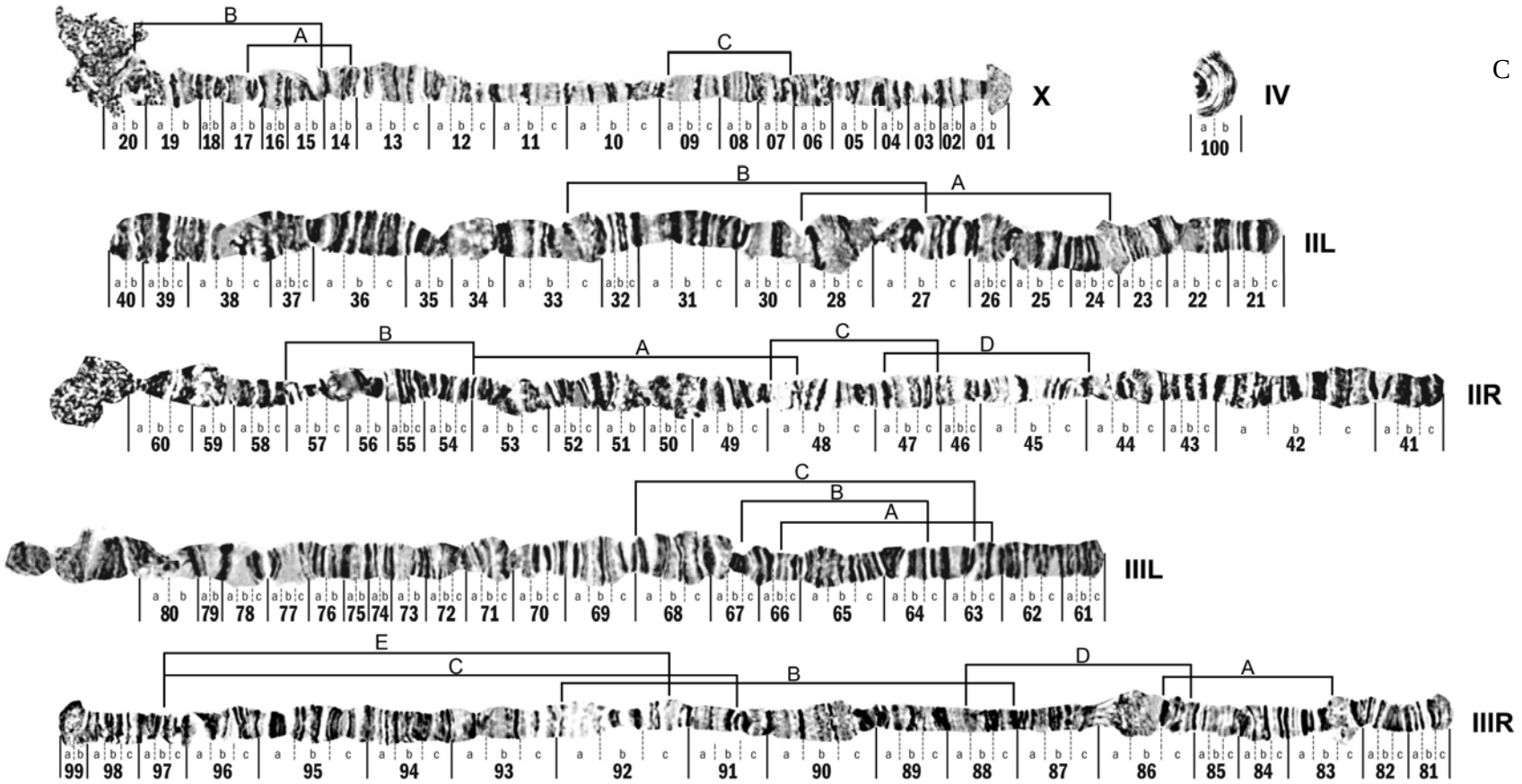


Figure 1SC: Improved reference photomap of *Drosophila polymorpha.* Legend is in the previous page.

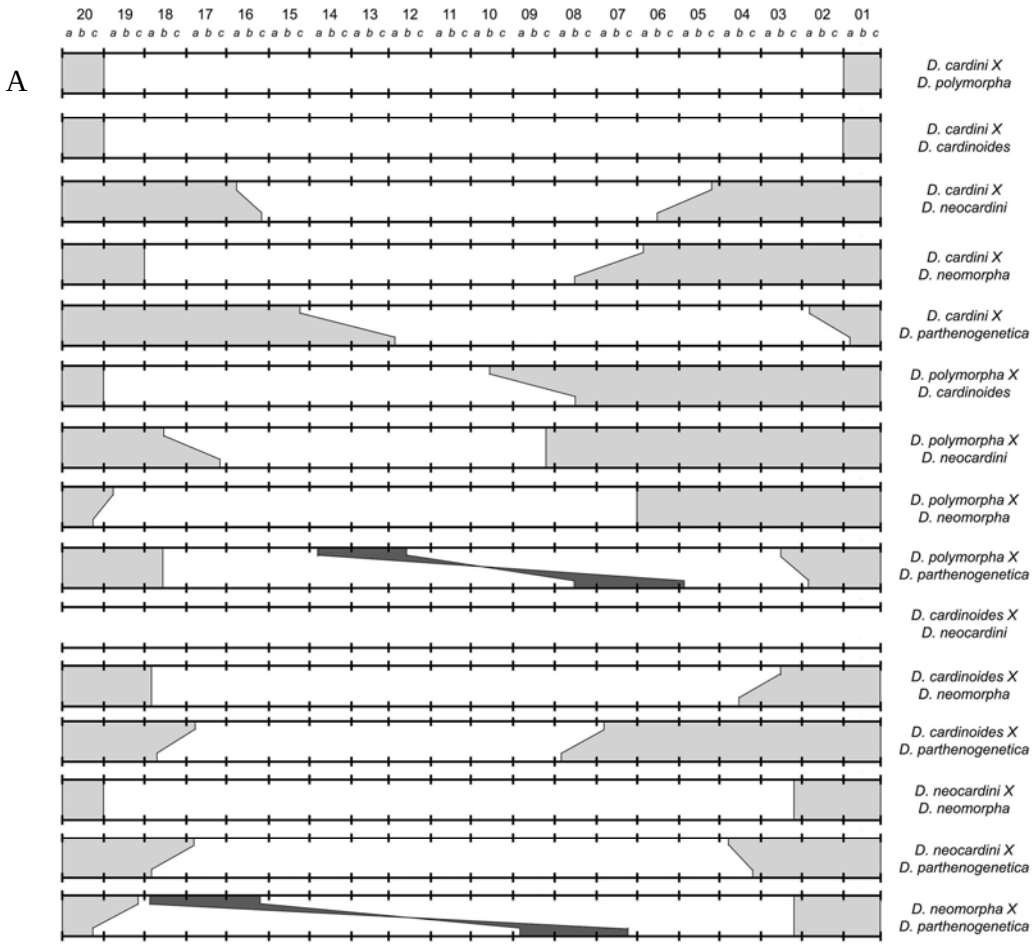


Figure 2S. Schematic representation of the analysis of 75 pairwise comparisons informed in Table 4. The lines represent the chromosomes, where in (A) it is represented the X chromosomes of all species; in (B) the chromosome II for *Drosophila cardini* and IIL chromosomal arms for the other five species; in (C) the chromosome III and IIR chromosomal arms; in (D) the chromosome IV and IIIL chromosomal arms; and in (E) the chromosome V and IIIR chromosomal arms. The distal regions of the chromosomes (base) are in the right side. Above the figures it is indicated the section and subsection for all chromosomes. The gray part of the comparisons indicates the regions with similarities. Inverted regions are filled by dark gray.

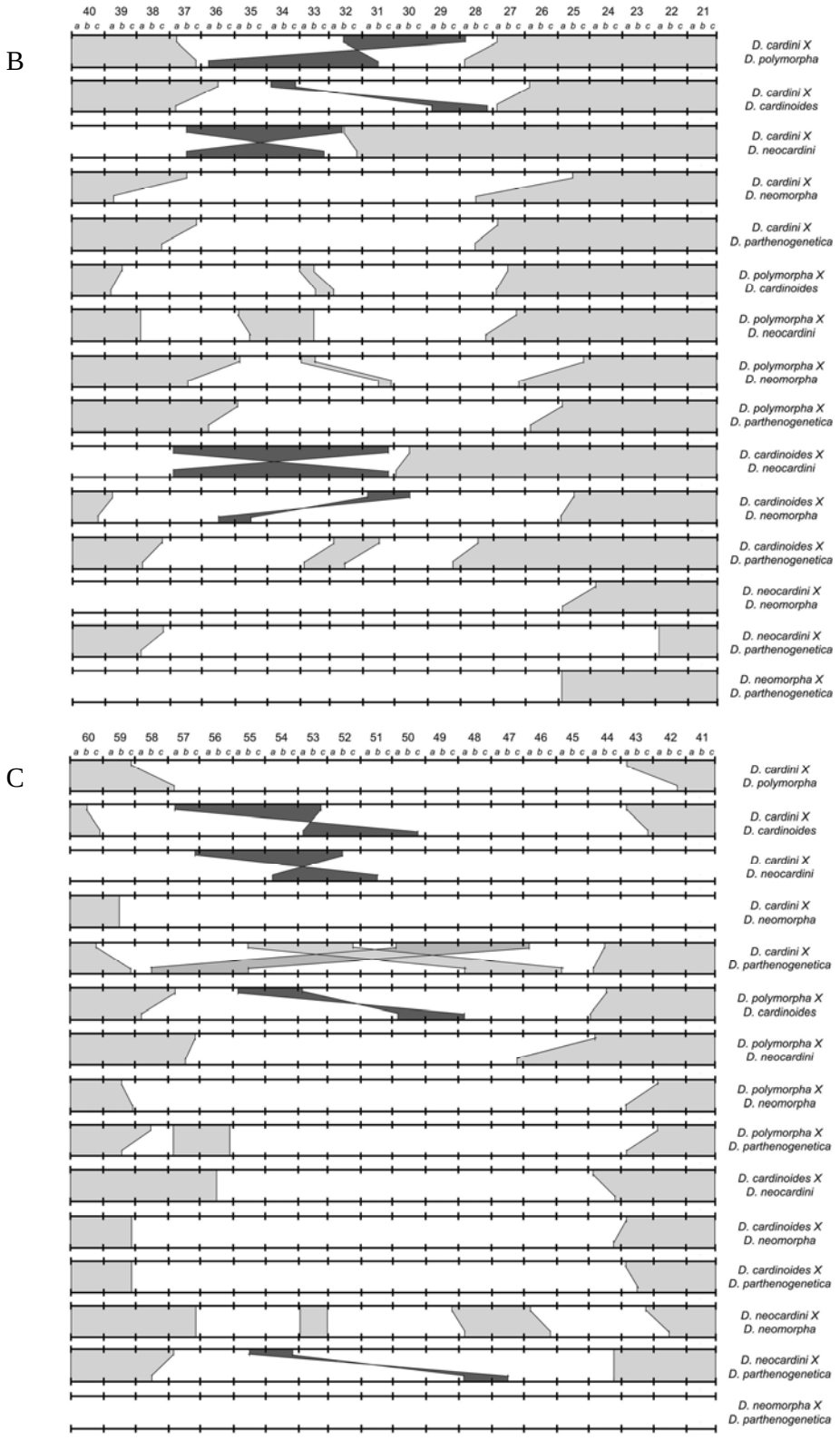


Figure 2S B and C: Schematic representation of the 75 pairwise comparisons analysis informed in Table 4. See previous page for legend.

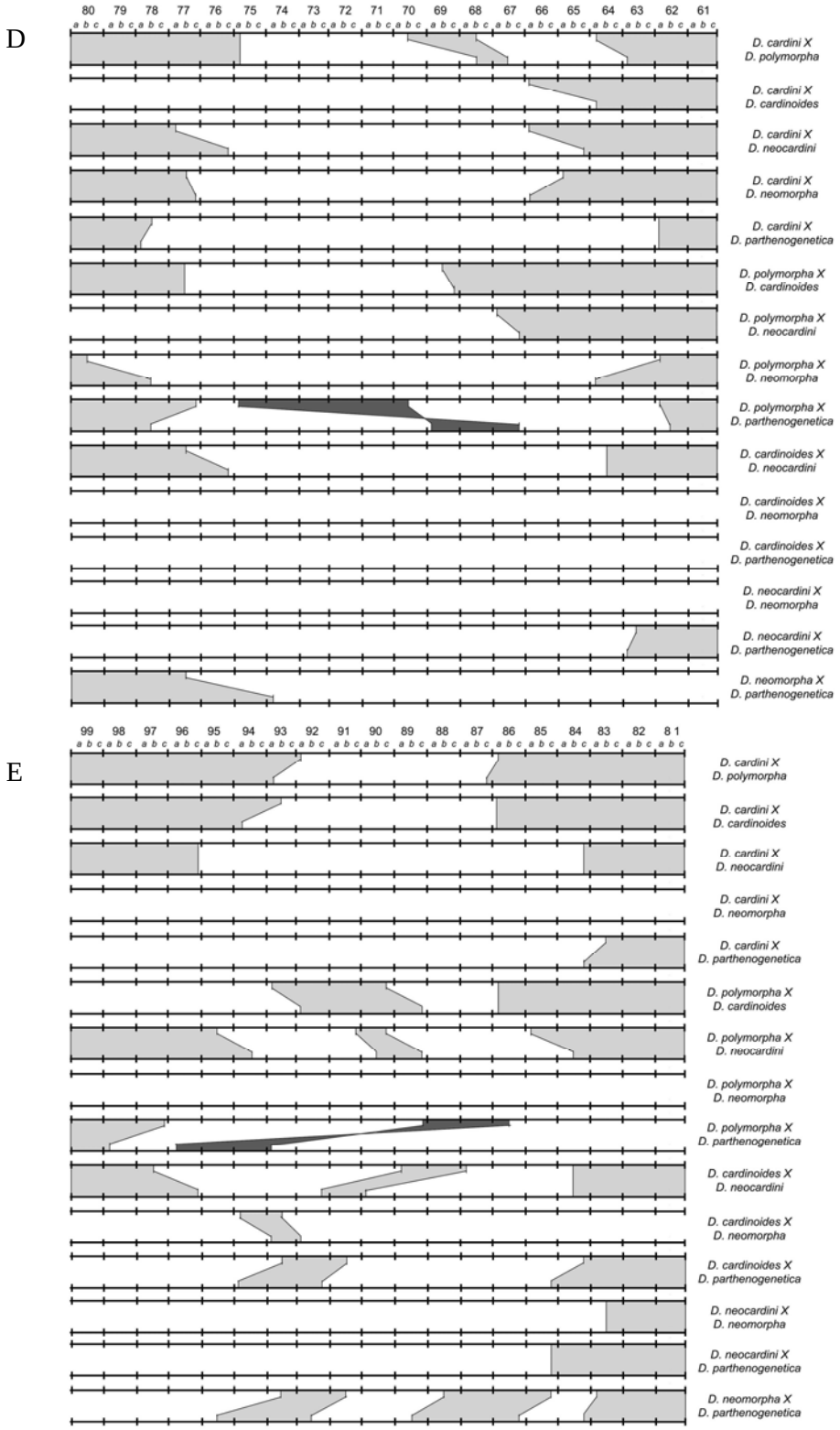


Figure 2S D and E: Schematic representation of the 75 pairwise comparisons analysis informed in Table 4. See previous page for legend.

CAPÍTULO VI

Discussão Geral

Nos anos 80 o retroelemento *micropia* foi descoberto intimamente relacionado às estruturas "plumosas" (*lampbrush*) presentes no cromossomo Y de *Drosophila hydei* (Hennig *et al.*, 1983; Huijser *et al.*, 1988). Após a sua descoberta, os trabalhos realizados até 1994 estudaram a organização molecular deste retroelemento nos genomas de *D. hydei* e *D. melanogaster* (Hennig *et al*., 1989; Lankenau *et al*., 1988, 1989, 1990, 1994) observando que ele estava presente em alto número de cópias nos genomas dessas duas espécies, além de possuir características muito relacionadas com elementos virais. Neste período o elemento *micropia* já havia sido identificado como um retroelemento com LTR. Mais recentemente os trabalhos de Almeida *et al*. (2001) e Almeida e Carareto (2004) ampliaram o registro da presença desse elemento para os genomas de espécies dos grupos *repleta* e *saltans*, revelando a ampla distribuição deste elemento no gênero *Drosophila*. Porém os dados gerados por esta Tese mostram que, apesar da distribuição do elemento *micropia* ter sido ampliada também para espécies do grupo *cardini* (**Capítulo II** e **III**), esta distribuição é descontínua tanto neste grupo quanto nos grupos *repleta* e *melanogaster*. O interessante é que utilizando uma região da sequência do clone *micropia*-dhMiF2 (entre os nucleotídeos 1777 e 2570 do total de aproximadamente 5kb) diferente da região utilizada por Almeida e Carareto (2004) (entre os nucleotídeos 2813 e 3198 da mesma sequência) não encontramos sequências similares no genoma completo de *D. mojavensis*, pertencente ao grupo *repleta*, disponível no banco de dados (Consórcio 12 Genomas de *Drosophila*, 2007), apesar das autoras anteriormente citadas terem conseguido amplificar o fragmento

esperado. A escolha do par de *primers* Mic1777 e Mic2570 levou em consideração a temperatura de anelamento (por volta de 60ºC) e principalmente o tamanho do fragmento potencialmente amplificado (por volta de 800pb). As diferenças obtidas entre os dados desta Tese e os dados daquele trabalho podem ser devidas à existência de cópias divergentes do elemento *micropia* no genoma dessas espécies, como foi sugerido por Almeida e Carareto (2004).

Ainda, em consideração à distribuição descontínua no grupo *cardini*, este fator parece refletir a própria distribuição geográfica deste grupo. As espécies do subgrupo *dunni* do grupo *cardini*, as quais não apresentam sequências similares ao elemento *micropia*, possuem distribuição restrita às ilhas do Caribe. Essas espécies insulares do grupo *cardini* são derivadas de um ancestral presente no continente, e a divergência dessas espécies aconteceu por volta de 1,6 a 2,6 milhões de anos atrás (Wilder e Hollocher, 2003). No **Capítulo III** calculamos que as cópias do retroelemento *micropia* presentes no grupo *cardini* parecem ter divergido há 3,7 milhões de anos atrás envolvendo mecanismos de HT. Considerando essas datas de divergência, pode-se supor que o evento de invasão desta cópia de *micropia* não atingiu o ancestral do subgrupo *dunni*, devido à perda de área de contato ou qualquer outro motivo. Contrariamente, se o genoma da espécie ancestral foi invadido por *micropia*, essas cópias divergiram rapidamente dificultando a detecção dessas sequências, representadas pelos fracos sinais no *Dot Blot* (Figura 1S, **Capítulo III**).

Ainda nos **Capítulos II** e **III** foi verificada uma baixa divergência entre as cópias de *micropia* presentes no grupo *cardini* e *repleta*, assim como entre as cópias presentes no grupo *melanogaster* (nomeadas de *nanopia* no **Capítulo III** de acordo com critérios estabelecidos por Wicker *et al*., 2007 já discutidos). No **Capítulo IV** foi estimado o número de cópias do elemento *micropia* para seis espécies do grupo *cardini* (que variaram

de cinco a 18 cópias) que parece estar dentro da média encontrada para espécies do grupo *repleta*, variando entre sete e 17 cópias (Lankenau *et al*., 1994; Almeida e Carareto, 2004). Algumas espécies do grupo *repleta*, incluindo *D. hydei*, apresentam um mecanismo de controle da transposição por meio da inibição de transcritos por interação RNA-RNA. No grupo *cardini*, a presença deste elemento tanto no cromossomo X quanto nos autossomos, assim como a sua distribuição aleatória (sem preferência por regiões de baixa recombinação como a heterocromatina), pode sugerir que, se este elemento está ativo nas espécies do grupo *cardini*, como parece estar em *D. hydei*, é possível que o mecanismo de regulação de transposição já identificado no grupo *repleta* também atue no genoma das espécies do grupo *cardini*.

Apesar de no **Capítulo IV** ter sido identificada a presença de *micropia* em pontos de quebra para inversões cromossômicas em três das espécies estudadas (*D. cardinoides*, *D. parthenogenetica* e *D. polymorpha*), não podemos afirmar que *micropia* esteja envolvido em mecanismos de geração de inversões, podendo corresponder a inserções secundárias à entrada de outros TEs da Classe II, estes sim promotores das quebras e rearranjos, como por exemplo, *Galileo* (Cáceres *et al*., 1999, 2001; Casals *et al.* 2003, 2005) em *D. buzzatii*. É no **Capítulo V** que é mostrado que *D. polymorpha* é a espécie do grupo *cardini* com o maior número de inversões heterozigotas analisadas, seguida por *D. cardinoides*. Essas duas espécies são as mais amplamente distribuídas e freqüentemente coletadas do grupo *cardini*. Esta observação talvez esteja conectada ao resultado obtido no **Capítulo IV**, onde essas duas espécies foram as que apresentaram maior número de inserções do elemento *micropia*. Como foi observado para *D. melanogaster* e *D. simulans* (Vieira *et al*., 1999), parece que a aquisição de novos TEs acompanha o processo de colonização de novos territórios, possivelmente propiciando maior plasticidade aos

genomas dos indivíduos dessas populações. Um estudo populacional identificando os sítios de inserção de *micropia* nos diferentes genomas poderia informar se o mesmo ocorre no grupo *cardini*; que apesar de não ter espécies cosmopolitas, apresenta espécies que ainda estão em processo de ocupação de novos ambientes, como é o caso de *D. neomorpha* e *D. parthenogenetica* (De Toni *et al*., 2005).

O registro da variabilidade de inversões em diferentes populações das seis espécies do grupo *cardini* anteriormente citadas, assim como a análise comparativa dos padrões de bandas e interbandas dos seus cromossomos politênicos (**Capítulo V**), tendem a ser refinados, permitindo compreender melhor os dados obtidos por Heed e Russel (1971). Como discutido no **Capítulo V**, o primeiro trabalho que construiu o fotomapa de espécies do grupo *cardini* foi publicado em 1996 (Rohde e Valente, 1996a), e o trabalho de Heed e Russel (1971) não apresenta os pontos de quebra das inversões analisadas, nem desenhos ou fotografias dessas variantes cromossômicas. A evolução dos cromossomos neste grupo de espécies parece ter envolvido uma grande série de rearranjos, dificultando o estabelecimento de regiões similares, invertidas ou não, principalmente na porção mediana dos cromossomos e braços cromossômicos. Com isso, a ausência da identificação das inversões obtidas por Heed e Russel (1971) prejudica a comparação com os dados obtidos no **Capítulo V**.

O estabelecimento correto e fiel de fotomapas de referência dos cromossomos politênicos para as espécies do grupo *cardini* é de extrema importância para os estudos genéticos e moleculares futuros. No **Capítulo V** além de melhorarmos os fotomapas já existentes de algumas espécies do grupo *cardini*, através do uso de recursos computação gráfica, produzimos os fotomapas cromossômicos para duas outras espécies que ocorrem no sul do Brasil. Com exceção de *D. parthenogenetica*, onde encontramos dificuldade em

obter bons cromossomos, o fotomapa aqui construído para *D. cardini* e os fotomapas remontados de *D. polymorpha, D. cardinoides* e *D. neocardini* permitem uma definição mais acurada das diferentes regiões cromossômicas e a descrição detalhada dos pontos de quebra das inversões, como consequência da divisão das seções em subseções.

Conclusões

Com os resultados obtidos e discutidos nos capítulos anteriores podemos chegar às seguintes conclusões:

- O elemento *micropia* apresenta sequências muito similares no genoma de espécies distantes evolutivamente (espécies dos grupos *cardini*, *repleta*, *melanogaster* e *willistoni*) e o padrão de evolução deste elemento no gênero *Drosophila* parece ser baseado na existência de polimorfismo ancestral, herança vertical, perdas estocásticas, mas principalmente por eventos de transmissão horizontal.

- As seis espécies do grupo *cardini* em que foram analisados os sítios de inserção do elemento *micropia*, *D. cardini*, *D. cardinoides*, *D. neocardini*, *D. neomorpha*, *D. parthenogenetica* e *D. polymorpha* apresentaram seis, 14, 12, cinco, seis, 18 cópias deste retroelemento nos braços eucromáticos dos cromossomos politênicos, respectivamente e não foram encontrados sinais de hibridação nos cromocentros de todas elas. Foram encontradas cópias localizadas em pontos de quebra de inversões em *D. cardinoides, D.*

polymorpha e *D. parthenogenetica*: duas, quatro e uma, respectivamente. Porém, o papel dessas coincidências entre sítios de inserção e quebra de inversões deverá ser testado em estudos futuros, já que elas podem ser devidas ao acaso ou representar inserções secundárias à entrada de outros TEs, especialmente os da Classe II, como *Galileo*.

- O fotomapa de *D. cardini* foi construído e os fotomapas de *D. cardinoides*, *D. neocardini* e *D. polymorpha* já existentes foram melhorados. *Drosophila polymorpha* se mostrou a espécies com o maior número de inversões heterozigotas do grupo *cardini*, e em *D. neomorpha* não foram encontradas inversões heterozigotas. A análise comparativa dos cromossomos politênicos mostrou que os cromossomos dessas espécies são extremamente reorganizados, dificultando o estabelecimento de uma relação filogenética precisa entre as espécies do grupo *cardini* baseada neste marcador. O uso de sondas de genes cuja localização cromossômica em outras espécies de *Drosophila* seja conhecida será útil para resolver este tipo de dificuldade.

Perspectivas

Tomando como base os resultados obtidos nos **Capítulos II**, **III**, **IV** e **V** desta Tese, tem-se como perspectiva de trabalho o que se segue:

- Analisar a presença de sequências relacionadas a *micropia* no genoma de outras espécies do gênero *Drosophila* verificando a potencial doadora das sequências para os grupos *cardini* e *repleta*;

- Verificar a presença de sequências relacionadas a *micropia* no genoma de espécies neotropicais com a finalidade de determinar a real distribuição deste retroelemento no gênero *Drosophila* .

- Analisar a sequência completa de *micropia* no grupo *cardini* verificando o seu padrão de controle de transposição com a finalidade de verificar se existe semelhança com o padrão observado em *D. hydei*, determinando, assim, a eficiência desta regulação no controle da transposição das cópias do elemento.

- Analisar a integridade de cópias divergentes encontradas nas espécies do grupo *cardini* e nas espécies dos demais grupos durante as análises tanto *in vivo* quanto *in silico*.

- Analisar os sítios de inserção das cópias de *micropia* nos cromossomos politênicos de diferentes populações de espécies do grupo *cardini* com o intuito de estudar a dinâmica deste retroelemento.

- Analisar a presença do transposon *Galileo* nas espécies do grupo *cardini* verificando a sua presença nas proximidades dos sítios de inserção onde *micropia* já foi encontrado, investigando a relação de *Galileo* com geração de inversão e a relação de *micropia* com uma possível invasão secundária dos pontos de quebra.

- Refinar o estudo da evolução cromossômica do grupo *cardini* por meio da técnica de hibridação *in situ* de genes nucleares de cópia única com a finalidade de monitorar de

forma mais precisa os rearranjos ocorridos durante o processo evolutivo, bem como estabelecer as relações filogenéticas das espécies do grupo.

REFERÊNCIAS BIBLIOGRÁFICAS

Agrawal A, Eastman QM, Schatz DG (1998) Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature 394:744–751.

Almeida LM, Carareto CMA (2004) Identification of two subfamilies of *micropia* transposable element in species of the *repleta* group of *Drosophila*. Genetica 121:155-164

Almeida LM, Carareto CM (2005) Multiple events of horizontal transfer of the *Minos* transposable element between *Drosophila* species. Mol Phylogenet Evol35(3):583-94.

Almeida LM, Carareto CMA (2006) Sequence heterogeneity and phylogenetic relationships between the *copia* retrotransposon in *Drosophila* species of the *repleta* and melanogaster groups. Genet Sel Evol 38:535-50

Almeida LM, Castro JP, Carareto CMA (2001) *micropia* transposable element occurrence in *Drosophila* species of the *saltans* group. Drosoph Inf Serv 84:114-117

Ashburner M (1976) Aspects of polytene chromosome structure and function. Allfrey:85-91.

Bächli G (2009) Taxodros: The database on taxonomy of Drosophilidae. Available at http://www.taxodros.unizh.ch. Acessado em agosto de 2009

Bartolomé C, Bello X, Maside X (2009) Widespread evidence for horizontal trnafer of transposable elements across *Drosophila* genomes. Genome Biol. 10(2):R22

Bicudo HEMC (1973) Chromosomal polymorphism in *saltans* group of *Drosophila*. Genetica 44:520-552.

Biémont C, Vieira C (2005) What transposable elements tell us about genome organization and evolution: the case of *Drosophila*. Cytogenet Genome Res 110:25-34.

Biémont C, Vieira C (2006) Genetics: junk DNA as an evolutionary force. Nature 443(7111):521-524.

Biémont C, Vieira C, Hoogland C, Cizeron G, Loevenbruck C, Arnault C, Carante JP (1997) Maintenance of transposable element copy number in natural populations of *Drosophila melanogaster* and *D. simulans*. Genetica 100:161-166.

Bowen NJ, Jordan IK (2002) Transposable elements and the evolution of eukaryotic complexity. Curr Issues Mol Biol 4(3):65-76.

Brisson JA, De Toni DC, Duncan I, Templeton AR (2005) Abdominal pigmentation variation in *Drosophila polymorpha*: geographic variation in the trait, and underlying phylogeography. Evolution 59(5):1046-1059.

Brisson JA, Wilder J, Hollocher H (2006) Phylogenetic analysis of the *cardini* group of *Drosophila* with respect to changes in pigmentation. Evolution 60:1228-1241

Brunet F, Godin F, Bazin C, Capy P (1999) Phylogenetic analysis of *Mos1*-like transposable elements in the Drosophilidae. J Mol Evol 49: 760–768.

Brunet F, Godin F, David JR, Capy P (1994) The *mariner* transposable element in the Drosophilidae family. Heredity 73: 377–385.

Bhutkar A, Schaeffer SW, Russo SM, Xu M, Smith TF, Gelbart WM (2008) Chromosomal rearrangement inferred from comparisons of 12 *Drosophila* genomes. Genetics 179(3):1657-80.

Cáceres M, Puig M, Ruiz A (2001) Molecular characterization of two natural hotspots in the Drosophila buzzatii genome induced by transposon insertions. Genome Res 11(8):1353-64.

Cáceres M, Ranz JM, Barbadilla A, Long M, Ruiz A (1999) Generation of a widespread *Drosophila* inversion by a transposable element. Science 285:415-418.

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

Capy P, Bazin C, Higuet D, Langin T (1998) Dynamics and evolution of transposable elements. Texas: Landes Bioscience 197p

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

Carson HL, Kaneshiro KY (1976) *Drosophila* of Hawaii: Systematics and ecological genetics. Ann Rev Ecol Syst7: 311-345.

Casals F, González J, Ruiz A (2006) Abundance and chromosomal distribution of six *Drosophila buzzatii* transposons: *BuT1*, *BuT2*, *BuT3*, *BuT4*, *BuT5*, and *BuT6*. Chromosoma 115:403-12.

Charles M, Tang H, Belcram H, Paterson A, Gornicki P, Chalhoub B (2009) Sixty million years in evolution of soft grain trait in grasses: emergence of the softness locus in the common ancestor of Pooideae and Ehrhartoideae, after their divergence from Panicoideae. Mol Biol Evol 26(7):1651-1661.

Chaves NB, Tidon R (2005) Drosophilidae of the Brazilian Savanna, the forgotten ecosystem. Drosoph Inf Serv 88: 25-27

Clark JB, Kidwell MG (1997) A phylogenetic perspective on *P* transposable element evolution in *Drosophila*. Proc Natl Acad Sci USA 94:11428-11433.

Craig N, Craigie R, Gellert M, Lambowitz A (2002) Mobile DNA II. Washington, DC: ASM Press.

Da Cunha AB (1949) Genetic analysis of the polymorphism of color pattern in *Drosophila polymorpha*. Evolution 3(3):239-251.

Da Cunha AB, Brncic D, Salzano FM (1953) A comparative study of chromosomal polymorphism in certain South American species of *Drosophila*. Heredity 2 (7): 193–202.

Da Cunha AB, Burla H, Dobzhansky T (1950) Adaptive chromosomal polymorphism in *Drosophila willistoni*. Evolution 4: 212-235.

Da Cunha AB, Dobzhansky T, Pavlovsky O, Spassky B (1959) Genetics of natural populations. XXVIII. Supplementary data on the chromosomal polymorphism of *Drosophila willistoni* in its relation to the enviroment. Evolution 13: 389-404.

Daniels SB, Chovnic A, Boussy IS (1990) Distribution of *hobo* transposable elements in the genus *Drosophila*. Mol Biol Evol 7:589–606.

De Toni DC, Araujo CB, Morales NB, Valente VLS (2001b) Reference photomap of the salivary gland polytene chromosomes of *Drosophila neocardini* (Streisinger, 1946). Drosoph Inf Serv 84: 88-91.

De Toni DC, Brisson JA, Hofmann PRP, Martins M, Hollocher H (2005) Frist record of *Drosophila parthenogenetica* and *D. neomorpha*, *cardini* group, Heed, 1962 (Drosophila, Drosophilidae), in Brazil. Drosoph Inf Serv 88:33-38.

De Toni DC, Herédia FO, Valente VLS (2001a) Chromosomal variability of *Drosophila polymorpha* populations from Atlantic Forest remnants of continental and insular environments in the State of Santa Catarina, Brazil. Caryologia 4 (54): 329–337.

De Toni DC, Hofmann PRP (1995) Preliminary taxonomic survey of the genus *Drosophila* (Diptera, Drosophilidae) at Morro da Lagoa da Conceição, Santa Catarina Island, Brazil. Revta Bras Biol 3 (55): 347-350.

De Toni DC, Loureiro MA, Hofmann PRP, Valente VLS (2006) Reference photomap of the salivary gland polytene chromosomes of *Drosophila neomorpha* (Streisinger, 1946) Drosoph Inf Serv 89: 73-77.

Dimitri P, Junakovic N (1999) Revising the selfish DNA hypothesis: new evidence on accumulation of transposable elements in heterochromatin. Trends Genet 15(4):123-124

Diniz NM (1998) Filogenia cromossômica de espécies do subgrupo *fasciola* do grupo *repleta* do gênero *Drosophila*. Tese de Doutorado, Universidade de São Paulo. Ribeirão Preto/SP.

Dobzhansky T (1937) Genetics and the origin of species. New York: Columbia University Press. 364pp.

Dobzhansky T, Pavan C, (1943) Studies on Brazilian species of *Drosophila*. Bolm Fac Filos Cienc S Paulo 36 (4): 7-72.

Feschotte C, Jiang N, Wessler SR (2002) Plant transposable elements: where genetics meets genomics. Nat Rev Genet 3(5):329-341

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

Frydenberg O (1956) Two new species from Peru (Diptera: Drosophilidae). Drosoph Inf Serv 30: 115.

González J, Petrov DA (2009) The adaptive role of transposable elements in the *Drosophila* genome. Gene 23 doi:10.1016/j.gene.2009.06.008

González J, Lenkov K, Lipatov M, Macpherson JM, Petrov DA (20080 High rate of recent transposable element-induced adaptation in *Drosophila melanogaster*. PLoS Biol 6(10):e251

González J, Macpherson JM, Petrov DA (2009) A recent adaptive transposable element insertion near highly conserved developmental loci in Drosophila melanogaster. Mol Biol Evol 20

Gottschalk MS, De Toni DC, Valente VLS, Hofmann PRP (2007) Changes in Brazilian Drosophilidae (Diptera) assemblages across an urbanisation gradient. Neotropical Ent. 36(6)**:** 848-862.

Grimaldi DA (1990) A phylogenetic, revised classification of the genera in the Drosophilidae (Diptera). Bull Am Mus Natl Hist 197:1-139.

Guerreiro MPG, Chávez-Sandoval BE, Balanyà J, Serra L, Fontdevila A (2008) Distribution of the transposable elements *bilbo* and *gypsy* in original and colonizing populations of *Drosophila subobscura.* BMC Evol Biol 8:234.

Heed WB (1962) Genetic characteristics of island populations. Univ Texas Publ Stud Genet 6205:173-206

Heed WB (1963) Density and distribution of *Drosophila polymorpha* and its color alleles in South America. Evolution 17: 502–518.

Heed WB, Blake PR (1963) A new color allele at the locus of *Drosophila polymorpha* from northern south America. Genetics 2 (48): 217-234.

Heed WB, Krishnamurthy NB (1959) Genetic studies on the *cardini* group of *Drosophila* in the West Indies. Univ Texas Publ Stud Genet 5914: 155–178.

Heed WB, Russel JS (1971) Phylogeny and population structure in island and continental species of the *cardini* group of *Drosophila* studied by inversion analysis. Univ. Texas Publs 7103:91-130

Heed WB, Wheeler MR (1957) Thirteen new species in the genus *Drosophila* from the neotropical region. Univ Texas Publ Stud Genet 5721: 17-38.

Hennig W, Huijser P, Vogt P, Jackle H, Edstrom JE (1983) Molecular cloning of microdissected lampbrush loop DNA sequences of *Drosophila hydei*. EMBO J 2 (10): 1746-1983.

Herédia FO, Loreto ELS, Valente VLS (2004) Complex evolution of *gypsy* in drosophilid species. Mol Biol Evol 21:1-12

Hollocher H (1996) Island hopping in *Drosophila*: patterns and processes. Philos Trans R Soc Lond B Biol Sci 351(1341):735-743

Hollocher H, Hatcher JL, Dyreson EG (2000) Evolution of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. Evolution 54 (6): 2046-2056.

Hoogland C, Biémont C (1996) chromosomal distribution of transposable elements in *Drosophila melanogaster*: test of the ectopic recombination model for maintenance of insertion site number. Genetics 144:197-204.

Huijser P, Kirchhoff C, Lankenau D-H, Hennig W (1988) Retrotransposon-like sequences are expressed in Y chromosomal lampbrush loops of *Drosophila hydei*. J Mol Biol. 203:689-697

Hurst GD, Werren JH (2001) The role of selfish genetic elements in eukaryotic evolution. Nat Rev Genet 2(8):597-606

International Human Genome Sequencing Consortium, Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, *et al*. (2001) Initial sequencing and analysis of the human genome. Nature 409(6822):860-921.

Jordan IK, Maayunina LV, McDonald JF (1999) Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. Proc Natl Acad Sci USA 22:12621-12625

Kapitonov VV, Jurka J (2003) Molecular paleontology of transposable elements in the *Drosophila melanogaster* genome. Proc Natl Acad Sci USA 100(11):6569-6574

Kastritsis CD (1966) Cytological studies on some species of the *tripunctata* group. Univ Texas Publ Stud Genet 6615: 413-474.

Kastritsis CD (1969) The chromosomes of some species of the *guarani* group of *Drosophila*. J. Hered 60: 51-57.

Kidwell MG, Lisch D (1997) Transposable elements as sources of variation in animals and plants. Proc Natl Acad Sci USA 94(15):7704-7711

Kidwell MG, Lisch DR (2001) Transposable elements, parasitic DNA, and genome evolution. Evolution 55:1-24

Krimbas CD, Powell JR (1992) *Drosophila* inversion polymorphism. Boca Raton, Florida: CRC Press. 560pp.

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

Ladevèze V, Aulard S, Chaminade N, Périquet G, Lemeunier F (1998) *Hobo* transposons causing chromosomal breakpoints. Proc. Biol. Sci. 265:1157-1159.

Lankenau D-H (1990) molecular structure and evolution of a retrotransposon family in *Drosophila*. PhD Thesis, University of Nijmegen, The Netherlands. 167pp.

Lankenau D-H (1993) The retrotransposon family *micropia* in *Drosophila* species. In: McDonald J (ed) Transposable Elements and Evolution. Amsterdam: Kluwer Publishers. pp 232-241

Lankenau D-H, Huijser P, Jansen E, Miedema K, Hennig W (1988) *micropia*: a retrotransposon of *Drosophila* combining structural features of DNA viruses, retroviruses and non-viral transposable elements. J Mol Biol. 2:233-246

Lankenau D-H, Huijser P, Jansen E, Miedema K, Hennig W (1990) DNA sequence comparison of *micropia* transposable elements from *Drosophila hydei* and *Drosophila melanogaster*. Chromosoma 99:111-117

Lankenau S, Corces GV, Lankenau D-H (1994) The *Drosophila micropia* retrotransposon encodes a testis-specific antisense RNA complementary to reverse transcriptase. Mol Biol Evol. 17:1542-1557

Le Rouzic A, Capy P (2005) The first steps of transposable elements invasion: parasitic strategy vs. genetic drift. Genetics 169(2):1033-1043

Lemeunier F, Ashburner M (1976) Relationships within the *melanogaster* species group of the genus *Drosophila* (*Sophophora*). II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences. Proc r Soc Lond 193: 275-294.

Lemeunier F, Ashburner M (1984) Relationships within the *melanogaster* species group of the genus *Drosophila* (*Sophophora*). IV. The chromosome of two new species. Chromosoma 89: 343-351.

Lemeunier F, David JR, Tsacas L, Ashburner M (1986) The *melanogaster* species group. In: Ashburner M, Carson HL, Thompson Jr JN (eds) The genetics and biology of *Drosophila*. New York: New York Academic Press. 147-256.

Lewis SM, Wu GE (1997) The origins of V(D)J recombination. Cell 88(2):159-62.

Linheiro RS, Bergman CM. (2008) Testing the palindromic target site model for DNA transposon insertion using the *Drosophila melanogaster P*-element. Nucleic Acids Res 36(19):6199-6208

Lisch D (2008) A new SPIN on horizontal transfer. Proc Natl Acad Sci USA 105(44):16827-16828

Loreto EL, Carareto CM, Capy P (2008) Revisiting horizontal transfer of transposable elements in *Drosophila*. Heredity 100:545-554

Loreto EL, Valente VL, Zaha A, Silva JC, Kidwell MG (2001) *Drosophila mediopunctata P* elements: a new example of horizontal transfer. J Hered 92(5):375-831

Lozovskaya ER, Hartl DL, Petrov DA (1995) Genomic regulation of transposable elements in *Drosophila*. .Curr Opin Genet Dev 5(6):768-773

Ludwig A, Loreto ELS (2007) Evolutionary pattern of the *gtwin* retrotransposon in the *Drosophila melanogaster* subgroup. Genetica 130:161–168.

Ludwig A, Valente VLS, Loreto ELS (2008) Multiple invasions of *errantivirus* in the genus *Drosophila*. Ins Mol Biol 17:113-124

Lyttle TW, Haymer DS (1992) The role of the transposable element *hobo* in the origin of the endemic inversions in wild populations of *Drosophila melanogaster*. Genetical Res 86:113-126.

Machado MX, De Toni DC, Hofmann PRP (2001) Abdominal pigmentation polymorphism of *Drosophila polymorpha* (Dobzhansky and Pavan, 1943) collected on Ilha de Santa Catarina and neighboring islands. Biotemas 14 (1): 87-107.

Malogolowkin C (1953) Sobre a genitália dos drosofilídeos. IV. A genitália masculina no subgênero *Drosophila* (Diptera: Drosophilidae). Revta Bras Biol 13: 245-264.

Martinez MN, Cordeiro AR (1970) Modifiers of color pattern genes in *Drosophila polymorpha*. Genetics 64: 573–587.

Miller WJ, Capy P (2006) Applying mobile genetic elements for genome analysis and evolution. Mol Biotechnol 33:161-174.

Mizrokhi LJ, Mazo AM (1990) Evidence for horizontal transmission of the mobile element *jockey* between distant *Drosophila* species. Proc Natl Acad Sci USA 87(23):9216-9220

Morales-Hojas R, Vieira CP, Vieira J (2006) The evolutionary history of the transposable element *Penelope* in the *Drosophila virilis* group of species. J Mol Evol 63(2):262-273

Napp M, Cordeiro AR (1981) Interspecific relationship in the *cardini* group of *Drosophila* studied by electrophoresis. Revta Bras Genet 4: 537–547.

Pardue ML, Rashkova S, Casacuberta E, DeBaryshe PG, George JA, Traverse KL (2005) Two retrotransposons maintain telomeres in *Drosophila*. Chromosome Res (5):443-453

Piednoël M, Bonnivard E (2009) *DIRS1*-like retrotransposons are widely distributed among Decapoda and are particularly present in hydrothermal vent organisms. BMC Evol Biol 9:86 doi:10.1186/1471-2148-9-86

Powell JR, DeSalle R (1995) *Drosophila* molecular phylogenies and their uses. In: Hecht MK et al. (eds) Evolutionary Biology. Plenun, New York, pp. 87-139.

Powell JR, Sezzi E, Moriyama EN, Gleason JM, Caccone A (2003) Analysis of a shift in codon usage in *Drosophila*. J Mol Evol 57:214-225

Puig M, Cáceres M, Ruiz A (2004) Silencing of a gene adjacent to the breakpoint of a widespread *Drosophila* inversion by a transposon-induced antisense RNA. Proc Natl Acad Sci USA 101:9013-9018.

Quentin Y (1992) Origin of the *Alu* family: a family of *Alu*-like monomers gave birth to the left and the right arms of the *Alu* elements. Nucleic Acids Res 20(13):3397-3401

Ranz JM, Maurin D, Chan YS, von Grotthuss M, Hillier LW, Roote J, Ashburner M, Bergman CM (2007) Principles of genome evolution in the Drosophila melanogaster species group. PLoS Biol. 5(6):e152.

Regner LP, Pereira MSO, Alonso CEV, Abdelhay E, Valente VLS (1996) Genomic distribution of *P* elements in *Drosophila willistoni* and a search for their relationship with chromosomal inversions. J Hered 87(3):191-198.

Rij RPV, Berezikov E (2009) Small RNAs and the control of transposons and viruses in *Drosophila*. Trends Microbiol 17(4):163-171

Robe LJ, Valente VLS, Budnik M, Loreto ELS (2005) Molecular phylogeny of the subgenus *Drosophila* (Diptera, Drosophilidae) with an emphasis on Neotropical species and groups: A nuclear versus mitochondrial gene approach. Mol Phylogenet Evol. 3:623-640

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

Rohde C, Valente VLS (1996a) Cytological maps and chromosomal polymorphism of *Drosophila polymorpha* and *Drosophila cardinoides*. Braz J Genet 19: 27–32.

Rohde C, Valente VLS (1996b) Ecological characteristics of urban populations of *Drosophila polymorpha* Dobzhansky e Pavan and *Drosophila cardinoides* Dobzhansky e Pavan (Diptera, Drosophilidae). Revta Bras Ent 1 (40): 75–79.

Salzano FM (1954) Chromosomal relations in two species of *Drosophila*. Am Nat 88: 399-405.

Sanchez-Gracia A, Maside X, Charlesworth B (2005) High rate of horizontal transfer of transposable elements in *Drosophila*. Trends Genet. 21(4):200-3.

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z *et al*. (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274(5288):765-768

Schaeffer SW, Bhutkar A, McAllister BF, Matsuda M, Matzkin LM, O'Grady PM, Rohde C, Valente VL, Aguadé M, Anderson WW, *et al*. (2008) Polytene chromosomal maps of

11 Drosophila species: the order of genomic scaffolds inferred from genetic and physical maps. Genetics 179(3):1601-1655.

Seberg O, Petersen G (2009) A unified classification system for eukaryotic transposable elements should reflect their phylogeny. Nat Rev Genet 10(4):276

Silva JC, Loreto ELS, Clark JB (2004) Factors that affect the horizontal transfer of transposable elements. Curr Issues Mol Biol. 6:57-72

Sinzelle L, Izsvák Z, Ivics Z (2009) Molecular domestication of transposable elements: from detrimental parasites to useful host genes. Cell Mol Life Sci 66(6):1073-1093

Staginnus C, Iskra-Caruana ML, Lockhart B, Hohn T, Richert-Pöggeler KR (2009) Suggestions for a nomenclature of endogenous pararetroviral sequences in plants. Arch Virol. 154(7):1189-1193.

Stalker HD (1953) Taxonomy and hybridization in the *cardini* group of *Drosophila*. Ann Ent Soc Am 46: 343-358.

Streisinger G (1946) The *cardini* species group of the genus *Drosophila*. J. N. Y. Ent. Soc. 54: 105-113.

Sturtevant AH (1916) Notes on North American Drosophilidae with descriptions of twenty-three new species. Ann Ent Soc Am 9: 323-343.

Sturtevant AH, Dobzhansky T (1936) Inversions in the third chromosome of wild races of *Drosophila pseudoobscura* and their use in the study of the history of the species. Proc Natl Acad Soc USA 22: 448-450.

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

Sturtevant, A.H. (1942). The classification of the genus *Drosophila*, with descriptions of nine new species. Univ. Texas Publs 4213: 5--51.

Tamura K, Subramanian S, Kumar S (2004) Temporal patterns of fruit fly (Drosophila) evolution revealed by mutation clocks. Mol Biol Evol 21(1):36-44.

Throckmorton L (1975) The phylogeny, ecology and the geography of *Drosophila*. In: King, R. C. (ed.) Handbook of Genetics. New York: Plenum Publ. Corp. 874pp.

Throckmorton LH (1982) The *virilis* species group. In: Ashburner, M.; Carson, H. L.; Thompson, Jr., J. N. (ed.) The genetics and Biology of *Drosophila*. 3b. London: Academic Press. p. 227-296.

Townsend JI, Wheeler MR (1955) Notes on Puerto Rican Drosophilidae, including descriptions of two new species of *Drosophila*. J. Agric. Univ. P. R. 39: 57-64.

Val FC, Vilela CR, Marques MD (1981) Drosophilidae of the Neotropical region. In: Ashburner M, Carson HL and Thompson Jr JN (ed) The Genetics and Biology of *Drosophila*. 3a. Academic Press, London. pp 123-168.

Vidal N, Ludwig A, Loreto ELS (2009) Evolution of *Tom, 297 17.6* and *rover* retrotransposons in Drosophilidae species. Mol Genet Genomics, doi 10.1007/s00438-009-0468-0.

Vieira C, Biémont C (1996) Selection against transposable elements in *D. simulans* and *D. melanogaster*. Genet Res 68(1):9-15.

Vieira C, Nardon C, Arpin C, Lepetit D, Biémont C (2002) Evolution of genome size in Drosophila. Is the invader's genome being invaded by transposable elements? Mol Biol Evol 19(7):1154-61.

Vilela CR (1984) Notes on the holotypes of four Neotropical species of the genus Drosophila (Diptera, Drosophilidae) described by A.H.Sturtevant. Revta Bras Ent 28(3):245-256.

Vilela CR, da Silva AFG, Sene FM (2002) Preliminary data on the geographical distribution of *Drosophila* species within morphoclimatic domains of Brazil. Revta Bras Ent 2 (46): 139-148.

Wasserman M (1982) Evolution of *repleta* group. In: Ashburner, M.; Carson, H. L.; Thompson, Jr., J. N. (ed.) The genetics and Biology of *Drosophila*. 3b. London: Academic Press. p. 61-139.

Wasserman M (1992) Cytological evolution of the *Drosophila repleta* group. In: Krimbas, C. D. e Powell, J. R. (ed.) *Drosophila* inversion polymorphism. Boca Raton, Florida: CRC Press. 560pp.

Wasserman M, Wilson FD (1957) Further studies on the *repleta* group. Univ Texas Publs 5721:132-156.

Wheeler MR (1981) The Drosophilidae: a taxonomic overview. In: Ashburner, M.; Carson, H. L.; Thompson, J. N. (ed.) The Genetics and Biology of *Drosophila*. 3a. New York: Academic Press. p. 1-97.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8(12):973-82.

Wilder JA, Dyreson EG, O'Neill RJ, Spangler ML, Gupta R, Wilder AS, Hollocher H (2004) Contrasting modes of natural selection acting on pigmentation genes in the *Drosophila dunni* subgroup. J Exp Zoolog B Mol Dev Evol 302(5):469-82

Williston SW (1896) Manual of the families and genera of North American Diptera. 2nd edition rewritten and enlarged. J.T. Hathaway (ed.). New Haven. 167pp.

Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J 9(10):3353-3362

Yeates DK, Wiegmann BM (2005) Phylogeny and Evolution of Diptera: Recent insights and new perspectives. In: Yeates DK, Wiegmann BM (eds) The evolutionary biology of flyes. Columbia University Press, New York, pp 14-44