

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
DEPARTAMENTO DE GENÉTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

**Resolução do conteúdo gênico dos elementos cromossômicos
e delimitação dos pontos de quebra de inversões: uma
abordagem citogenômica no organismo modelo *Drosophila
willistoni*.**

Carolina Flores Garcia

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).

Prof^a. Dr^a. VERA LÚCIA S. VALENTE GAIESKY – Orientadora

Porto Alegre, Dezembro de 2015.

Este trabalho foi desenvolvido no Laboratório de *Drosophila*, Departamento de Genética do Instituto de Biociências na Universidade Federal do Rio Grande do Sul, com recursos financeiros do Programa de Pós Graduação em Genética e Biologia Molecular da UFRGS, FAPERGS (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul – Processo PRONEX 10/0028-7) e Bolsa de Doutorado CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico); e no Grupo de Genómica, Bioinformática y Evolución - Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, sob orientação do Prof. Dr. ALFREDO RUIZ PANADERO, com bolsa do Programa Institucional de Doutorado Sanduíche no Exterior (PDSE) da CAPES (Coordenação de Aperfeiçoamento de Pessoal Nível Superior).

A compaixão pelos animais está intimamente ligada à bondade de caráter, e quem é cruel com os animais não pode ser um bom homem.

Arthur Schopenhauer – On The Basis of Morality (1840).

Dedico esta Tese à minha tia Dalva, a estrela D'alva da família, à minha mãe guerreira Marisa e ao meu tio Enio. O trio de toda a minha vida!

AGRADECIMENTOS

À Professora Vera, muito obrigada pela credibilidade em me passar um dos seus sonhos científicos, e assim, torná-lo meu sonho real. O seu sim para o aceite da minha entrada no laboratório de *Drosophila* agora chega neste momento tão raro em muitos sentidos. Sigo sendo muito grata e feliz por poder aprender com a senhora.

Amigos e colegas de Laboratório, já se passaram épocas distintas e este laboratório sempre se manteve como nosso local de trabalho com muitas alegrias, parcerias, grandes resultados obtidos e festas.

Ju Cordeiro, Ju WG, Ana, Hermes, Marícia, Liz, Adri Ludwig, Adri Sassi, Maríndia, Mário, Dirleane, Gisele, Gilberto, éramos um turbilhão!!! Foi, e sempre será, um prazer o trabalho, a convivência e a troca de conhecimentos com vocês.

Ju Cordeiro, Marícia, Hermes, nem preciso dizer que a saudade ali no lab da frente sempre vem! Tantas coisas, que aquele livro de histórias do Laboratório de *Drosophila* terá que ter alguns volumes. Obrigada pelas parcerias.

Obrigada Ju WG e Maríndia, que sempre me deram também apoio técnico em todos os momentos.

À turma Gilberto, Rebeca, Brenda, Paula, Lucas, que comigo éramos os seis no lab.

Obrigada Maríndia pela atenção, pela parceria e pelos debates e ensinamentos que me destes e que sempre me ajudaram.

A todos vocês, obrigada pelo compartilhamento de idéias, pelas conversas, pela mão amiga e pela intriga da oposição. A mim vocês deram a oportunidade de aprendizado em nível profissional e pessoal.

À Berê, Dani, Helena e Bibi pelo apoio técnico e convivência no Laboratório. Bibi muito obrigada pelo auxílio em diversos momentos, ainda mais agora no final do meu Doutorado com a manutenção dos estoques.

A las tres personas, sin las cuales este trabajo no sería posible:

Profesor Alfredo Ruiz, para mi una inspiración científica de mucho rigor y esmero. Gracias por toda la oportunidad y por las tantas veces que en diez minutos me enseñabas lo que a mi, me llevaría mucho tiempo.

Gracias Dra. Alejandra Delprat por la amabilidad en enseñarme las técnicas científicas, las cuales llevaré por toda mi vida. Gracias por la amistad, el cariño y por dejares Jordi ser mi perro prestado en Barcelona.

Gracias Profesora Beatriz Goñi, de la Universidad de la República del Uruguay, por la atención. Su trabajo y esfuerzo que lograron la línea SG12.00 fueran las bases de este estudio.

Obrigada à minha família: mãe Marisa, tia Dalva, tio Enio, Vó Eva (*in memorian*), meus irmãos Giuliano e Rafael. Tudo que tenho conquistado na minha vida é devido a este apoio incomensurável que tenho de vocês. Tomara um dia eu possa fazer por vocês, o mesmo tanto que fazem por mim.

Ao Leonardo, por estar ao meu lado me ajudando em todos os momentos. Por cuidar de mim em qualquer lugar do mundo. Pelo companheirismo, o carinho e o grande amor.

Ao PPGBM na figura de seus professores e funcionários. Ao Elmo pela presteza e gentileza que se expandem além das funções profissionais.

Ao CNPq e a CAPES pelo auxílio financeiro

SUMÁRIO

LISTA DE ABREVIATURAS, SÍMBOLOS E UNIDADES	10
LISTA DE FIGURAS	12
LISTA DE TABELAS	15
RESUMO.....	17
ABSTRACT	19
CAPÍTULO I	21
I – Introdução	21
I.1 - Rearranjos cromossômicos estruturais no gênero <i>Drosophila</i>	21
I.2 – Estudos populacionais-evolutivos de inversões cromossômicas no gênero <i>Drosophila</i>	29
I.3 - Pontos de quebra das inversões em <i>Drosophila</i> : origem e distribuição cromossômica.....	34
I.4 – Caracterização dos pontos de quebra de inversões em <i>Drosophila</i> e mecanismos que originam inversões cromossômicas.....	40
I.5 - <i>Drosophila willistoni</i> como organismo modelo para os estudos de inversões cromossômicas.....	56
I.5.1 - <i>Drosophila willistoni</i> na Era Genômica	58
OBJETIVOS	61
CAPÍTULO II	62
II - Materiais e Métodos.....	62
II.1 - Populações utilizadas e manutenção dos estoques de <i>Drosophila willistoni</i>	62
II.2 – Análises citogenéticas	64
II.3 - Extração de DNA genômico	72
II.4- Estabelecimento das sondas no genoma	72
II.5 – Desenho de <i>primers</i>	73
II.6 – Técnica da Reação em Cadeia da Polimerase (<i>PCR</i>) e medição dos fragmentos.....	73

II.7 – Clonagem	74
II.8 - Sequenciamento.....	75
II.9 - Análise das sequências.....	76
II.10 – Determinação da diferença do padrão de bandas do braço cromossômico IIL da linhagem sequenciada de <i>Drosophila willistoni</i> com relação ao fotomapa da espécie.....	76
II.11 – Cruzamentos entre as linhagens Gd-H4-1 e SG12.00 de <i>Drosophila willistoni</i>	76
II.13 - Mapeamento de marcadores genéticos no cromossomo II de <i>Drosophila willistoni</i>	77
II.14 – Delimitação dos pontos de quebra da inversão IIL-H do cromossomo II de <i>Drosophila willistoni</i>	78
Apêndice I	90
CAPÍTULO III	101
III - Resultados e Discussão	101
III.1 – Estudo dos pontos de quebra de inversões de <i>Drosophila willistoni</i> . I. Comparação do arranjo cromossômico do braço IIL entre a linhagem sequenciada Gd-H4-1 e o fotomapa da espécie.....	101
III.2 – Estudo dos pontos de quebra de inversões de <i>Drosophila willistoni</i> . II. Determinação dos diferentes arranjos cromossômicos do braço IIL entre as linhagens Gd-H4-1 e SG12.00 (do Uruguai) por meio de cruzamentos.	107
III.3 – Estudo dos pontos de quebra de inversões de <i>Drosophila willistoni</i> . III. Redefinição da ordem e da orientação dos <i>scaffolds</i> estabelecidos na montagem do genoma do cromossomo II de <i>Drosophila willistoni</i>	112
III.4 – Estudo dos pontos de quebra de inversões de <i>Drosophila willistoni</i> . IV. Mapeamento físico de alta resolução e delimitação gênica dos pontos de quebra da inversão IIL-H.	128
CAPÍTULO IV	161
Conclusões e Perspectivas.....	161
CAPÍTULO V	164
Reassignment of <i>Drosophila willistoni</i> genome scaffolds to chromosome II arms.....	164

REFERÊNCIAS BIBLIOGRÁFICAS DA TESE.....	179
Anexo I.....	194
Rearranjos cromossômicos estruturais em organismos diploides: Consequências meióticas e mecanismos de origem.....	194
Referências Bibliográficas Anexo I	213
Anexo II.....	217
Tabela Anexo II: Sumário histórico dos estudos sobre cromossomos e polimorfismo cromossômico de <i>Drosophila willistoni</i> , e demais espécies crípticas do subgrupo <i>willistoni</i> de <i>Drosophila</i>	217
Referências Bibliográficas Anexo II.....	222
Anexo III	227
Evolução cromossômica em eucariotos – Rearranjos cromossômicos e elementos de transposição	227

LISTA DE ABREVIATURAS, SÍMBOLOS E UNIDADES

μL = microlitros

Adh = gene Álcool desidrogenase (*Alcohol dehydrogenase*)

BACs = Cromossomos artificiais de bactéria (*Bacterial Artificial Chromosomes*)

bw = gene *brown*

cM = cetimorgan, unidade de medida de mapa de ligação

DNA = ácido desoxirribonucleico

dNTP = desoxirribunocletídeo 5'- trifosfato

EDTA = Ácido etilenodiamino tetra-cético

g = gramas

GL = Grau Gay Lussac (quantidade em mililitros de álcool absoluto contida em 100 mililitros de mistura hidro-alcoólica)

h = horas

HIS = Hibridação *in situ* não fluorescente (*In Situ Hybridization*)

ITRs = Repetições Terminais Invertidas (*Inverted Repeats Terminals*)

kb = kilobase (1.000 pares de base)

LTRs = Repetições Terminais Longas (*Long Terminal Repeats*)

M = Molar

Mb = Megabase (1.000.000 pares de base)

mM = Milimolar

Mya = Milhões de anos

N = Normal

ng = nanogramas (bilionésima parte do grama)

NaOH = Hidróxido de Sódio

NAHR = Recombinação homóloga não alélica (*Non allelic homologous recombination*)

NHEJ = União das extremidades não homólogas (*Non homologous end joining*)

pág = página

pb = pares de base (pair base)

px = gene *plexus*

PCR = Reação de polimerização em cadeia (*Polymerase Chain Reaction*)

rpm = rotações por minuto

SDS = dodecilsulfato de sódio

SINEs = *Short Interspersed Nuclear Elements*

TA = Temperatura Ambiente

TE(s) = Elemento(s) transponívei(s) (*Transposable elements*)

TSD = Duplicação do sítio alvo (*Target Site Duplication*)

U = Unidade

V = Volts

v/v = volume/volume (concentração percentual)

LISTA DE FIGURAS

- Figura I.1:** Inversões cromossômicas paracêntricas do braço cromossômico IIL de *Drosophila willistoni*: estado heterozigoto, com formação da alça de inversão; e estado homozigoto (homo) _____ 23
- Figura I.2:** Representação esquemática dos resultados de um único evento de recombinação meiótica (*crossing over*) dentro da alça de uma inversão, ao fim da meiose I e II _____ 25
- Figura I.3:** Eliminação seletiva de gametas anormais na meiose de fêmeas heterozigotas de *Drosophila melanogaster* _____ 27
- Figura I.4:** Delimitação citológica das inversões paracêntricas do braço cromossômico IIL de *Drosophila willistoni* _____ 39
- Figura I.5:** Representação esquemática dos dois principais mecanismos que originam inversões no gênero *Drosophila* _____ 41
- Figura II.1:** Representação esquemática do sistema indireto de detecção e revelação utilizados na hibridação *in situ* não fluorescente _____ 71
- Figura II.2:** Delimitação gênica do ponto de quebra distal da inversão IIL-H, pela técnica da *PCR* _____ 80
- Figura III.1.1:** Análise do arranjo cromossômico da linhagem sequenciada Gd-H4-1 de *Drosophila willistoni* _____ 102
- Figura III.2.1:** Arranjo cromossômico do braço IIL resultante do cruzamento entre as linhagens ♀ Gd-H4-1 x ♂ SG12.00 de *Drosophila willistoni* _____ 109
- Figura III.2.2:** Arranjo cromossômico do braço IIL resultantes do cruzamento entre as linhagens ♀ SG12.00 x ♂ Gd-H4-1 *Drosophila willistoni* _____ 110

- Figura III.2.3:** Representação esquemática das alças de inversões da dupla heterozigose (Inversões IIL-F e IIL-H) decorrentes dos cruzamentos recíprocos entre a linhagem sequenciada Gd-H4-1 e a linhagem SG12.00 de *Drosophila willistoni* _____ 111
- Figura III.3.1:** Esquematização da associação dos *scaffolds* do genoma do cromossomo II de *Drosophila willistoni*, com os respectivos cromossomos politênicos, com base nos marcadores disponíveis _____ 114
- Figura III.3.2:** Localização *in situ* da sonda do gene *bw* (*scaffold* 4954) em *Drosophila willistoni* _____ 115
- Figura III.3.3:** Localização *in situ* da sonda do gene *px* (*scaffold* 4558) em *Drosophila willistoni* _____ 116
- Figura III.3.4:** Localização *in situ* da sonda do gene *Adh* (*scaffold* 4577) em *Drosophila willistoni* _____ 117
- Figura III.3.5:** Mapa de ligação do cromossomo II de *Drosophila willistoni* inferido por Spassky & Dozhansky (1950) _____ 121
- Figura III.3.6:** Redefinição dos *scaffolds* e marcadores gênicos do cromossomo II de *Drosophila willistoni* (Schaeffer *et al.*, 2008), de acordo com este estudo (Tabela II.1) _____ 124
- Figura III.4.1:** Delimitação citológica da inversão IIL-H em *Drosophila willistoni* _____ 130
- Figura III.4.2:** Representação esquemática da denominação dos pontos de quebra distal e proximal da inversão IIL-H utilizada no presente estudo _____ 131
- Figura III.4.3:** Resultados dos mapeamentos físicos obtidos na etapa inicial da caminhada cromossômica do ponto de quebra distal da inversão IIL-H em *Drosophila willistoni* _____ 132
- Figura III.4.4:** Esquematização dos resultados obtidos por caminhada cromossômica, para a delimitação do ponto de quebra distal da inversão IIL-H, na linhagem SG12.00 de *Drosophila willistoni* _____ 136

- Figura III.4.5:** Visualização em gel de agarose 0,7% das cinco regiões intergênicas obtidas pela técnica da *PCR*, para a delimitação do ponto de quebra distal da inversão IIL-H de *Drosophila willistoni* _____ 137
- Figura III.4.6:** Mapeamento físico dos genes que delimitam o ponto de quebra distal da inversão IIL-H em *Drosophila willistoni* _____ 138
- Figura III.4.7:** Esquematisação dos resultados obtidos na primeira etapa para a delimitação do ponto de quebra proximal da inversão IIL-H de *Drosophila willistoni* _____ 141
- Figura III.4.8:** : Esquematisação dos resultados obtidos na segunda etapa para a delimitação do ponto de quebra proximal da inversão IIL-H de *Drosophila willistoni* _____ 143
- Figura III.4.9:** Esquematisação dos resultados obtidos na terceira etapa para a delimitação do ponto de quebra proximal da inversão IIL-H em *Drosophila willistoni* _____ 147
- Figura III.4.10:** Mapeamento físico das sondas 50 e 52b na linhagem SG12.00 de *Drosophila willistoni* _____ 149
- Figura III.4.11:** Mapeamento físico das sondas intergênicas 62iA e 62iB na linhagem SG12.00 de *Drosophila willistoni* _____ 152
- Figura III.4.12:** Esquematisação dos principais resultados obtidos na caminhada cromossômica para a delimitação gênica dos pontos de quebra da inversão IIL-H de *Drosophila willistoni* _____ 153
- Figura III.4.13:** Possíveis cenários para a origem das inversões IIL-H e IIL-F na linhagem SG12.00, de acordo com os resultados obtidos nesse estudo _____ 155
-

LISTA DE TABELAS

Tabela I.1: Estudos de caracterização molecular dos pontos de quebras de inversões em espécies do gênero <i>Drosophila</i>	43
Tabela II.1: Local e ano de coleta das linhagens de <i>Drosophila willistoni</i> analisadas no presente estudo	62
Tabela II.2: <i>Primers</i> e região genômica dos genes clonados <i>Dwil\GK21049</i> e <i>Dwil\GK21115</i>	75
Tabela II.3: Marcadores genéticos utilizados para a redefinição dos <i>scaffolds</i> do cromossomo II de <i>Drosophila willistoni</i>	81
Tabela II.4: Marcadores genéticos utilizados como sondas para os cromossomos X e III de <i>Drosophila willistoni</i>	83
Tabela II.5: Marcadores genéticos utilizados como sondas para a delimitação do ponto de quebra distal da inversão IIL-H de <i>Drosophila willistoni</i> via caminhada cromossômica	84
Tabela II.6: Regiões intergênicas e <i>primers</i> do <i>scaffold</i> scf2_1100000004851 utilizados para delimitação do ponto de quebra distal da inversão IIL-H de <i>Drosophila willistoni</i> , pela técnica da <i>PCR</i>	85
Tabela II.7: Marcadores genéticos utilizados como sondas para a delimitação do ponto de quebra proximal da inversão IIL-H de <i>Drosophila willistoni</i> via caminhada cromossômica, e pela técnica da <i>PCR</i>	86
Tabela Apêndice I: Especificações da reação de ligação plasmidial do vetor <i>pGEM -T Easy</i>	95
Tabela III.3.1: Redefinição dos <i>scaffolds</i> do cromossomo II de <i>Drosophila willistoni</i>	119

Tabela III.3.2: Posições genômicas dos genes utilizados para o estabelecimento de sondas e posições citológicas obtidas no presente estudo por meio da análise da designação dos *scaffolds* dos cromossomos X e III de *Drosophila willistoni* _____ 125

Tabela III.4.1: Marcadores físicos utilizados para a delimitação do ponto de quebra distal da inversão IIL-H, por caminhada cromossômica, na linhagem SG12.00 de *Drosophila willistoni* _____ 134

Tabela III.4.2: Intervalos gênicos (denominados de A-E) utilizados para a delimitação do ponto de quebra distal da inversão IIL-H, pela técnica da *PCR*, em ambas as linhagens de *Drosophila willistoni*, de acordo com a Figura II.2 _____ 135

Tabela III.4.3: Marcadores físicos utilizados na primeira e segunda etapas da caminhada cromossômica, para o estabelecimento do ponto de quebra proximal da inversão IIL-H na linhagem SG12.00 de *Drosophila willistoni* _____ 142

Tabela III.4.4: Marcadores físicos utilizados na terceira etapa da caminhada cromossômica, para o estabelecimento do ponto de quebra proximal da inversão IIL-H na linhagem SG12.00 de *Drosophila willistoni* _____ 146

Tabela III.4.5: Marcadores físicos utilizados na quarta etapa da caminhada cromossômica, para o estabelecimento do ponto de quebra proximal da inversão IIL-H na linhagem SG12.00 de *Drosophila willistoni* _____ 150

RESUMO

Drosophila willistoni é uma espécie pertencente ao subgênero *Sophophora* com origem e distribuição Neotropical. Esta espécie é um intrigante modelo biológico para diferentes pesquisas em genética evolutiva e de populações, evolução molecular e ecologia. Sua característica mais proeminente é a elevada ocorrência de polimorfismo cromossômico para inversões paracêntricas segregantes, sendo considerada por muitos especialistas como a espécie mais polimórfica do gênero *Drosophila*. O genoma da linhagem Gd-H4-1 de *Drosophila willistoni*, oriunda da Ilha Guadalupe (Caribe), foi sequenciado no Consórcio *Drosophila* 12 Genomes (2007). Nosso grupo de pesquisa do Laboratório de *Drosophila* da UFRGS é referência mundial no estudo da caracterização do polimorfismo cromossômico de *Drosophila willistoni*. A disponibilização do genoma desta espécie trouxe novos desafios e motivações para as análises envolvendo este organismo-modelo tão peculiar, especificamente no que concerne à gênese de suas inversões cromossômicas.

O presente estudo visa delimitar, pela primeira vez, os pontos de quebra da inversão IIL-H do cromossomo II de *Drosophila willistoni*, na linhagem sequenciada Gd-H4-1, e na linhagem SG12.00 coletada no Uruguai e portadora da inversão IIL-H fixada. Esta delimitação é o primeiro passo para a subsequente caracterização molecular desta região, a fim de tentar inferir o possível mecanismo que originou esta inversão e as consequências genômicas que esta possa acarretar.

A primeira análise comparou o padrão cromossômico do braço IIL entre a linhagem sequenciada Gd-H4-1 e o fotomapa de *Drosophila willistoni*, mostrando que o braço IIL da linhagem sequenciada apresenta os arranjos IIL-A e IIL-F fixados que o diferencia do arranjo do fotomapa. Dada a devida caracterização cromossômica da linhagem sequenciada, estabeleceu-se que esta seria o padrão para a análise dos pontos de quebra das inversões em *Drosophila willistoni* (**Capítulo III, Tópico III.1**).

Para o estudo dos pontos de quebra da inversão IIL-H faz-se necessária a comparação destas regiões genômicas, entre a linhagem sequenciada padrão e outra linhagem de *Drosophila willistoni* que possua a inversão IIL-H fixada. Sendo assim, escolheu-se a linhagem SG12.00, a qual tinha seu padrão cromossômico caracterizado. Já o estabelecimento das diferenças do padrão cromossômico no braço IIL entre as duas linhagens foi obtido por cruzamentos recíprocos entre estas, mostrando que o braço IIL da linhagem SG12.00 difere do arranjo cromossômico da Gd-H4-1 pela ocorrência dos arranjos IIL-A, IIL-F e IIL-H fixados (**Capítulo III, Tópico III.2**).

As análises acerca da montagem dos *scaffolds* do genoma do cromossomo II de *Drosophila willistoni* foram feitas a partir do estabelecimento de 18 sondas mapeadas fisicamente neste cromossomo da linhagem Gd-H4-1, por hibridação *in situ* não fluorescente. Adicionalmente, quatro sondas foram estabelecidas, uma para o braço cromossômico XL, uma para o cromossomo III, e duas para o braço cromossômico XR. Os resultados obtidos mudam a tradicional inferência dos Elementos de Muller respectivos aos braços IIL e IIR do cromossomo II de *Drosophila willistoni*, bem como mostram que a orientação dos *scaffolds* do braço cromossômico IIR estava invertida (**Capítulo III, Tópico III.3 e Capítulo V**).

Para a delimitação dos genes flanqueadores dos pontos de quebra distal e proximal da inversão IIL-H de *Drosophila willistoni* foram executadas diferentes etapas de planejamento e estabelecimento de sondas gênicas e intergênicas mapeadas fisicamente por hibridação *in situ* não fluorescente, juntamente com a tentativa de delimitação dos pontos de quebra pela técnica da *PCR*. O ponto de quebra distal da inversão IIL-H se encontra entre os genes *Dwil\GK21048* e *Dwil\GK21115*. Já a delimitação do ponto de quebra proximal da inversão IIL-H foi mais laboriosa e complexa, mostrando que este ponto está envolvido com o reuso de uma sequência de aproximadamente 1.212 pb pelo ponto de quebra distal da inversão IIL-F, bem como a ocorrência da duplicação desta mesma sequência (**Capítulo III, Tópico III.4**).

ABSTRACT

Drosophila willistoni is a Neotropical species member of the *Sophophora* subgenus. This species is an interesting model organism for several researchers on evolutionary biology, population genetics, molecular evolution and ecology. This species is mainly characterized by its huge chromosomal polymorphism of paracentric segregating inversions, being considered by several experts as the most polymorphic *Drosophila* species. The genome of the Gd-H4-1 strain of *Drosophila willistoni*, collected in the Caribbean Guadeloupe Island, was sequenced by the *Drosophila* 12 Genomes Consortium (2007). Our research group at *Drosophila* Laboratory of the UFRGS, member of such Consortium, is worldwide recognized in the study of the chromosomal polymorphism of *Drosophila willistoni*. The availability of the *Drosophila willistoni* fully sequenced genome created new challenges and incentivized the study of this special model organism, mainly respect to the genesis of its chromosomal inversions.

For the first time, the present Thesis aimed to characterize detect and delineate the breakpoints of a chromosomal inversion in *Drosophila willistoni* - the IIL-H inversion in the Gd-H4-1 and in the Uruguayan SG12.00, a homozygous strain for the IIL-H inversion. This is the first step, for a further molecular characterization of this region, in order to reveal the possible mechanism that generated that inversion and the genomic consequences of this event.

The first analysis here performed, compared the chromosomal banding pattern of the IIL arm with that of the sequenced Gd-H4-1 and those of the photomap of *Drosophila willistoni*, showing that IIL of the sequenced strain presents the IIL-A and IIL-F in homozygosis. Such characteristic differentiate Gd-H4-1 of the assumed standard arrangement. Considering the characterization of the sequenced strain, it was established that it should be considered as standard for the analysis of the breakpoints of the *Drosophila willistoni* inversions (**Chapter III, Topic III.1**).

To determine the breakpoints of the IIL-H inversion, it was necessary to compare such regions with those of other strain of *Drosophila willistoni*, with the IIL-H inversion fixed in homozygosis, the Uruguayan strain SG12.00. The establishment of the precise difference between the chromosomal arrangements of the IIL arm of both strains was obtained through reciprocal crossings between them. It was observed that the IIL arm in the SG12.00 strain differ of that of the Gd-H4-1 strain, by the occurrence of the IIL-A, IIL-F e IIL-H fixed arrangements (**Chapter III, Topic III.2**).

Analyses of the *scaffolds* assemblage of the chromosome II of *Drosophila willistoni* were performed through the use of 18 probes mapped by non-fluorescent *in situ* hybridization in the chromosomes of the Gd-H4-1 strain. Additionally, four other probes were hybridized, one in the XL chromosomal arm, one in the third chromosome and two in the XR arm. The results obtained in this study changed the traditional inference of the genic content of Muller Elements for the IIL and IIR chromosomal arms of *Drosophila willistoni*, and demonstrated that the orientation of the *scaffolds* of the chromosomal arm IIR was inverted (**Chapter III, Topic III.3**).

To determine the genes flanking the distal and proximal breakpoints of the IIL-H inversion of *Drosophila willistoni* we followed several methodological steps. The first was the planning and the choice of probes of genic and intergenic sequences, and their physical mapping by non-fluorescent *in situ* hybridization and an attempt to define the breakpoints by *PCR*. It was found that the distal breakpoint of the IIL-H inversion occurs between the genes *Dwil\GK21048* and *Dwil\GK21115*. The delimitation of the proximal breakpoint of IIL-H was laborious and more complex, showing that this breakpoint is involved with the reuse of a 1.212 pb sequence by the distal breakpoint of the IIL-F inversion, and is also involved with the duplication of this same sequence (**Chapter III, Topic III.4**).

CAPÍTULO I

I – Introdução

I.1 - Rearranjos cromossômicos estruturais no gênero *Drosophila*

Rearranjos cromossômicos estruturais originam-se de quebras cromossômicas em diferentes sítios, seguidas da reconstituição destas quebras em uma combinação distinta. Compreendem as deleções, duplicações, inversões e translocações. Estes rearranjos são reconhecidos quando afetam a morfologia cromossômica ou mudam a posição de certos marcadores, tais como o centrômero, as faixas dos cromossomos politênicos, diferentes tipos de bandas obtidas pelas técnicas de bandeamento, entre outros; envolvendo grandes quantidades de material genético e que são visíveis em microscopia óptica (Guerra, 1988) (Veja Anexo 1, pág. 194).

Altas taxas de rearranjos cromossômicos são notavelmente abundantes em *Drosophila* (Krimbas & Powell, 1992). A análise destes no cariótipo das espécies deste gênero foi bastante favorecida devido à presença dos cromossomos politênicos (Painter, 1933).

Os cromossomos politênicos de *Drosophila* formam-se em núcleos interfásicos e são os produtos finais de sucessivos ciclos de replicação sem a consequente separação das cromátides filhas, resultando em uma grande estrutura que apresenta bandamento natural, formado pela sinapse precisa dos cromômeros paralelos das cromátides irmãs politenizadas. Estima-se que os cromossomos politênicos encontrados nas glândulas salivares sofrem 2^{10} eventos de replicação, ou seja, chegam a gerar até 1024 filamentos em cada par cromossômico de uma célula diploide (Swift, 1962), originando assim uma magnitude de visualização ímpar. Tecidos e órgãos contendo células com cromossomos politênicos estão, de maneira geral, envolvidos em intensas

funções secretoras realizadas durante um curto espaço de tempo, num contexto de crescimento rápido. Outra peculiaridade dos cromossomos politênicos interfásicos é que, como estes não sofrem segregação após a replicação, os cromossomos parentais permanecem unidos e pareados na mesma conformação visualizada somente na meiose I da maioria dos outros organismos (Zhimulev & Koryakov, 2009).

A estrutura física dos cromossomos politênicos permite a análise precisa dos diferentes rearranjos cromossômicos em *Drosophila*, destacando-se as inversões, as quais são o tipo mais frequente no gênero (Powell, 1997). Este rearranjo consiste na quebra simultânea em dois sítios de um cromossomo e a reorganização desta região com a ordem 180° invertida.

Em organismos diploides, as inversões são classificadas em dois tipos: paracêntricas (as quais não envolvem o centrômero na sua formação, ocorrendo no mesmo braço cromossômico) e pericêntricas (as quais envolvem o centrômero e mais de um braço cromossômico). Durante o pareamento dos cromossomos homólogos na meiose I, este rearranjo pode ser visualizado em estado heterozigoto, quando somente um dos cromossomos parentais porta a inversão; neste caso formando uma alça de inversão para o correto pareamento dos cromossomos homólogos; ou pode ser visualizado em estado homozigoto, quando ambos os cromossomos parentais portam a inversão. Estas conformações cromossômicas podem ser visualizadas nos cromossomos politênicos interfásicos de *Drosophila* (**Figura I.1**) (Hartl & Jones, 2009).

As inversões cromossômicas, em comparação com os demais rearranjos cromossômicos estruturais, costumam ser mais bem suportadas pelos organismos que as portam, uma vez que ela não implica, teoricamente, aumento ou redução do material genômico. Entretanto, uma inversão que ocorra dentro de um gene pode resultar em mutação, muitas vezes letal para o organismo. A alteração da posição dos genes com relação uns aos outros e às suas sequências controladoras, o que se denomina Efeito de Posição, também é outra consequência da inversão; podendo resultar em alterações na expressão gênica, e conseqüentemente alterações em nível fenotípico (Griffiths *et al.*, 2008)

Um exemplo clássico de Efeito de posição vem dos estudos feitos com fêmeas de *Drosophila melanogaster* heterozigotas para o alelo que confere a mutação de cor de olho *white* (w^+/w) no cromossomo X.

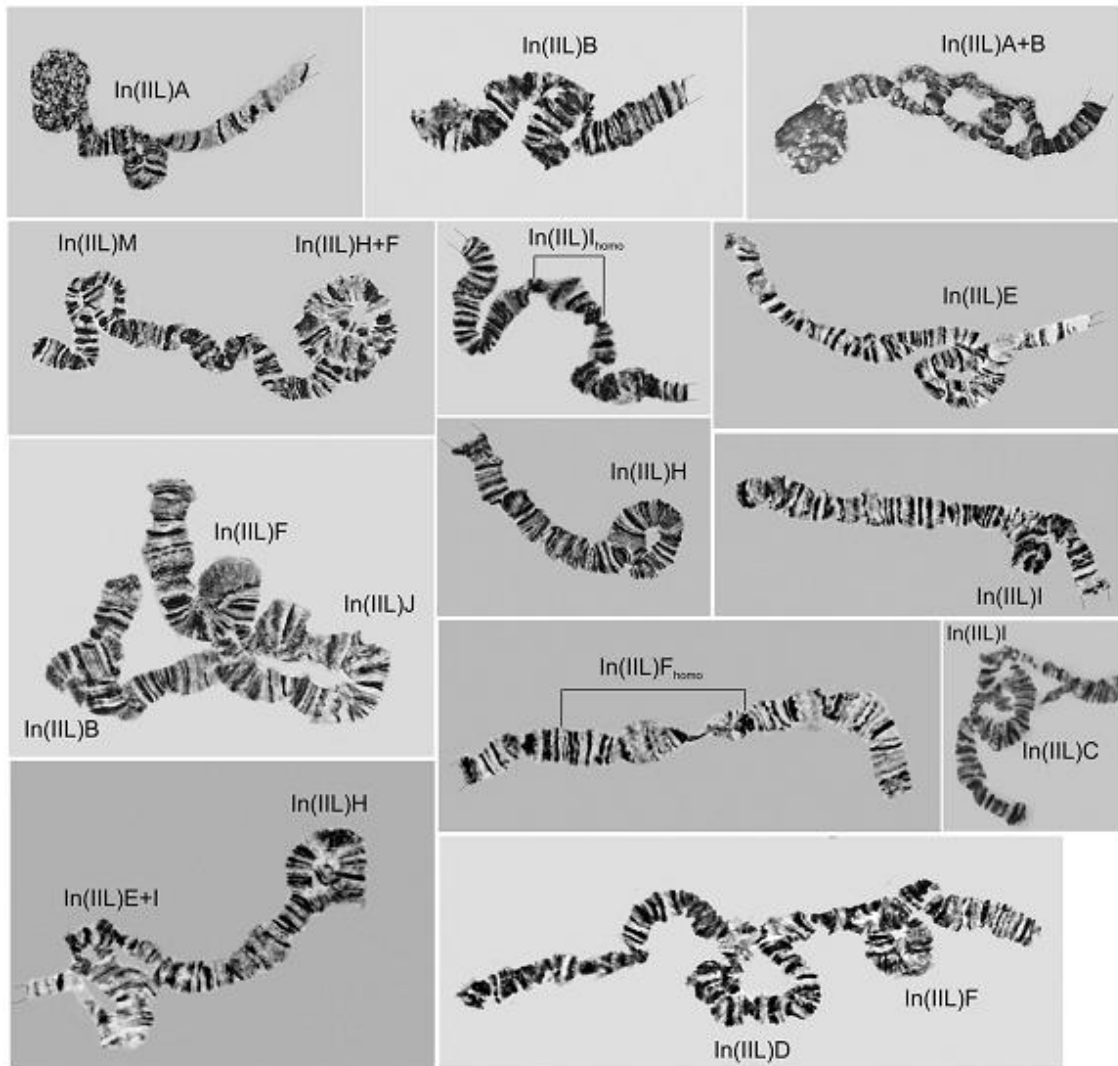


Figura 1.1: Inversões cromossômicas paracêntricas do braço cromossômico IIL de *Drosophila willistoni*: estado heterozigoto, com formação da alça de inversão; e estado homozigoto (homo). Retirado de Rohde & Valente (2012).

A presença de uma inversão heterozigota que leve o alelo selvagem (w^+) para a proximidade do centrômero, de tal maneira anula a expressão do alelo selvagem e permite a expressão do alelo mutante recessivo (w) em diferentes

níveis, resultando em prole com olhos variegados com manchas vermelhas e brancas (Schultz, 1936; Talbert & Henikoff, 2000).

O Efeito de Posição também foi constatado por Puig *et al.* (2004), analisando o efeito da inversão $2j$, do cromossomo 2 de *Drosophila buzzatii*, no gene *CG13617*. Este gene se encontra muito próximo ao ponto de quebra proximal desta inversão, e os embriões de linhagens homozigotas apresentam a expressão 5x mais baixa comparada às linhagens sem a presença da inversão. Pela caracterização dessa região no genoma de *D. buzzatii*, e análises dos níveis de mRNA, os autores apontaram que o elemento de transposição (*TE*) denominado *Kepler* é responsável por originar um RNA antisense, o qual se complexa ao mRNA do gene *CG13617*, tornando-o inativo; exercendo assim uma regulação pós-transcricional do gene. O *TE Kepler* está inserido adjacente ao ponto de quebra proximal nas linhagens que portam a inversão $2j$, e não se encontra nessa mesma região do ponto de quebra nas linhagens sem a inversão. Os resultados deste estudo mostram que os mecanismos de influência das inversões nas regiões adjacentes podem se apresentar das mais variadas formas, contando inclusive com a atuação de sequências a elas associadas.

O comportamento de uma inversão heterozigota e as consequências que ela pode acarretar difere durante a meiose e a mitose. Durante a meiose I, conforme já dito acima, para o pareamento correto dos cromossomos homólogos ocorre a formação de uma alça de inversão. A ocorrência de recombinação meiótica (*crossing over*) dentro de uma alça de inversão paracêntrica leva à formação de um cromossomo dicêntrico (presença de dois centrômeros), e um fragmento acêntrico (ausência do centrômero) na meiose I, resultando em gametas com deleções (**Figura I.2a**). Por sua vez, a ocorrência de recombinação meiótica na alça de uma inversão pericêntrica resulta na segregação normal dos cromossomos durante a meiose I, uma vez que os centrômeros estão contidos na inversão, porém, origina gametas com deleções e duplicações ao final da meiose II (**Figura I.2b**) (Griffiths *et al.*, 2008).

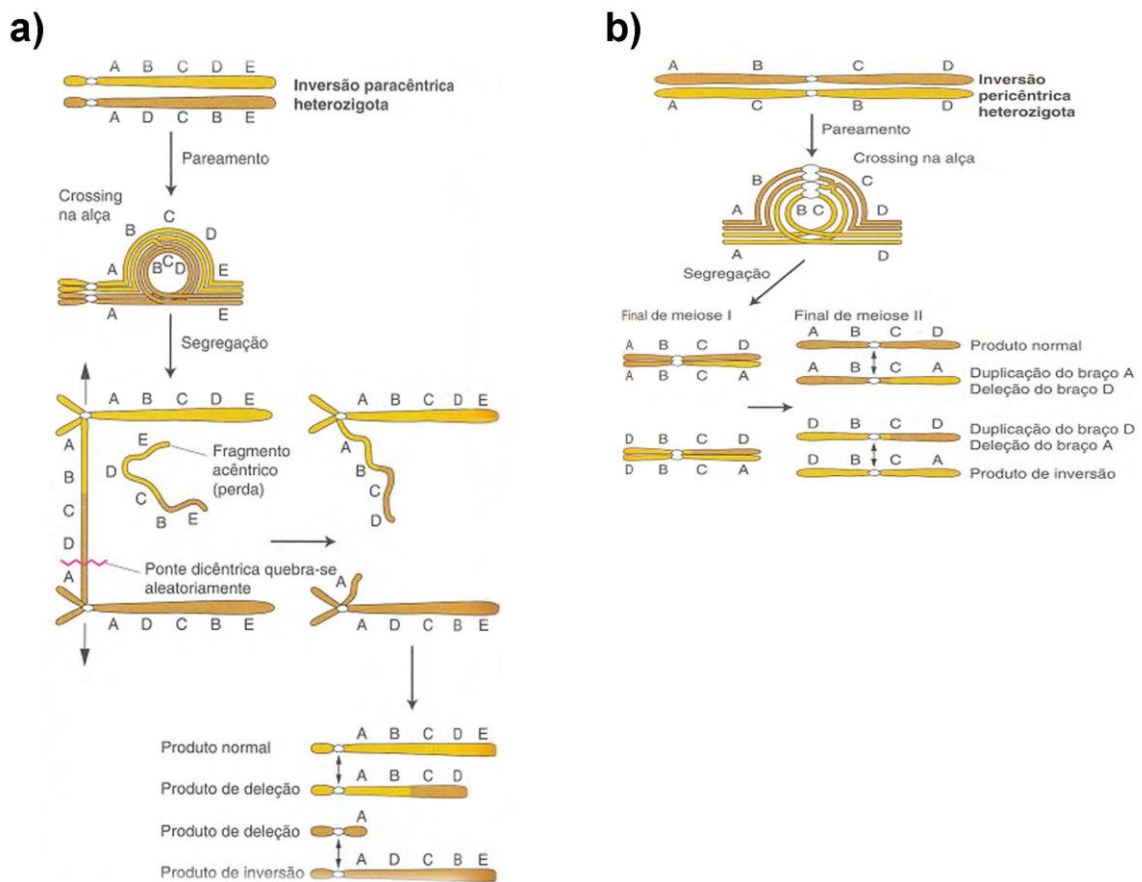


Figura I.2: Representação esquemática dos resultados de um único evento de recombinação meiótica (*crossing over*) dentro de uma alça de inversão, ao fim da meiose I e II. **a)** paracêntrica, com um fragmento acêntrico e um cromossomo dicêntrico na meiose I, resultando em gametas com deleções. **b)** pericêntrica, resultando em gametas com deleções e duplicações. Retirado de Griffiths *et al.* (2008).

Durante a mitose, a presença de uma inversão heterozigota não implica maiores dificuldades para o andamento do ciclo, uma vez que cada cromossomo se duplica, e as cromátides irmãs são direcionadas às células filhas resultantes.

A alta resolução de análise dos cromossomos politênicos, somada ao fato de que mais da metade das espécies de *Drosophila* estudadas são naturalmente polimórficas para inversões (Powell, 1997), caracterizaram as espécies desse gênero como organismos modelos para os estudos das inversões cromossômicas. Entretanto, com base no conhecimento das desestabilizações genômicas e efeitos na produção de gametas que as inversões podem originar, a alta ocorrência de polimorfismo cromossômico não é esperada *a priori* nos diferentes seres vivos. O que parece levar as espécies do gênero *Drosophila* a apresentarem tantas inversões paracêntricas, sem um maior efeito deletério no seu sucesso reprodutivo, é a presença de mecanismos de defesa, presentes em fêmeas e machos deste gênero, contra a produção de gametas portadores de rearranjos cromossômicos não balanceados (Orr-Weaver, 1995; Lake & Hawley, 2012).

Em fêmeas de *Drosophila* ocorre um mecanismo que elimina seletivamente os gametas recombinantes. Trata-se de um evento de permuta meiótica nas fêmeas durante a formação dos corpúsculos polares, de forma que o núcleo funcional do oócito (o qual será fecundado) contenha um cromossomo balanceado; ou com a ordem normal, ou com a ordem invertida (**Figura I.3**) (Hinton & Lucchesi, 1960).

Em machos, o mecanismo de proteção contra a produção de gametas inviáveis parece ser a supressão da recombinação na espermatogênese (Powell, 1997), conforme inferido por Morgan (1912) para machos de *D. melanogaster*. Este mecanismo de defesa parece ser preponderante também na espécie *Drosophila mediopunctata*, de acordo com a análise feita por Cavasini *et al.* (2010). Mutações em genes que afetam a segregação dos cromossomos que não tiveram permuta meiótica em fêmeas de *Drosophila*, não apresentam o mesmo efeito em machos, o que sugere que a permuta não é necessária para a correta segregação dos cromossomos homólogos na meiose I dos machos deste gênero (Orr-Weaver, 1995).

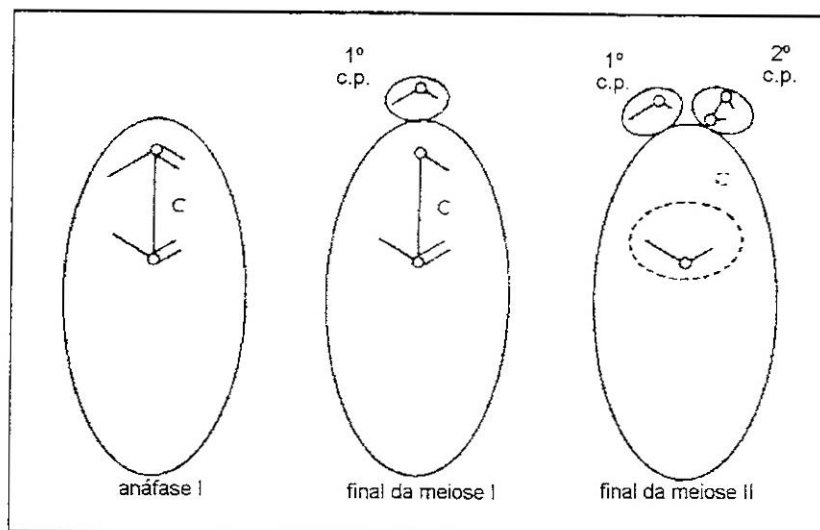


Figura 1.3: Eliminação seletiva de gametas anormais na meiose de fêmeas heterozigotas de *Drosophila melanogaster*. O primeiro corpúsculo polar (c.p.) a ser excluído é uma das cromátides balanceadas (ordem padrão, ou ordem invertida). O segundo corpúsculo polar eliminado é o cromossomo dicêntrico. O fragmento acêntrico não se orienta no fuso meiótico e posteriormente é degradado. O último corpúsculo polar a ser eliminado, o qual será efetivamente fecundado, apresenta também a ordem padrão, ou a ordem invertida. Extraído de Swanson *et al.* (1967).

Apesar da inferida supressão da recombinação em machos, relatos da sua ocorrência em nível meiótico estão presentes na literatura, evidenciando algumas particularidades. Dentre estas se destaca a alta ocorrência em machos que apresentam o fenômeno da disgenesia do híbrido de diferentes espécies. Este fenômeno se caracteriza também pela presença de altas frequências de prole inviável, mutações, alterações cromossômicas estruturais e distorção da razão de transmissão de alelos por um dos sexos (Kidwell *et al.*, 1977).

Outra particularidade é a ocorrência espontânea em machos de espécies com alto grau de polimorfismo para inversões paracêntricas, tais como *D. melanogaster* (Henderson *et al.*, 1978), *D. ananassae* (Kikkawa, 1937;

Goñi *et al.*, 2012) e *D. willistoni* (França *et al.*, 1968; Santos-Colares *et al.*, 2004).

No que concerne à *D. willistoni*, estudos clássicos feitos por Battaglia & Birch (1956) e França & Da Cunha (1968) apontaram a ocorrência de recombinação envolvendo as regiões situadas entre as inversões paracêntricas do cromossomo II em uma taxa de mais de 19%, e para inversões do cromossomo III em uma taxa de ocorrência de 11%. Já França *et al.* (1968) reportaram a ocorrência de um evento de recombinação em 207 cromossomos analisados em machos desta espécie.

O estudo de Santos-Colares *et al.* (2004) aponta a dificuldade de análises envolvendo a recombinação em machos de *D. willistoni*, devido às grandes limitações para obtenção de linhagens monomórficas estáveis. Este estudo, em uma tentativa de monitorar os efeitos da recombinação meiótica ao longo do tempo em machos desta espécie, analisou simultaneamente os cromossomos politênicos e meióticos dos mesmos machos pertencentes a duas linhagens: uma sendo híbrida entre uma população do nordeste versus uma população do extremo sul do Brasil; e uma linhagem que na época havia sido mais recentemente coletada (ano 2000) e que portava uma inversão pericêntrica envolvendo os braços XL e XR (Rohde *et al.*, 1998). A população híbrida apresentou pontes anafásicas em indivíduos que portavam as inversões IIL-F e IIL-D+E em heterozigose. Por sua vez, a população mais recentemente coletada apresentou a ocorrência de bivalentes na fase diplóteno em um indivíduo portando a inversão IIL-H desta população. Esses resultados apontam prováveis eventos de recombinação envolvendo estas inversões nos machos analisados. Os resultados das análises clássicas de Battaglia & Birch (1956) e França & Da Cunha (1968) somados aos resultados de Santos-Colares *et al.* (2004) ressaltam peculiaridades do cromossomo II de *D. willistoni* e seu provável envolvimento em eventos de recombinação meiótica associada a polimorfismos cromossômicos em machos e fêmeas. No entanto, o estudo de recombinação meiótica nesta espécie permanece pouco explorado.

Apesar das exceções, a presença de inúmeros casos de heterozigidade múltipla ocorrendo em diversas espécies de *Drosophila*

embasa a grande eficiência destes mecanismos, e nos direcionam para os aspectos evolutivos da ampla ocorrência e da adaptabilidade conferida pelo polimorfismo cromossômico aos portadores de inversões paracêntricas encontradas no gênero.

I.2 – Estudos populacionais-evolutivos de inversões cromossômicas no gênero *Drosophila*

O alto polimorfismo para inversões cromossômicas no gênero *Drosophila* tem servido como um sistema modelo para os estudos dos diferentes processos adaptativos, envolvidos na manutenção da variação genética. Os estudos pioneiros envolvendo a análise de inversões cromossômicas em populações naturais das espécies *Drosophila persimilis* e *D. pseudoobscura* vieram das inquietações do evolucionista Theodosius Dobzhansky e colaboradores, há mais de 60 anos (revisão em Dobzhansky, 1970). Seus achados foram catalisadores para muitas das descobertas que constituem a base da síntese evolutiva moderna (Dobzhansky, 1937; 1941; 1951), a qual mescla de maneira intrincada a teoria da evolução das espécies moldada pela seleção natural de Charles Darwin (1859), com os padrões de herança mendelianos e a genética de populações.

A obra de Dobzhansky "*Genetics and the Origin of Species*" (1937 e edições posteriores) foi um grande estímulo para o desenvolvimento da experimentação na genética evolutiva e de populações (revisão em Powell, 1997).

Os numerosos experimentos com a espécie *D. pseudoobscura* realizados por Dobzhansky, e colaboradores, foram a base para a postulação do modelo de coadaptação dos genes contidos nas inversões (Dobzhansky, 1949). Dobzhansky estabeleceu que a reduzida recombinação nas inversões cromossômicas nesta espécie (conforme visto, devido a diferentes mecanismos de defesa contra a produção de gametas portadores de rearranjos cromossômicos não balanceados) é capaz de manter combinações positivas

de genes em epistasia com outros arranjos gênicos existentes na população. Sendo assim, complexos de genes ligados em uma inversão nos diferentes tipos cromossômicos são herdados como blocos e raramente são corrompidos por recombinação meiótica, e a heterozigosidade seria preferível à homozigosidade, conforme predito pela seleção balanceadora (revisão em Krimbas & Powell, 1992; Powell, 1997; Hoffmann *et al.*, 2004). A partir daí, a análise e caracterização do polimorfismo cromossômico para inversões de populações naturais de outras espécies tornou-se extensivamente explorada, e evidências indiretas para a associação de inversões cromossômicas conferindo maior adaptação aos indivíduos que as portam foram reportadas.

Por mais de 40 anos Dobzhansky e colaboradores analisaram o polimorfismo cromossômico de *D. pseudoobscura* e suas variações em resposta às mudanças sazonais, anuais e geográficas. Esta espécie apresenta ampla distribuição geográfica na América do Norte, estando presente desde o Canadá ocidental, os Estados Unidos e parte da América Central, com a presença de uma subespécie na Colômbia (*D. pseudoobscura* Bogotana). Os arranjos *ST*, *AR*, *CH* e *TL*, ocorrentes no cromossomo 3 desta espécie são amplamente monitorados e tradicionalmente apresentam uma clina em suas frequências, estando a inversão *ST* com altas frequências na costa do Pacífico e nas populações do oeste de Sierra Nevada, onde *AR*, *CH*, e *TL* também ocorrem. Já a inversão *AR* torna-se a inversão predominante a partir de Sierra Nevada até as *Rocky Mountains* (Montanhas rochosas). Por sua vez, a inversão *CH* é predominante no sul da Califórnia, enquanto que o oeste da Califórnia apresenta alta diversidade para todas as inversões. Os arranjos *ST*, *AR* e *CH* têm-se apresentado com similares frequências ao longo dos anos relacionadas às adaptações locais que elas conferem, entretanto a inversão *TL* apresentou uma alta na sua frequência a partir da década de 80, sendo relacionada às mudanças ambientais (Dobzhansky & Sturtevant, 1938; Dobzhansky & Epling, 1944; Anderson *et al.*, 1991; McGaugh & Noor, 2012).

Tradicionalmente, a associação entre variáveis climáticas e o polimorfismo cromossômico de *D. subobscura* tem sido interpretada como o resultado de adaptação. Esta espécie de origem Paleártica, nas décadas de 70/80 invadiu o continente americano (Brncic *et al.*, 1981). Estudos envolvendo

a frequência de inversões nas populações destas regiões mostram uma relação inversamente proporcional do aumento na frequência de inversões ocorrentes em baixas latitudes (áreas mais quentes) e um decréscimo na frequência de inversões ocorrentes em latitudes mais altas (áreas de clima frio) (Zivanovic & Mestres, 2012; Fragata *et al.*, 2014). Nesta espécie o polimorfismo cromossômico também tem sido associado à contaminação ambiental por metais pesados (Kenig *et al.*, 2015).

Na espécie de origem africana *D. melanogaster* clinas de variações na frequência das inversões cromossômicas cosmopolitas *In(2L)t*, *In(2R)A7S*, *In(3L)P* e *In(3R)P* têm sido associadas em diferentes populações do mundo à características ecológicas importantes, tais como altitude e clima (Singh & Das, 1992; Aulard & David, 2002, Reinhardt *et al.*, 2014).

Drosophila buzzatii é um membro das espécies cactofílicas do grupo *repleta*. Sua origem é reportada ao sul da América Latina, tendo sua ocorrência sido relatada na década de 1970 na região mediterrânea, Ilhas Canárias, África equatorial e Austrália, em associação com espécies do gênero de cactus *Opuntia*, as quais têm sido disseminadas por interferência humana (Fontdevila *et al.*, 1982). Nesta espécie, clinas latitudinais na frequência de algumas inversões têm sido inferidas para as populações das áreas originais e das áreas colonizadas. O polimorfismo descrito para o segundo cromossomo, como por exemplo, o arranjo *2st*, tem sido relacionado à resistência térmica e à viabilidade da larva (Betrán *et al.*, 1995; Hoffmann *et al.*, 2004).

As espécies do grupo Neotropical da *D. willistoni* também foram classicamente estudadas quanto ao seu polimorfismo cromossômico de populações naturais (Da Cunha *et al.*, 1950; Pavan, 1946; Da Cunha & Dobzhansky, 1954; Dobzhansky & Pavlovsky, 1972; Kastritsis, 1966; De Souza *et al.*, 1970; Valente & Morales, 1985; Valente & Araújo, 1985; 1986; Santos & Valente 1990; Valiati & Valente, 1997). Neste grupo figura a espécie *D. willistoni*, a qual é foco deste estudo e que será abordada no **Tópico 1.5, Capítulo I** desta Tese.

Evidências diretas entre a associação de inversões cromossômicas e pressões seletivas passaram a ser apresentadas com o avanço das técnicas

moleculares e o sequenciamento dos genomas. Uma considerável quantidade de dados parece confirmar o maior valor adaptativo dos indivíduos portadores de inversões em *Drosophila*. A inibição da recombinação em heterocariótipos mediada por inversões se mostra fundamental na manutenção de seu papel adaptativo. Os padrões de desequilíbrio de ligação localizados dentro das regiões invertidas refletem a história da inversão e fluxo gênico desde sua origem (Schaeffer *et al.*, 2003; Pegueroles *et al.*, 2010; McGaugh & Noor, 2012; Fuller *et al.*, 2014).

Embora um grande número de características seja associado de maneira indireta, ou direta, com as inversões, pouco progresso se obteve na definição das bases genéticas destas associações (Hoffmann & Rieseberg, 2008). Um exemplo de caso neste sentido vem do estudo da variação em nível genômico da inversão *In(3R)P*, em duas populações da Austrália (uma da região tropical, e outra da região temperada) da espécie *D. melanogaster* (Rane *et al.*, 2015). Esta inversão é conhecida por estar associada às adaptações climáticas e ao sucesso de um evento migratório evolutivamente recente (100 anos) desta espécie na Austrália. Os resultados desta análise suportam a hipótese de que a inversão *In(3R)P* está associada com a captura de alelos adaptados localmente, agindo de forma aditiva ou epistática, mostrando elevadas taxas de desequilíbrio de ligação na região dentro da inversão, e alta diferenciação desta região genômica entre a população tropical e a população da região temperada (Rane *et al.*, 2015).

Evidências da seleção atuando na manutenção de inversões têm sido também obtidas em outros organismos, incluindo ratos (Lyon, 2003) a mosca praga agrícola *Rhagoletis pomonella* (Feder *et al.*, 2003), o mosquito transmissor da malária *Anopheles funestus* (Ayala *et al.*, 2012) e seres humanos (Stefansson *et al.*, 2005).

Entretanto, cabe ressaltar que muitos polimorfismos cromossômicos ocorrentes nas diferentes espécies são adaptativamente neutros, e assim, sofrem menor pressão seletiva (ou nenhuma), tendo a sua fixação, ou perda, dependente do tamanho da população e migração. Estas inversões também podem alcançar altas frequências por meio de outros mecanismos, como por

exemplo, a inversão *In(1)Be* do cromossomo X de *D. melanogaster*, a qual provavelmente deve sua manutenção à distorção da razão de transmissão em machos da espécie (Corbett-Detig & Hartl, 2012).

Apesar da confirmada associação das inversões cromossômicas com a manutenção de combinações de alelos que se situam dentro desta; a recombinação gênica na região invertida de um cromossomo é possível porque gametas recombinantes viáveis surgem por recombinações meióticas múltiplas e também em consequência da conversão gênica (Chovnick, 1973). Navarro *et al.* (1997) e Navarro & Ruiz (1997) estudaram o efeito da recombinação gênica ao longo da região genômica envolvida com as inversões em *D. melanogaster*, espécie que na época tinha o genoma disponível. Os autores fornecem equações que predizem a taxa de recombinação em diferentes locais ao longo do cromossomo invertido. Três conclusões principais emergiram destas análises. Primeiro, a redução na taxa de recombinação é mais forte nas inversões menores. Segundo, o grande efeito das inversões dentro do segmento invertido é a redução da taxa de recombinação, mas não a da conversão gênica. A supressão da recombinação pode ser grande o suficiente para a conversão gênica dominar como o principal modo de fluxo gênico, mesmo a grandes distâncias físicas. A terceira, e mais importante, é que a redução da recombinação não é uniforme ao longo do cromossomo invertido. Assim, uma inversão heterozigota não só reduz a recombinação gênica no sítio genômico específico onde esta ocorre, mas também redistribui os eventos de recombinação ao longo de todo o cromossomo. De acordo com Andolfatto *et al.* (2001) as inversões reduzem o fluxo gênico entre diferentes arranjos para quase zero perto dos pontos de quebra, mas as taxas de recombinação podem ser elevadas nas regiões centrais da região invertida.

I.3 - Pontos de quebra das inversões em *Drosophila*: origem e distribuição cromossômica

Paralelamente aos estudos evolutivo-populacionais das inversões cromossômicas em *Drosophila*, a inquietude acerca dos mecanismos que originam estes polimorfismos nas populações já se fazia presente. Esta questão foi classicamente analisada por Sturtevant & Dobzhansky (1936). Neste estudo, os autores ressaltam que as inversões de um mesmo cromossomo podem em certos casos, esclarecer algum aspecto da história evolutiva da espécie analisada, ou espécies distintas quando estas permitem intercruzamentos. Sendo assim, este seria o princípio para construção de filogenias baseadas em inversões cromossômicas do tipo sobrepostas.

As inversões do cromossomo 3 de *D. pseudoobscura* foram o primeiro conjunto de dados genéticos usados para criar uma filogenia com base nestes rearranjos cromossômicos (Dobzhansky & Sturtevant, 1938). Para isto, um cromossomo de uma dada população foi inferido de maneira arbitrária como o arranjo padrão, sendo nomeada de *ST*. Os demais arranjos cromossômicos analisados foram nomeados *Arrowhead (AR)*, *Pikes Peak (PP)*, *Santa Cruz (SC)*, *Chiricahua (CH)*, e *Tree Line (TL)*, de acordo com o nome da localidade onde estes arranjos prevaleciam. O cruzamento de machos coletados na natureza (bem como machos descendentes das fêmeas coletadas) com fêmeas da linhagem *ST*, mostrava na descendência obtida na F1 as diferenças dos arranjos cromossômicos existentes entre as populações, devida a formação de alças de inversões. Nesta análise foi necessária a inferência de um arranjo central hipotético na filogenia, o qual nunca foi encontrado na natureza em trabalhos posteriores. Esta metodologia comparativa das inversões cromossômicas permitiu relacionar os diferentes rearranjos encontrados, oriundos de inversões heterozigotas sobrepostas, em uma árvore filogenética não enraizada. Entretanto, o ponto chave para esta análise era que todas as cópias de uma inversão particular teriam uma origem única, sendo o arranjo visto nos indivíduos de uma população uma réplica do arranjo único surgido no passado, em um único indivíduo ancestral comum; ou seja, sua

origem seria monofilética. Posteriormente, análises filogenéticas com bases moleculares corroboraram que as inversões do cromossomo 3 de *D. pseudoobscura* são de origem única (Wallace *et al.*, 2011).

A partir do estudo de Dobzhansky & Sturtevant (1938), o ponto de vista tradicional para a gênese das inversões cunhou-se de que estas são resultantes de duas quebras independentes, ocorrendo ao mesmo tempo, seguidas pela religação das partes quebradas do cromossomo em orientação invertida, em relação às regiões vizinhas. Sendo assim, as múltiplas inversões sobrepostas encontradas em muitas espécies de *Drosophila*, e que servem de base para a construção de filogenias (Dobzhansky & Sturtevant, 1938), teriam ocorrido sequencialmente, e não devido a ocorrência simultânea de múltiplas quebras. Quanto às inversões *in tandem* (inversões lado a lado), a coincidência de pontos de quebra é atribuída ao acaso, em eventos que ocorreram em diferentes tempos (revisão em Krimbas & Powell, 1992).

O que reforça a hipótese da origem monofilética das inversões é a raridade que uma inversão cromossômica se origine, e uma maior raridade de que dois eventos que originem a mesma inversão ocorram espontaneamente em um mesmo local do mesmo cromossomo (Sperlich & Pfriem, 1986).

A análise das relações filogenéticas das inversões sobrepostas proposta por Sturtevant & Dobzhansky (1936) considera a rota mais parcimoniosa (ou seja, aquela onde ocorreu menor número de inversões) para a inferência evolutiva. Com base nestas condições, filogenias foram construídas para distintos grupos de espécies, tais como *melanogaster* (Lemeunier & Ashburner, 1984), *cardini* (Heed & Russell, 1971) as drosófilas havaianas (Carson & Kaneshiro, 1976); *virilis* (Throckmorton, 1982), grupo *repleta* (Wasserman, 1960) e subgrupo *fasciola* (Diniz & Sene, 2004), subgrupo *willistoni* (Rohde *et al.*; 2006), entre outros.

Inversões cromossômicas também podem ser introduzidas em uma população via hibridização. O exemplo melhor embasado acerca disto vem da inversão introduzida por *Anopheles gambiae* como resultado de seu cruzamento na natureza com *A. arabiensis* (Besansky *et al.* 2003).

Levando em conta o ponto de vista tradicional da gênese de uma inversão, diversos trabalhos consideram que a distribuição das inversões ao longo dos cromossomos ocorre de forma aleatória, tal como o estudo clássico de Olvera *et al.* (1979). Neste estudo os autores revisaram e analisaram o polimorfismo cromossômico de *D. pseudoobscura* ao longo da distribuição geográfica da espécie. De fato, esta era uma época em que o modelo *Random breakage model*, proposto por Ohno (1973) estava sendo embasado em diferentes estudos com mamíferos (Nadeau & Taylor, 1984; Schoen, 2000).

Entretanto, estudos cada vez mais consistentes evidenciando a ocorrência de repetidas quebras no mesmo local para diferentes inversões em uma quantidade considerável de espécies, trouxeram dúvidas quanto à aleatoriedade da distribuição dos pontos de quebra. Estes locais, os quais apresentam esta recorrência, foram denominados de “*hot spot*” e podem implicar instabilidades estruturais particulares destas regiões.

Cáceres *et al.* (1997) analisaram 86 inversões paracêntricas descritas para as espécies do complexo da *D. buzzatii* e 18 inversões induzidas por introgressão em *D. buzzatii* por meio de cruzamentos com *D. koepferae*. Os autores constataram que as inversões que apresentam maior sucesso para a fixação nas espécies são as de tamanho intermediário, e que a distribuição dos pontos de quebra das inversões, levando em consideração a localização da banda envolvida na quebra, não é aleatória, encontrando em certos segmentos cromossômicos, até oito pontos de quebra na mesma banda. Resultados similares foram observados em *D. subobscura* (Krimbas & Loukas, 1980), drosófilas havaianas (Tonzetich *et al.* 1988) e *D. melanogaster* (Lemeunier & Aulard, 1992).

A disponibilização do genoma humano e de outros mamíferos mostraram os efeitos das limitações do modelo *Random breakage model*, uma vez que este não considerava (por não estarem disponíveis) incontáveis regiões dos genomas destes organismos. Pevzner & Tesler (2003) ao analisarem 281 blocos sintênicos de até 1 Mb compartilhados entre humanos e ratos (a maioria ainda não conhecidos até então) encontraram, por meio da comparação dos rearranjos cromossômicos ocorridos durante a divergência

entre as duas espécies, uma grande quantidade de pontos de quebras destes rearranjos localizados em alta proximidade uns com os outros. Esta característica não se encaixava na teoria *Random breakage model*, sendo então proposto pelos autores o modelo *Fragile breakage model* (Pevzner & Tesler, 2003). Este modelo se baseia na inferência de que os pontos de quebra dos rearranjos cromossômicos ocorrem principalmente dentro de curtos sítios frágeis do genoma (*hot spots*), ou seja, regiões com características mais propensas às quebras. Estes sítios frágeis podem corresponder às duplicações segmentares, às regiões com uma concentração elevada de elementos transponíveis (*TEs*), ou a uma estrutura palindrômica (Murphy *et al.*, 2005). Com a denominação desta característica de "reuso", os autores enfatizam que este não implica no uso da mesma posição genômica (em nível de nucleotídeo) repetidas vezes, mas sim que o ponto de quebra apresenta múltiplas regiões genômicas que originam rearranjos cromossômicos.

A hibridação *in situ* de 180 sondas em *D. repleta*, oriundas de clones da biblioteca genômica de BACs dos cromossomos X e 2 da espécie *D. buzzatii*, somados a 249 marcadores físicos previamente analisados nestes cromossomos para ambas as espécies (Ranz *et al.*, 2003) foram capazes de cobrir toda a extensão genômica destes cromossomos, proporcionando assim uma mapa físico de alta resolução (González *et al.*, 2007). Este estudo mostrou que os cromossomos X e 2 de *D. repleta* e *D. buzzatii* diferem por 12 inversões paracêntricas. Dos três casos de reuso descritos anteriormente (um no cromossomo X e dois no cromossomo 2; Ruiz & Wasserman, 1993), somente um presente no cromossomo 2 foi confirmado. Entretanto, outro reuso de ponto de quebra neste mesmo cromossomo, até então desconhecido, foi inferido (González *et al.*, 2007). Similar abordagem foi feita entre as espécies *D. buzzatii* e *D. uniseta*, encontrando dois casos de reuso de pontos de quebra no cromossomo 2 de *D. uniseta* (Prada *et al.*, 2011).

A mais recente revisão acerca do polimorfismo cromossômico de *D. willistoni* foi feita por Rohde & Valente (2012). Neste estudo os casos de reuso de pontos de quebra de inversões, em nível citológico nessa espécie, foram inferidos com maior precisão. O estudo reporta o reuso no braço cromossômico XL da subseção 3C pelas inversões XLXRp (ponto proximal da inversão

pericêntrica descrita por Rohde *et al.*, 2005) e XL-I, subseção 4A pelas inversões XL-B e XL-F₁, subseção 5B pelas inversões XL-D e XL-F₁, subseção 9C pelas inversões XL-B e XL-K, da subseção 11B pelas inversões XL-C e XL-F₂, da subseção 13C pelas inversões XL-D e XL-E e da subseção 14A pelas inversões XL-B e XL-F₂. Já no braço cromossômico XR há a ocorrência de reuso da subseção 19B pelas inversões XLXRp e XR-G, subseção 23B pelas inversões XR-D e XR-F₁, subseção 26A pelas inversões XR-B e XR-F₁, subseção 28A pelas inversões XR-C e XR-D, e subseção 30C pelas inversões XR-B, XR-C e XR-H.

No braço IIL do cromossomo II foram apontados reusos de pontos de quebras na subseção 41B pelas inversões IIL-B e IIL-M, subseção 42B pelas inversões IIL-D e IIL-E, subseção 50C pelas inversões IIL-J e IIL-F e subseção 53A pelas inversões IIL-F e IIL-H (**Figura I.4**). Já para o braço IIR há a ocorrência de reuso na subseção 66A pelas inversões IIR-L e IIR-M (Rohde & Valente 2012).

Por fim, o cromossomo III de *D. willistoni* apresenta reuso de pontos de quebra em nível citológico na subseção 88A pelas inversões III-B e III-P, subseção 92C pelas inversões III-C e III-W e subseção 94A pelas inversões III-G, III-H e III-Z (Rohde & Valente 2012). Sem dúvida, as inferências citológicas acerca do reuso nesta espécie a ressaltam como um organismo modelo altamente promissor neste quesito (Veja **Tópico I.5** desta Tese).

Embora o reuso dos pontos de quebra de inversões em nível citológico seja comum e bastante documentado no gênero *Drosophila*, a caracterização em nível da sequência de DNA ainda é limitada (Richards *et al.*, 2005; Ranz *et al.*, 2007; Calvet *et al.*, 2012; Puerma *et al.*, 2014). Apesar desta defasagem com relação à inferência citológica, mesmo com poucas explorações, já é possível destacar o quão difundido é no gênero *Drosophila* o fenômeno da reutilização dos pontos de quebra de inversões, sejam estas intra ou interespecíficas (Ranz *et al.*, 2007).

Atualmente a ideia preponderante nos diferentes estudos acerca da origem das inversões em *Drosophila* (apesar da grande inferência do reuso dos pontos de quebra), é de que as mesmas são monofiléticas (Wesley & Eanes,

1994; Cáceres *et al.*, 2001; Wallace *et al.*, 2011). Isto se relaciona também às características genômicas das espécies deste gênero, uma vez que estas apresentam, na sua grande maioria, DNA repetitivo de tamanho pequeno e bastante espaçado, o que não favorece tanto fisicamente, como estatisticamente, que estas regiões pertencentes à fração altamente repetitiva do genoma dêem origem, por meio de mecanismos de recombinação ou reparo, à mesma alteração cromossômica em distintos tempos (Sachs *et al.*, 2000).

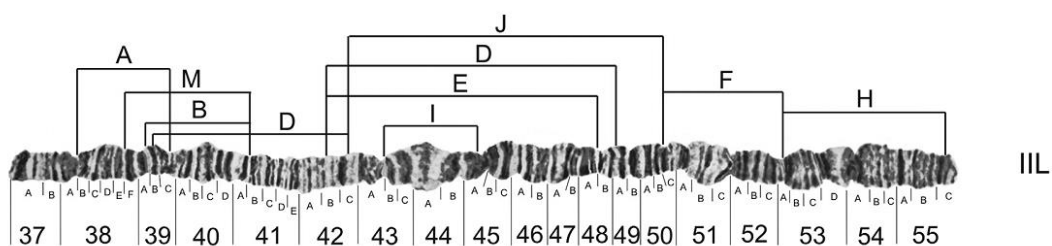


Figura I.4: Delimitação citológica das inversões paracêntricas do braço cromossômico III L de *Drosophila willistoni*. Reuso de pontos de quebra em nível citológico inferidos nas seções 41B, 42B, 50C e 53A. Retirado de Rohde & Valente (2012).

No entanto, casos de reinversão são documentados em outros organismos, destacando-se os mamíferos. Cáceres *et al.* (2007) analisaram em 27 espécies de mamíferos placentários um segmento do cromossomo Xq28, o qual contém os genes *FLNA* e *EMD*, e que está envolvido em uma inversão cromossômica. A comparação genômica apontou que uma duplicação quase idêntica flanqueia esta inversão em humanos e em outras 15 espécies, e sua mais provável origem seria via mecanismo de recombinação *NAHR* (Veja Anexo I, pág. 202) entre estas repetições invertidas ocorrentes no ancestral comum. Estima-se que esta inversão tenha se originado (desde o último ancestral comum, há aproximadamente 100 milhões de anos) pelo menos em 10 ocasiões distintas ao longo da história evolutiva destas espécies. Um ponto

importante é que o tamanho desta inversão recorrente é de aproximadamente 40 kb, o que reforça a característica da proximidade das sequências repetitivas sendo mais favoráveis fisicamente e estatisticamente à fenômenos recorrentes que originam alterações cromossômicas estruturais (Cáceres *et al.*, 2007).

I.4 – Caracterização dos pontos de quebra de inversões em *Drosophila* e mecanismos que originam inversões cromossômicas

A caracterização detalhada dos pontos de quebras das inversões é um passo crucial para a determinação dos mecanismos que as originam. Diferentes estudos apontam que as inversões cromossômicas em *Drosophila* podem ser originadas por dois principais mecanismos.

O primeiro deles é a recombinação ectópica (recombinação homóloga não alélica, *NAHR* - Veja Anexo I, pág. 202) entre sequências repetitivas, destacando-se entre essas os *TEs* (Cáceres *et al.*, 1999; Casals *et al.*, 2003; Delprat *et al.*, 2009). Quando a recombinação ectópica ocorre entre duas cópias de um *TE* inseridas em dois locais cromossômicos diferentes e em orientações opostas, o segmento cromossômico invertido resultante é flanqueado por duas cópias quiméricas do *TE* delimitadas por duplicações do sítio alvo (*TSD*) trocadas (Lim & Simons, 1994; Delprat *et al.*, 2009) (**Figura I.5a**).

O segundo mecanismo é a quebra cromossômica e o reparo das extremidades livres pela União das extremidades não homólogas (*NHEJ* - Veja Anexo I, pág. 204) (Sonoda *et al.*, 2007). Nesse caso, as quebras da fita dupla podem ser simples (do Inglês *Simple double-strand breaks* - *DSB*), ou quebras escalonadas (do Inglês *Staggered single-strand breaks* - *SSB*). No segundo caso, a consequência é a geração de duplicações invertidas em ambos os lados do segmento invertido (Ranz *et al.*, 2007; Papaceit *et al.*, 2012) (**Figura I.5b**). Assim, as inversões geradas desta maneira podem ser reconhecidas por segmentos duplicados de DNA (originalmente cópia única), em orientações opostas flanqueando o segmento cromossômico.

Considerando inversões relativamente recentes na história evolutiva das espécies que as portam, a presença de *TEs* em ambos os pontos de quebra de uma determinada inversão poderia suportar sua origem via recombinação entre essas cópias invertidas; enquanto que a ausência destas sequências, junto com a presença de fragmentos duplicados em um ou ambos os pontos de quebra, poderiam suportar a origem de uma inversão por meio de “quebras escalonadas” (Guillén & Ruiz, 2012).

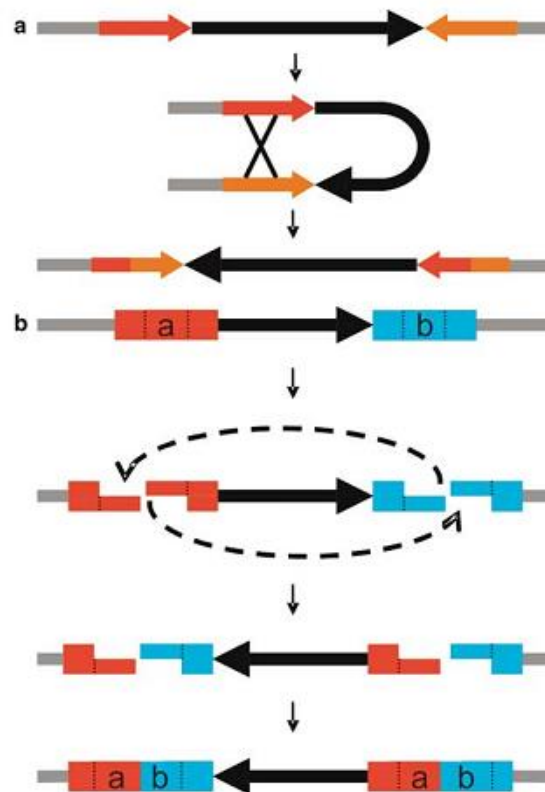


Figura I.5: Representação esquemática dos dois principais mecanismos que originam inversões no gênero *Drosophila*. **a)** Recombinação ectópica entre duas sequências iguais, ou muito similares, em ordem invertida. **b)** Quebras escalonadas e a geração de duplicações invertidas em ambos os lados do segmento. No primeiro caso, a presença de mais de uma cópia de uma determinada sequência (duplicação) é o motivo primário da origem da inversão; no segundo caso, a presença de mais de uma cópia de uma determinada sequência é consequente ao evento de inversão. Retirado de Casals & Navarro (2007).

A contribuição relativa dos dois mecanismos para a gênese das inversões em *Drosophila* ainda não é bastante clara, uma vez que muitas questões acerca da origem das inversões permanecem em aberto, tais como: se estes mecanismos são, ou não, generalizados entre as espécies do gênero; se há implicações funcionais ao longo da evolução cromossômica mantidas por estas inversões, e se há em qual nível ela opera (Ranz *et al.*, 2007).

A **Tabela I.1** apresenta uma compilação dos diferentes estudos que caracterizaram os pontos de quebra de inversões em nível molecular, em diferentes espécies do gênero *Drosophila*. Como se pode constatar, nem sempre as análises dos pontos de quebra permitem inferir a provável origem das inversões, com base nestes dois mecanismos.

Tabela I.1: Estudos de caracterização molecular dos pontos de quebra de inversões em espécies do gênero *Drosophila*.

Espécie	Inversão	Descrição	Referência
<i>D. melanogaster</i>	<i>In(3L)Payne</i>	Análise por microdissecção e sequenciamento da região da inversão no cromossomo. Ausência de sequências repetitivas no ponto de quebra.	Wesley & Eanes (1994)
<i>D. melanogaster</i> x <i>D. subobscura</i>	Inversão fixada no cromossomo X de <i>D. subobscura</i>	Sequências ricas em timinas de aproximadamente 30-50 pb flanqueado os pontos de quebra.	Cirera <i>et al.</i> (1995)
<i>D. melanogaster</i>	<i>In(2L)t</i>	Análise do ponto de quebra proximal, e presença de um <i>TE</i> tipo - <i>LINE</i> .	Andolfatto <i>et al.</i> (1999)
<i>D. buzzatii</i>	Inversão <i>2j</i>	Presença de cópias homólogas de um <i>TE</i> denominado <i>Galileo</i> nos pontos de quebra, e origem da inversão por <i>NAHR</i> entre as cópias invertidas deste <i>TE</i> .	Caceres <i>et al.</i> (1999)
<i>D. buzzatii</i>	Inversão <i>2q⁷</i>	Presença de cópias homólogas do <i>TE Galileo</i> nos pontos de quebra e origem da inversão por <i>NAHR</i> entre as cópias invertidas deste <i>TE</i> .	Casals <i>et al.</i> (2003)
<i>D. pseudoobscura</i>	Inversão <i>Arrowhead</i>	Presença de motivos repetitivos de 128 e 315 pb em orientação oposta nos pontos de quebra da inversão. Origem da inversão por <i>NAHR</i> entre as cópias invertidas destas repetições.	Richards <i>et al.</i> (2005)
<i>D. melanogaster</i>	<i>In(3R)Payne</i>	Pequenas duplicações em ambos os pontos de quebra da inversão.	Matzkin <i>et al.</i> (2005)
<i>D. melanogaster</i> x <i>D. yakuba</i>	29 inversões	59% das inversões apresentavam duplicações invertidas nos pontos de quebra. Origem destas inversões pelo mecanismo de quebras escalonadas.	Ranz <i>et al.</i> (2007)

Tabela I.1: continuação

<i>D. americana</i>	<i>In(4)a</i>	Sequências repetitivas em orientação oposta, de um <i>TE</i> tipo - <i>MITE</i> em ambos os pontos de quebra da inversão.	Evans <i>et al.</i> (2007)
<i>D. mojavensis</i> x <i>D. arizonae</i>	Inversão no cromossomo X	Ausência de sequências repetitivas significativas nos pontos de quebra da inversão.	Cirulli & Noor (2007)
<i>D. pseudoobscura</i> x <i>D. persimilis</i>	Inversões nos cromossomos X e II	Repetições <i>in tandem</i> de um motivo de 319 pb nos pontos de quebra da inversão no braço XR de <i>D. persimilis</i>	Machado <i>et al.</i> (2007)
<i>D. buzzatii</i>	Inversão $2z^3$	Presença de cópias homólogas do <i>TE GalileoN</i> nos pontos de quebra e origem da inversão por <i>NAHR</i> entre as cópias invertidas deste <i>TE</i> .	Delprat <i>et al.</i> (2009)
<i>D. buzzatii</i>	Inversão <i>5g</i>	Ausência de sequências repetitivas significativas nos pontos de quebra.	Prazeres da Costa <i>et al.</i> (2009)
<i>D. mojavensis</i>	Inversão <i>Xe</i>	Ausência de sequências repetitivas significativas nos pontos de quebra. Origem provável por quebras simples.	Runcie & Noor (2009)
<i>D. americana</i> x <i>D. virillis</i>	Inversões <i>(In)Xa</i> e <i>(In)5a</i>	Presença de cópias do <i>TE</i> tipo- <i>MITE DAIBAM</i> nos pontos de quebra das inversões em <i>D. americana</i> . Origem das inversões por <i>NAHR</i> entre as cópias invertidas deste <i>TE</i> .	Fonseca <i>et al.</i> (2012)
<i>D. buzzatii</i>	Inversões <i>2m</i> e <i>2n</i>	Inversão <i>2m</i> com duplicações de 13 kb em ambos os pontos de quebra; origem da inversão por quebras escalonadas.	Calvete <i>et al.</i> (2012)

Tabela I.1: continuação

<i>D. melanogaster</i>	Inversões <i>In(2L)t</i> , <i>In(2R)NS</i> , <i>In(3R)K</i> , <i>In(3R)Mo</i> , <i>In(3R)P</i> , <i>In(3L)P</i> , <i>In(1)A</i> , <i>In(1)Be</i>	Presença de duplicações invertidas nos pontos de quebra das inversões <i>In(2R)NS</i> , <i>In(3R)K</i> , <i>In(3R)P</i> , <i>In(1)A</i> , <i>In(1)Be</i> . Origem das inversões por quebras escalonadas.	Corbett-Detig <i>et al.</i> (2012)
<i>D. mojavensis</i>	Inversões <i>2c</i> , <i>2f</i> , <i>2g</i> , <i>2h</i> , <i>2q</i> , <i>2r</i> e <i>2s</i>	Presença de cópias do <i>TE But-5</i> em ambos os pontos de quebra da inversão <i>2s</i> ; origem da inversão por <i>NAHR</i> entre as cópias invertidas deste <i>TE</i> . Presença de duplicações invertidas nos pontos de quebra das inversões <i>2h</i> e <i>2q</i> ; origem das inversões por quebras escalonadas.	Guillén & Ruiz (2012)
<i>D. pseudoobscura</i>	Inversão 3 do cromossomo O	Sequência de 300 pb em ambos os pontos de quebra; origem da inversão por quebras escalonadas.	Papaceit <i>et al.</i> (2012)
<i>D. subobscura</i>	Inversões E_1 e E_2	Provável origem da inversão E_1 por quebras escalonadas e duplicação do motivo β ; origem da inversão E_2 por <i>NAHR</i> entre motivos α .	Puerma <i>et al.</i> (2014)

I.4.1 - Origem das inversões via o mecanismo de “quebras escalonadas”

Neste mecanismo as inversões são iniciadas por dois pares de quebras (da fita dupla) escalonadas simples, podendo ocorrer entre cromossomos homólogos durante a Prófase meiótica I, ou entre cromátides irmãs (isocromátide) durante as mitoses pré-meióticas. Esta via de origem não exige sequências similares flanqueando os dois pontos de quebra no cromossomo. Ao invés disto, as quebras fisicamente próximas no cromossomo causam falhas no pareamento de bases entre estas quebras, e as regiões cromossômicas se separam (**Figura I.5b**). As extremidades finais salientes geradas por este processo podem ser reparadas de duas maneiras: ou com a perda dos nucleotídeos salientes na extremidade 5', ou com o preenchimento da região faltante seguido de *Junção das extremidades não homólogas (NHEJ* - Veja Anexo I, pág. 205). Este último origina inversões quando não estabelece a junção da extremidade 5' com a sua própria extremidade 3', mas sim com a extremidade 3' do outro ponto de quebra. As inversões, geradas então, desta maneira podem deixar sinais evidentes deste processo na sequência genômica do descendente. O reparo de ambas as extremidades 5' salientes do mesmo ponto de quebra de maneira idêntica, resulta ou em duplicações invertidas flanqueando os pontos de quebra, ou em deleções recíprocas destas regiões no cromossomo do descendente. Já, se as duas extremidades 5' salientes de um ponto de quebra são reparadas de maneiras opostas, então, muito provavelmente só se encontrará uma única duplicação (ou deleção) em um dos pontos de quebra, ou nenhum sinal evidente (Ranz *et al.*, 2007; Runcie & Noor, 2009).

Drosophila melanogaster, *D. simulans* e *D. yakuba* são membros do subgrupo da *D. melanogaster*. O estudo de Ranz *et al.* (2007) por meio de métodos experimentais e computacionais analisou os pontos de quebra de inversões interespecíficas nestas espécies.

Cinco inversões paracêntricas são conhecidas por diferenciarem o padrão de bandas dos cromossomos politênicos de *D. melanogaster* e *D.*

simulans (Horton, 1939). Já *D. yakuba*, por sua vez, possui 28 inversões paracêntricas diferenciando seus cromossomos dos de *D. melanogaster*.

A análise dos pontos de quebra da inversão *In(3R)84F1;93F6-7*, a qual diferencia o braço cromossômico 3R de *D. melanogaster* do de *D. simulans*, evidenciou a presença de três sequências (*HDC14862*, *pdf800* e *HDC12400*) duplicadas e em orientação oposta, em ambos os pontos de quebra da inversão em *D. melanogaster*, sendo confirmado também por hibridação *in situ*. Estas sequências se apresentam como cópias únicas em *D. simulans* e em *D. yakuba*, indicando que a inversão *In(3R)84F1;93F6-7* é um estado derivado com relação ao arranjo cromossômico padrão destas espécies (Ranz *et al.*, 2007).

Tomando o cromossomo 3 de *D. simulans* como portador de um arranjo ancestral, comparações em nível molecular do braço cromossômico 3R desta espécie, com o de *D. yakuba*, evidenciaram uma inversão fixada na última espécie (*In3R(7)*), a qual reutiliza um dos pontos de quebra da inversão *In(3R)84F1;93F6-7*. Em ambos os pontos de quebra dessa inversão havia a presença de duas sequências (*CG34034* e *CG31286*) duplicadas, e em orientação oposta (Ranz *et al.*, 2007)

Devido à presença de duplicações invertidas associadas aos pontos de quebra destas inversões, Ranz *et al.* (2007) propuseram como mais parcimonioso o mecanismo de quebras escalonadas como sendo o responsável pela origem das inversões *In(3R)84F1;93F6-7* e *In3R(7)*. Estas quebras escalonadas podem envolver isocromátides, ocorrendo durante a prófase mitótica e denominadas “quebras escalonadas de fita simples” (do Inglês *Staggered Single-Strand Breaks*); ou envolver cromátides irmãs, ocorrendo durante a prófase meiótica, denominadas “quebras escalonadas de fita dupla” (do Inglês *Staggered Double-Strand Breaks*).

O mesmo estudo analisou *in silico* os pontos de quebra das 28 inversões paracêntricas que diferenciam os cromossomos de *D. melanogaster* dos de *D. yakuba*, bem como uma inversão pericêntrica no cromossomo 2. As evidências genômicas e filogenéticas sugerem que virtualmente todas as inversões se originaram na linhagem de *D. yakuba*. A análise dos pontos de quebra das

inversões mostrou que em aproximadamente 62% dos casos (18 das 29 inversões), ocorria a presença de sequências duplicadas, as quais estavam presentes com apenas uma cópia no genoma de *D. melanogaster*. Em seis destas inversões as duplicações eram em ambos os pontos de quebras, em 12 estavam presentes em apenas um dos pontos de quebras, o que pode ser devido ao modo de reparo, ou outras modificações ocorridas na sequência. Nenhuma destas duplicações correspondia a sequências de *TEs* já classificados. Fica claro nesta análise que a maioria das inversões que diferenciam os cromossomos de *D. melanogaster* dos de *D. yakuba*, se originaram por “quebras escalonadas” na última espécie (Ranz *et al.*, 2007).

Outra análise envolvendo este mecanismo foi feita para a inversão 3 do cromossomo *O* de *D. subobscura*. Esta abordagem também permitiu a reconstrução dos pontos de quebra (denominados *AB* e *CD*) do extinto arranjo cromossômico ancestral O_3 (Elemento de Muller E), por meio de uma estratégia que “mixou” análises *in silico* e citomoleculares (Papaceit *et al.*, 2012). O arranjo O_3 deu origem a duas inversões nesta espécie: inversão 4, a qual deu origem ao arranjo cromossômico O_{3+4} (pontos de quebra denominados *AB* e *DC*, o que possibilitou a reconstrução dos pontos de quebra do arranjo O_3 , por apresentar a mesma ordem do ponto de quebra proximal (*AB*), e uma pequena inversão da ordem do ponto de quebra distal (*CD*); e inversão 3, a qual deu origem ao arranjo cromossômico O_{st} (pontos de quebra denominados *AC* e *BD*). Comparações entre os pontos de quebra das inversões 4 e 3, presentes nos arranjos O_{3+4} e O_{st} respectivamente, apontaram a presença de repetições de aproximadamente 300 pb nos pontos de quebra da inversão 3, fato que claramente dá suporte à origem desta inversão por meio de “quebras escalonadas na dupla fita” (Papaceit *et al.*, 2012).

A espécie *D. subobscura* também teve os pontos de quebras das inversões E_1 e E_2 (Elemento de Muller C) analisados por Puerma *et al.* (2014). Estas duas inversões compartilham citologicamente um dos pontos de quebra. As comparações dos pontos de quebra da linhagem padrão *Est* (pontos de quebra *AB*, *EF*, *GH*) com relação aos pontos de quebra da linhagem com as inversões E_{1+2} (pontos de quebra *AG*, *FB*, *EH*) apontaram a presença de dois motivos, denominados α e β , os quais compartilham a porção terminal

denominada δ , em orientações opostas. O motivo α estava nos pontos de quebra AB e AG , com a mesma orientação, mas com orientação invertida no ponto de quebra GH (duas cópias em orientação invertida no cromossomo *Est*, e uma cópia no cromossomo E_{1+2}). O motivo β estava presente nos pontos de quebra EF e EH com a mesma orientação, e com a orientação invertida no ponto FB (uma cópia no cromossomo *Est*, e duas cópias no cromossomo E_{1+2}). O motivo α exibe pequenos fragmentos com similaridades com o elemento *SGM*, enquanto que o motivo β não apresenta similaridade com nenhum *TE* descrito. Em vista deste cenário, a provável origem da inversão E_1 foi inferida por quebras escalonadas, que levaram à duplicação do motivo β presentes nos pontos de quebra FB e EH . Já a origem da inversão E_2 inferiu-se como devida à recombinação ectópica entre os dois motivos α , presentes em ambos os pontos de quebra AB e GH . O reuso foi inferido pela presença de repetições de 400-700 pb nos pontos de quebra, entretanto a direção evolutiva dos eventos de inversões não ficou esclarecida (Puerma *et al.*, 2014).

1.4.2 - Envolvimento dos elementos transponíveis na origem das inversões: Recombinação ectópica.

Elementos transponíveis (*TEs* – do Inglês *Transposable Elements*) são componentes interessantes e dominantes da maioria dos genomas eucariotos, o que significa que a compreensão da sua biologia é um assunto fundamental na genética. Desde a sua descoberta por McClintock (1946; 1947), muito tem se aprendido sobre as propriedades moleculares dos *TEs*. Estes elementos são classificados de acordo com seus modos de transposição e características.

Elementos de Classe I, também chamados de retrotransposons, replicam por meio de um mecanismo de "copiar e colar", que envolve a produção de um intermediário de RNAm, o qual é processado por transcrição reversa para DNA e inserido novamente no genoma (Eickbush, 2002; Sandmeyer *et al.*, 2002). Os retrotransposons se subdividem em retrotransposons com *LTRs* (do Inglês *Long Terminal Repeats*), como por

exemplo, os *elementos copia* e *Gypsy*, que são semelhantes à retrovírus; e os retrotransposons sem *LTRs*, tais como *LINEs* (do Inglês *Long Interspersed Elements*) e *SINEs* (do Inglês *Short Interspersed Nuclear Elements*), que são chamados de retroposons.

Elementos de Classe II, ou transposons de DNA, se replicam, genericamente por um mecanismo de "cortar e colar", no qual os elementos são fisicamente excisados do genoma e reinseridos em outro local. Nesse caso, um aumento no número de cópias ocorre durante a reparação dos sítios de excisão do transposon de DNA pelo hospedeiro durante a síntese DNA, ou pela inserção do *TE* em um local diferente do genoma que ainda vai ser replicado (Finnegan, 1989; Capy *et al.*, 1998).

Ainda dentro da Classe II de *TEs*, há um grupo de elementos não autônomos denominados *MITEs* (do Inglês *Miniature Inverted-repeat Transposable Elements*). Esses elementos são sequências curtas, presentes normalmente em alto número de cópias no genoma, e com nenhuma capacidade codificadora; conforme sugerido para o elemento *Mar*, cuja distribuição parece ser restrita ao subgrupo *willistoni* de *Drosophila* (Deprá *et al.*, 2012).

Os *TEs* de ambas as classes são ainda classificados em Classe, Subclasse, Ordem, Superfamília, Família e subfamília com base no seu compartilhamento de determinadas estruturas e semelhanças de sequências (Wicker *et al.*, 2007).

No gênero *Drosophila*, os estudos relacionados à associação dos *TEs* com pontos de quebra de rearranjos cromossômicos começaram principalmente com as análises em linhagens apresentando disgenesia híbrida, as quais se caracterizam por alta frequência de prole inviável, recombinação em machos, mutações e aberrações cromossômicas (Kidwell, 1977). A causa da disgenesia híbrida tem sido reportada à ativação de diversas famílias de *TEs*, incluindo elementos *P*, *I* e *hobo* em *D. melanogaster* (Kidwell, 2005) e *Penelope*, *Ulysses*, *Helena*, *Paris* e *Telemac* em *D. virilis* (Evgen'ev *et al.*, 1997). Estes estudos, por meio de cruzamentos programados e análises citogenéticas da prole acompanhavam a movimentação dos *TEs* envolvidos e o

surgimento de rearranjos cromossômicos associados a esta movimentação (Engels & Preston, 1984; Lim, 1988).

Desde então, diferentes estudos têm mostrado a associação de inserções de *TEs* com pontos de quebra de inversões em populações naturais de *Drosophila*. Dentre esses, destacam-se os estudos de Lyttle & Haymer (1992) analisando o transposon *hobo* em *D. melanogaster*, e os de Zelentsova *et al.* (1999) e Evgen'ev *et al.* (2000) analisando os retroelementos *Penelope* e *Ulysses* em espécies do grupo da *D. virilis*.

Neste contexto, o trabalho de Regner *et al.* (1996), do nosso grupo de pesquisa, foi um dos pioneiros nas análises de populações naturais. Neste estudo, os autores analisaram a coincidência dos sítios de inserção do elemento *P* com pontos de quebra de inversões em *D. willistoni*, encontrando associação deste elemento com 10 pontos de quebra de inversões paracêntricas segregantes na espécie.

De acordo com Gray *et al.* (1996), Gray (2000), Zhang *et al.* (2009) e Delprat *et al.* (2009), há dois possíveis mecanismos pelos quais rearranjos cromossômicos associados a *TEs* podem ocorrer: diretamente por um processo de transposição alternativa, ou indiretamente por recombinação ectópica.

Os *TEs* induzem rearranjos cromossômicos diretamente por uma versão alternativa do evento de transposição. Neste caso, as regiões finais dos *TEs* envolvidos são provenientes de elementos homólogos separados (dando origem a um elemento híbrido), ao invés de pertencerem à mesma cópia de um elemento, como na transposição tradicional. Este mecanismo também é conhecido como transposição aberrante, não convencional ou alternativa (Veja Anexo I, pág. 207); e até o presente momento tem sido descrito para os elementos de Classe II. Sendo assim, dois *TEs* podem participar de um evento de transposição alternativa, onde a região final de *TEs* separados formam sinapses, devido à complementaridade de suas sequências. Com isto, o elemento híbrido formado se transpõe para um novo sítio cromossômico alvo, o qual pode estar até mesmo dentro de outro *TE*. Os tipos de rearranjos resultantes deste processo dependem do tipo de molécula de DNA envolvida

(circular ou linear), da localização do sítio alvo e da orientação dos *TEs* envolvidos nos cromossomos parentais. Já a visualização do resultado do evento de transposição alternativa depende da viabilidade do organismo e de sua capacidade de suportar alterações cromossômicas estruturais. Este mecanismo pode explicar os diferentes tipos de rearranjos cromossômicos que não se originam por recombinação ectópica (Delprat *et al.*, 2009).

O processo de origem de rearranjos cromossômicos por transposição alternativa foi apresentado em *Drosophila*, por meio de experimentos com o elemento *P* biotecnologicamente modificado por deleções *in situ* de suas extremidades. A combinação de elementos com as mesmas extremidades deletadas (elementos em *cis*) resultaram em baixas taxas de recombinação, enquanto que a combinação de elementos com as extremidades opostas deletadas (elementos em *trans*) resultou em altas frequências de recombinação (Svoboda *et al.*, 1995; Gray *et al.*, 1996). A transposição alternativa é um fenômeno muito bem embasado em milho, onde o sistema *Ac/Ds*, descrito por Barbara McClintock (1941), atua via este mecanismo na origem de diversos rearranjos cromossômicos estruturais associados a alterações fenotípicas.

A ação indireta dos *TEs* promove rearranjos cromossômicos mediante a presença no genoma de múltiplas cópias semelhantes, se não idênticas, entre as quais podem ocorrer recombinações. Estas trocas podem envolver recombinação de sequências homólogas, ou reparo defeituoso das quebras em fitas duplas formadas durante a excisão do elemento transponível. Após um evento de recombinação ectópica entre dois *TEs* é esperado que duas cópias quiméricas destes *TEs* flanqueassem a inversão, bem como suas *TSDs* (do Inglês - *Target Site Duplication*) estejam trocadas (Delprat *et al.*, 2009).

A primeira análise que evidenciou diretamente o envolvimento de um *TE* na origem de uma inversão, em uma população natural de *Drosophila*, foi a de Cáceres *et al.* (1999). Este estudo caracterizou molecularmente os pontos de quebra da inversão polimórfica *2j* da espécie *D. buzzatii*, a qual se originou do arranjo cromossômico *2st* (padrão). Para a análise dos pontos de quebra da inversão *2j*, estes foram delimitados por caminhada cromossômica, clonados e sequenciados. Primeiramente, a região proximal do ponto de quebra do cromossomo (denominada *CD*) foi obtida de uma linhagem homocigota para o

arranjo $2st$ ($st-1$). Após, os pontos de quebra distal e proximal no cromossomo $2j$ (AC e BD respectivamente) foram isolados de uma linhagem homocigota para o arranjo $2j$ ($j-1$). Com isto, a região distal do ponto de quebra do cromossomo $2st$ (AB) foi obtida da linhagem $st-1$. O alinhamento das sequências AB e CD em $st-1$ (perfazendo 2.086 pb), com as sequências AC e BD em $2j$ (com 8.668 pb), mostrou que haviam inserções presentes em ambos os pontos de quebra da inversão $2j$ que não estavam presentes no arranjo $2st$. A inserção entre A e C possuía 392 pb com longas *ITRs* (do Inglês - *Inverted Repeats Terminals*) de 106 pb. A inserção entre B e D possuía 4.319 pb, com uma estrutura interna complexa, incluindo *ITRs* como aquelas da inserção AC de 106 e 47 pb; esta última correspondendo ao começo de uma repetição da *ITR* de 106 pb. Os 180 pb centrais da inserção AC e da sequência BD que segue a *ITR* proximal, apresentaram 95% de homologia; porém, estão em direções opostas. As sequências flangeadoras de cada uma destas inserções possuem 7 pb, e assemelham-se a sítios alvos duplicados que foram produzidos durante um evento de inserção. Entretanto, os sítios alvos duplicados são separados e invertidos, o que sugere fortemente que a inversão $2j$ se originou via pareamento intracromossomal e recombinação entre as duas sequências homólogas inseridas em sítios distantes e em orientações opostas. Segundo os autores, pode-se considerar que estas inserções eram homólogas em aproximadamente 274 pb, e mesmo essa homologia sendo relativamente pequena, parece ter tamanho suficiente para sustentar uma recombinação ectópica em *Drosophila*. A duplicação da sequência de 7 pb após cada inserção, a presença de *ITRs*, e o padrão de hibridação *in situ* moderadamente repetitivo identificam as inserções dos pontos de quebra da inversão $2j$ como cópias de um *TE* de Classe II, o qual foi denominado *Galileo*.

O elemento *Galileo* é um transposon pertencente à superfamília *P* de *TEs*, e está presente em seis dos 12 genomas de *Drosophila* sequenciados. Dentre estas espécies *D. willistoni* se ressaltou por ter inferida a maior quantidade de cópias defectivas deste elemento no seu genoma (Marzo *et al.*, 2008).

Subsequentemente, o elemento *Galileo* e outros membros de sua família (*GalileoN* e *GalileoK*, Delprat *et al.*, 2009) tiveram seu envolvimento apontado

na origem de mais duas inversões polimórficas do cromossomo 2 de *D. buzzatii*: $2q^7$ (Casals *et al.*, 2003) e $2z^3$ (Delprat *et al.*, 2009). Estes trabalhos mostraram, por meio de análises citológicas, moleculares e *in silico*, que a origem destas inversões deveu-se à ocorrência de recombinação ectópica entre duas cópias do *TE Galileo*, presentes nos pontos de quebra dessas inversões.

O cromossomo *neo-X* de *D. americana* (do grupo *virilis*), o qual é derivado de uma fusão centromérica segregante entre o cromossomo X (elemento de Muller A) e o cromossomo 4 (Elemento de Muller B) nesta espécie, está relacionado à latitude da distribuição geográfica da espécie (McAllister, 2002). A análise da inversão 4a deste *neo-X* por Evans *et al.* (2007) apontou a provável origem desta inversão por meio de recombinação ectópica entre duas cópias de um elemento repetitivo do tipo-*MITE* disperso no genoma, as quais não estavam presentes na região correspondente da linhagem da espécie sem a inversão. As características desta sequência repetitiva que suportam sua identidade como *TE* é a presença de *TIRs* de 240 pb flanqueando uma região interna de 869 pb. Em comparação com as outras cópias presentes no genoma, a cópia presente no ponto de quebra proximal aparenta ser um elemento canônico, enquanto que a cópia presente no ponto de quebra distal trata-se de um elemento rearranjado.

Posteriormente, o sequenciamento com baixa cobertura do genoma de duas linhagens de *D. americana*, permitiu a comparação com *D. virilis* (outra espécie pertencente ao grupo *virilis*). A análise da inversão *Xa* fixada em *D. americana* e ausente em *D. virilis*, e da inversão polimórfica 5a, por Fonseca *et al.* (2012) alinhou os pontos de quebra de ambas as inversões entre as duas espécies. Nas regiões onde o alinhamento era corrompido, ocorria sempre a presença de uma sequência variando entre 500-1130 pb, presentes somente nas linhagens portando as inversões. Estas sequências mostraram por BLASTN alta similaridade com uma sequência do tipo-*MITE* incompleta, com *TIRs* de 240 pb. Neste estudo, os autores o denominaram de *DAIBAM* (*Drosophila americana Inversion Breakpoints Associated MITE*). Na inversão *Xa* foi possível encontrar claras *TSDs* e cópias defectivas do *TE DAIBAM* flanqueando a inversão. Já na inversão 5a, as cópias do elemento *DAIBAM*

flanqueando a inversão apresentavam mais de 70% de similaridade nucleotídica. Considerando que as cópias do *TE DAIBAM* são defectivas e que as inversões analisadas são antigas, os autores inferem que os dados encontrados suportam a origem das inversões *Xa* e *5a* por recombinação ectópica entre os elementos *DAIBAM* presentes nos pontos de quebra destas inversões. Foi também constatado que este elemento era o mesmo que estava presente nos pontos de quebra da inversão *4a*, analisada por Evans *et al.* (2007). Sendo assim, o elemento *DAIBAM* é o responsável pela origem de pelo menos 20% das inversões ocorrentes no grupo *virilis* (Fonseca *et al.*, 2012).

A análise citológica dos cromossomos politênicos da espécie *Drosophila mojavensis* apresenta 10 inversões fixadas com relação ao arranjo Primirivo I do grupo *repleta*; uma no cromossomo X (*Xe*), sete no cromossomo 2 (*2c*, *2f*, *2g*, *2h*, *2q*, *2r* e *2s*) e duas no cromossomo 3 (*3a* e *3d*). Guillén & Ruiz (2012) caracterizaram molecularmente os pontos de quebra das sete inversões do cromossomo 2 desta espécie, comparando o mapa físico da biblioteca genômica de *D. buzzatii* com o genoma de *D. mojavensis* pela técnica *BAC-end sequences*. Comparações com o genoma de *D. virilis* (um grupo externo sem as inversões) também foram feitas. A partir disto, os pontos de quebra de todas as 10 inversões de *D. mojavensis* foram delimitados. Os pontos de quebra da inversão *2s* apresentaram cópias do transposon *BuT-5*, sendo que a cópia do ponto de quebra distal possuía 981 pb, delimitada por sequências de 9 pb AAGGCAAGT e CTGTATAAT; e no ponto de quebra proximal a cópia do *TE BuT-5* tinha 27 pb, sendo que 12 pb eram idênticos à uma das *TIRs* e os demais 15 pb eram idênticos à outra *TIR*, delimitada por sequências de 9 pb ACTTGCCTT e ATTATACAG. As sequências ACTTGCCTT e CTGTATAAT são as sequências complementares invertidas de AAGGCAAGT e ATTATACAG, respectivamente e estas características apontam que a origem da inversão *2s* em de *D. mojavensis* teve sua origem devido à recombinação ectópica entre as duas cópias do *TE BuT-5* (Guillén & Ruiz, 2012).

Apesar das tentativas de análise dos pontos de quebra de inversões em diferentes espécies (**Tabela I.1**), este campo ainda permanece escasso de caracterização. Muitas vezes a implicação do mecanismo envolvido na origem das inversões é ambígua (Richards *et al.*, 2005; Calvete *et al.*, 2012; Puerma *et*

al., 2014) e com isto, nem sempre é possível determinar a provável gênese destes polimorfismos.

Neste contexto, novamente a *D. willistoni*, com seu amplo polimorfismo cromossômico e disponibilidade de um genoma sequenciado, pode ser muito informativa acerca dos mecanismos envolvidos na gênese dos rearranjos cromossômicos em eucariotos.

1.5 - *Drosophila willistoni* como organismo modelo para os estudos de inversões cromossômicas

Drosophila willistoni pertence ao grupo *willistoni* do subgênero *Sophophora*, o qual é constituído por 24 espécies (Bächli, 2015), apresentando distribuição geográfica essencialmente Neotropical e sendo subdivididas em três subgrupos: o subgrupo *willistoni* de *Drosophila*, que compreende seis espécies crípticas (morfologicamente muito semelhantes): *D. willistoni*, *D. paulistorum* (constituída por seis semi-espécies), *D. tropicalis*, *D. equinoxialis*, *D. pavlovskiana* e *D. insularis* (provável ancestral que deu origem a todo o subgrupo (Rohde *et al.*, 2006; Robe *et al.*, 2010); subgrupo *bocainensis*, com 12 espécies e subgrupo *alagitans* com seis espécies (Bächli, 2015). A *D. willistoni* tem a maior distribuição dentre as espécies desse grupo, sendo encontrada desde a Flórida e México, até o sul do Brasil, Uruguai e Argentina; do Oceano Atlântico ao Pacífico, exceto em áreas de grande altitude e desertos (Spassky *et al.*, 1971; Ehrmann & Powell, 1982).

O cariótipo básico de *D. willistoni*, o qual se estende para as demais espécies do grupo, é constituído por dois pares de cromossomos metacêntricos (o par sexual e o segundo cromossomo) e um par acrocêntrico (terceiro cromossomo) (Dobzhansky, 1950). A nomenclatura dos cromossomos politênicos correspondentes, definidos por Dobzhansky (1950) é: XL e XR (braço direito e esquerdo do cromossomo X), IIL e IIR (braço esquerdo e direito do segundo cromossomo) e III (o cromossomo acrocêntrico).

Muller (1940), por meio de análises de mapas de ligação de distintas espécies, observou que os marcadores gênicos tendiam a se agrupar no mesmo cromossomo, ressaltando uma característica da organização genômica das espécies do gênero *Drosophila*: a de que o conteúdo gênico nos cromossomos entre estas espécies é altamente conservado, e a reorganização espécie-específica deve-se às inversões paracêntricas e eventos de fusões e fissões que ocorreram ao longo da história evolutiva das diferentes espécies (Bhutkar *et al.*, 2008). De acordo com isto, o complemento cromossômico das espécies do gênero *Drosophila* são organizados em Elementos de Muller que se denominam de A-E, conforme seus respectivos conteúdos gênicos, tomando como base o conteúdo gênico dos cromossomos da espécie *D. melanogaster*, a qual possui $n=4$; sendo que o cromossomo X é acrocêntrico (Elemento A), o cromossomo II é metacêntrico (Elementos B e C), o cromossomo III também metacêntrico (Elementos D e E), e o cromossomo pontual IV (Elemento F).

Em conformidade com esta designação, os cromossomos de *D. willistoni* vêm sendo tradicionalmente classificados como: braços cromossômicos XL e XR, Elementos de Muller A e D respectivamente; braços cromossômicos IIL e IIR, Elementos de Muller C e B respectivamente (**Capítulo III, Figura III.1**); e cromossomo III, Elementos de Muller E+F.

Do ponto de vista cromossômico, *D. willistoni* é claramente a espécie mais polimórfica do grupo e também uma das mais polimórficas do gênero *Drosophila* (revisões em Ehrman & Powell, 1982; Valente & Morales, 1985; Valente & Araújo, 1985, 1986; Valente *et al.*, 2001; Rohde & Valente 2012). Este polimorfismo diz respeito às inversões do tipo paracêntricas, embora um caso de inversão pericêntrica muito frequente tenha sido registrado em populações de *D. willistoni* no continente e ilhas de Santa Catarina, no sul do Brasil por Rohde *et al.* (2005).

A **Tabela Anexo 2** (pág. 217) apresenta um sumário histórico dos estudos realizados sobre os cromossomos e o polimorfismo cromossômico de *D. willistoni* e suas espécies crípticas. Claramente esta espécie resalta-se na quantidade de estudos e dados disponíveis. Entretanto, a descrição do polimorfismo cromossômico desta espécie se apresentava geralmente bastante

confusa e imprecisa, sendo que uma tentativa de sistematização foi feita por Rohde & Valente (2012).

I.5.1 - *Drosophila willistoni* na Era Genômica

O primeiro genoma disponível de uma espécie do gênero *Drosophila* foi o da *D. melanogaster*, há 15 anos (Adams *et al.*, 2000). Cinco anos depois *D. pseudoobscura* teve seu genoma sequenciado (Richards *et al.*, 2005). Subsequentemente, um Consórcio multinacional foi estabelecido para sequenciar o genoma de mais 10 espécies de *Drosophila*. A única espécie Neotropical escolhida foi a *D. willistoni*, e para tal, nosso grupo de pesquisa do Laboratório de *Drosophila*, do Departamento de Genética da UFRGS, foi convidado a participar em função da longa trajetória de conhecimento acerca desse organismo. O resultado deste enorme esforço foi a publicação "*Evolution of genes and genomes on the Drosophila phylogeny*" (*Drosophila* 12 Genomes Consortium, 2007).

Atualmente, cerca de 25 espécies de *Drosophila* já tiveram seus genomas sequenciados, embora alguns não estejam totalmente montados e com cobertura satisfatória; e apesar das similaridades entre as diferentes espécies, muitas peculiaridades e diferenças têm sido encontradas. Os estudos comparativos desses genomas estão permitindo um avanço sem precedentes no entendimento da evolução biológica, e das diferenças subjacentes às variáveis de comportamento e ecologia de organismos relacionados. Entre esses estudos, *D. willistoni* vem se destacando com diversas peculiaridades, ressaltando-se entre estas a inferência do menor nível de uso preferencial de códons (Vicario *et al.*, 2007), a virtual ausência de selenoproteínas (Chapple & Guigó, 2008) a segunda maior proporção de *TEs* de diferentes famílias (*Drosophila* 12 Genomes Consortium, 2007), o maior conteúdo de regiões potencialmente codificantes entre as demais espécies (65,4 Mb), a maior proporção de íntrons (23,5 Mb) e o maior conteúdo de tRNA, com 484 genes de tRNA anotados (*Drosophila* 12 Genomes Consortium, 2007).

Desde então, novas liberações (ou "releases") de vários desses genomas têm sido feitas, destacando-se as de *D. melanogaster*, para a qual já se tem dados de transcriptoma e anotação de *TEs*; e também grandes avanços na caracterização do genoma de *D. pseudoobscura*. Neste particular, o genoma da *Drosophila willistoni*, o último a ser sequenciado, ainda se apresenta sem muitos aprimoramentos desde seu sequenciamento inicial, sendo que desde 2007 até o presente momento, somente uma nova inferência de genes foi estabelecida em Fevereiro de 2015. Esta recente atualização se baseou na predição gênica desenvolvida pelo banco de dados *NCBI* (Geer *et al.*, 2010), a qual utiliza uma combinação de buscas por sequências homólogas e uma predição inicial (*ab initio*) (Thibaud-Nissen *et al.*, 2013); a inferência gênica final é estabelecida pelo programa *Gnomon* (Suvorov *et al.*, 2010). Com isso, uma ampla revisão das coordenadas dos diversos genes foi estabelecida, bem como alguns tiveram suas anotações retiradas, e outros foram inseridos.

Uma dificuldade inerente ao processo de sequenciamento de *D. willistoni* deve-se ao seu extenso polimorfismo cromossômico para inversões (revisões em Ehrman & Powell, 1982; Rohde & Valente, 2012), que ocorre em populações naturais, ao longo da sua ampla distribuição geográfica (do sul dos Estados Unidos até o norte da Argentina - Spassky *et al.*, 1971). Assim, a linhagem de *D. willistoni* primeiramente sequenciada foi a Gd-H4-1, procedente da Ilha Caribenha de Guadalupe; e resultado de várias gerações de cruzamentos irmã-irmão, a fim de obter uma linhagem monocariotípica (sem inversões segregantes). Este endocruzamento, produzido artificialmente, comprometeu a estabilidade da linhagem, já que a espécie é naturalmente polimórfica. Além disso, a atual linhagem cujo arranjo cromossômico foi considerado padrão ("*standard*") para a espécie (Regner *et al.*, 1996), e cujo fotomapa foi aperfeiçoado posteriormente por nosso grupo de pesquisa (Schaeffer *et al.*, (2008; Rohde & Valente, 2012), foi a estirpe Wip-4, coletada pelos pesquisadores A.R. Cordeiro e Helga Winge na cidade de Ibirapitanga (Bahia, Brasil) na década de 1960. Essa linhagem, de fácil manutenção em laboratório, apresenta o padrão de bandas dos cromossomos politênicos similar ao do mapa de referência de Dobzhansky (1950), desenhado em

câmara clara, com base em uma linhagem de Belém (Pará, Brasil) que não se encontra mais disponível.

Sendo a característica que mais se destaca em *D. willistoni* a ocorrência de elevado polimorfismo para inversões paracêntricas, esta espécie se torna um intrigante modelo biológico para diferentes pesquisas em genética evolutiva e de populações, especiação, ecologia e evolução molecular. A disponibilização do genoma desta espécie traz novos desafios e motivações para o prosseguimento do estudo do polimorfismo cromossômico neste organismo-modelo tão peculiar; agora com um enfoque na gênese de suas inversões cromossômicas, com o uso de modernas ferramentas da Genômica que se somam ao sólido conhecimento acerca de sua Citogenética.

OBJETIVOS

- Objetivo geral

Contribuir para o conhecimento dos mecanismos geradores de rearranjos cromossômicos em eucariotos, em especial do organismo modelo *Drosophila willistoni*, por meio de análises citogenômicas.

- Objetivos específicos

- Estabelecer a precisa diferenciação do arranjo cromossômico do braço IIL entre o fotomapa e a linhagem sequenciada Gd-H4-1 de *Drosophila willistoni*, a fim de que se possa estabelecer a linhagem Gd-H4-1 como a padrão para as análises dos pontos de quebra nesta espécie (**Capítulo III, Tópicos III.1 e III.2**).

- Aumentar o mapeamento físico de marcadores genéticos no genoma de *Drosophila willistoni*, estabelecendo sua posição nos cromossomos politênicos da espécie. Esta abordagem pretende aprimorar a primeira montagem do genoma de *Drosophila willistoni*, determinando com maior precisão a ordem e a orientação dos *scaffolds* estabelecidos (**Capítulo III, Tópico III.3 e Capítulo V**).

- Estabelecer no genoma da linhagem sequenciada Gd-H4-1, e na linhagem SG12.00 de *Drosophila willistoni*, os genes e outras sequências que delimitam os pontos de quebra da inversão IIL-H (no cromossomo II), a fim de que se possa posteriormente inferir a possível origem desta inversão por meio de uma análise comparativa dos pontos de quebra em ambas as linhagens, bem como a influência desta inversão na estrutura e função das demais sequências flanquadoras (**Capítulo III, Tópico III.4**).

CAPÍTULO II

II - Materiais e Métodos

II.1 - Populações utilizadas e manutenção dos estoques de *Drosophila willistoni*

As populações estoque de *D. willistoni* foram mantidas como culturas massais, em meio de cultura padrão estabelecido por Marques *et al.* (1966), dentro de câmaras de temperatura e umidade constante (a 18 °C ou a 22 °C \pm 1 grau e 60% de umidade relativa); no Laboratório de *Drosophila* da UFRGS. A manutenção dos estoques foi feita semanalmente.

As linhagens de *D. willistoni* utilizadas no presente estudo são descritas na **Tabela II.1**.

Tabela II.1: Local e ano de coleta das linhagens de *Drosophila willistoni* analisadas no presente estudo.

Linhagem	Local	Ano de Coleta
Gd-H4-1	Guadalupe, Caribe	1991
SG12.00	Montevideo, Uruguai	2000

A linhagem Gd-H4-1 de *D. willistoni* teve seu genoma totalmente sequenciado em 2006/2007. Dada a natureza altamente polimórfica para inversões cromossômicas da espécie (Rohde & Valente, 2012), o sequenciamento exigiu a padronização monomórfica desta linhagem por meio de endocruzamentos. Sendo assim, esta linhagem passou por nove gerações

de endocruzamentos programados, e com isso, virtualmente obteve-se uma linhagem sem a presença de inversões segregantes, e portando um arranjo específico fixado para cada cromossomo (*Drosophila* 12 Genomes Consortiun, 2007).

A linhagem SG12.00 ($cd^{SG12.00}$), com fenótipo mutante cardinal para a cor dos olhos (uma mutação recessiva - Soler & Goñi, 2012), foi estabelecida a partir de alguns indivíduos que emergiram dentro da primeira ou segunda geração da linhagem isofêmea SG2.00 (a qual foi estabelecida a partir de uma fêmea individual inseminada), coletada no ano 2000 na Facultad de Agronomía, Montevideo (Uruguai). Esta estirpe mutante tem sido mantida em condições de laboratório para análises de alelismo, desequilíbrio de ligação gênica (Soler & Goñi, 2012) e análises citogenéticas (Goñi & Valente, 2012).

A linhagem SG12.00 apresenta a inversão IIL-H em condição homozigótica (ver Tabela 1 e Figura 1E em Goñi & Valente, 2012), com relação ao fotomapa da espécie (Schaeffer *et al.*, 2008). Essa linhagem apresenta a sequência padrão nos braços XL, XR e 2R, e segrega três inversões paracêntricas no cromossomo III (III-J, III-B e III-A, Goñi & Valente, 2012); provavelmente derivadas de populações altamente polimórficas para inversões presentes na Facultad de Agronomia, Montevideo, Uruguai (Valente *et al.*, 2001; Valente *et al.*, 2003).

A disponibilidade de indivíduos com boa viabilidade, a partir de um estoque de *D. willistoni* com a inversão IIL-H fixada, fornece um excelente material para as abordagens moleculares dos pontos de quebra desta inversão que ocorre na natureza. A inversão IIL-H caracteriza-se por ser uma inversão telomérica, entre as seções 53A e 55B, e é amplamente distribuída ao longo da distribuição geográfica meridional da espécie (Rohde & Valente, 2012).

Estas duas populações foram empregadas por apresentarem arranjos cromossômicos distintos e fixados no braço cromossômico IIL.

II.2 – Análises citogenéticas

II.2.1 – Citologia clássica dos cromossomos politênicos

Para as análises citogenéticas, as quais necessitam de larvas em terceiro estágio de desenvolvimento, as duas linhagens receberam tratamento diferencial. As linhagens estoques foram mantidas em câmara com temperatura de $22\text{ }^{\circ}\text{C} \pm 1$ grau e 60% de umidade relativa. A cada três dias, os adultos foram repicados, e os vidros contendo somente os ovos e larvas de primeiro estágio foram passados para a câmara de temperatura $18\text{ }^{\circ}\text{C} \pm 1$ grau e 60% de umidade relativa. A alimentação das larvas se deu a cada intervalo de um dia, com levedura fresca (fermento comercial). Em aproximadamente 7/10 dias foram obtidas as larvas de terceiro estágio para o uso nas preparações cromossômicas.

A citologia clássica foi empregada para a obtenção de bons preparados citológicos dos cromossomos politênicos das populações utilizadas neste estudo (**Tópico II.1** deste capítulo), a fim de caracterizar o padrão de bandas e o polimorfismo cromossômico destas. Para obter preparados citológicos de boa qualidade, são necessárias diversas etapas essenciais. Estas incluem a obtenção e tratamento das larvas (conforme descrito acima), tratamento de lâminas e lamínulas (Veja Apêndice I, pág. 93), dissecação das glândulas salivares, e preparação cromossômica. As duas últimas etapas são descritas abaixo.

Dissecação das glândulas salivares: As glândulas salivares de larvas em terceiro estágio de desenvolvimento apresentam o mais elevado grau de politenia. Estas glândulas foram dissecadas em Solução fisiológica 0,9% de NaCl com o auxílio de pinça e agulha histológica, e depois transferidas para uma lâmina.

Preparação cromossômica: Na lâmina previamente tratada, as glândulas salivares foram fixadas com uma gota (20 μL) de solução de ácido acético 45%, por um período de 30 segundos a 1 minuto. Após, o excesso de fixador foi retirado e acrescentou-se uma gota de orceína-acéto-láctica 5% por 5

minutos. Retirado o excesso de corante, as glândulas salivares foram submetidas a estresse osmótico com solução de ácido acético-água-ácido láctico na razão 3:2:1 (v/v) por 3 minutos. Após esse tempo, uma lamínula extremamente limpa foi colocada sobre o material. A ruptura das células, e o espalhamento dos cromossomos, foram obtidos por batidas, com o auxílio de um objeto pontudo, sobre a lamínula. Por fim, com o auxílio de um papel absorvente, foi realizado o esmagamento (*squash*) para aplanar o material. Com exceção da solução do corante, as soluções fixadoras ácidas foram preparadas no momento do uso.

Subsequentemente, os preparados cromossômicos obtidos foram todos analisados com o uso de microscópio de contraste de fase *NIKON Optiphot-2* ou *Zeiss*. A comparação do padrão de bandas foi feita com relação ao fotomapa da espécie (Shcaeffler *et al.*, 2008). Os melhores preparados foram selados com esmalte comercial, e documentados via a captura de imagens por câmara digital *Moticam Package (Moticam3 com 3.0 MP e Motic Images Plus 2.0)* ou *Sony DSC-H50* acopladas ao microscópio. As imagens obtidas foram processadas com uso do programa *Adobe Photoshop® CS6*.

Preparações cromossômicas de boa qualidade são aquelas onde os cromossomos de cada linhagem apareçam espalhados, de forma que cada braço cromossômico possa ter seu padrão de bandas reconhecido e ser analisado.

II.2.2 - Hibridação *in situ* não fluorescente (HIS)

Para a localização precisa das sequências de interesse nos cromossomos politênicos foi utilizada a técnica de hibridação *in situ* não fluorescente (do Inglês *In situ Hybridization - HIS*), uma vez que ela preserva o padrão de bandas e é recomendada para mapeamento físico de alta qualidade (*Drosophila* 12 Genomes Consortium, 2007). A técnica se utiliza de uma propriedade inerente dos ácidos nucléicos que é a complementaridade de bases com formação de duplexes. Consiste basicamente no pareamento de determinado segmento de DNA ou RNA, com uma sequência de nucleotídeos complementar situada dentro da célula, visando verificar se a célula possui essa sequência e qual a sua exata localização (Guerra, 2004). O processo da

HIS consiste nas etapas de: preparação citológica, marcação da sonda, pré-hibridização, hibridação *in situ* propriamente dita, pós-hibridação, detecção, revelação e visualização.

- Preparação citológica

As glândulas salivares de larvas bem alimentadas e em terceiro estágio de desenvolvimento são dissecadas em Solução fisiológica 0,9% de NaCl. Em uma lamínula extremamente limpa coloca-se uma gota de solução ácido acético-água-acido láctico na razão 2:2:1 (v/v). A glândula salivar dissecada e limpa repousa neste fixador por 4 minutos. Após, acopla-se a lâmina tratada (Veja Apêndice 1, pág. 90) e com o auxílio de um objeto pontudo, faz-se o espalhamento do material. Por último, com o auxílio de um papel absorvente, faz-se o esmagamento para aplanar o material.

Os melhores preparados cromossômicos são armazenados em posição horizontal a 4°C pelo período de uma noite. No outro dia, estas lâminas são imersas em nitrogênio líquido até o resfriamento total. Subsequentemente, com o auxílio de uma lâmina de aço, as lamínulas são retiradas e então as lâminas são imersas em etanol 95 GL, por pelo menos 15 minutos. Após, é realizada a secagem em TA, e nova inspeção em microscópio de contraste de fase. Os preparados cromossômicos que permaneceram em condições ideais para o uso são armazenados a 4°C.

- Marcação da sonda

A marcação fez-se de maneira direta pela técnica da *PCR*, com biotina-14-dUTP do kit *Biotin PCR Labeling Core (Jena Bioscience)*. Este kit fornece *Taq DNA polimerase* (2 U/ μ L), tampão *PCR* 10x, mix de nucleotídeos (1 mM dATP; 1 mM dCTP; 1 mM dGTP; 0,5 mM dTTP; 0,5 mM Biotina-dUTP; pH 7.5) e água. Ele incorpora uma porcentagem aleatória de Biotina-dUTP ao longo das sínteses das novas fitas cópias da sequência de interesse na *PCR*. O volume da reação foi ajustado das recomendações do fabricante, de 20 μ L para 50 μ L finais.

5 μ L - Tampão 10x
 5 μ L - Mix Nucleotídeos
 1 μ L - Taq DNA polimerase
 2 μ L - *Primer* senso (estoque a 10 μ M)
 2 μ L - *Primer* antisenso (estoque a 10 μ M)
 50 ng - DNA
 H₂O – completar para 50 μ L

As condições do ciclo da *PCR* seguem as sugestões dadas pelo fabricante:

94°C – 2 minutos - Desnaturação inicial	} 30 vezes
94°C – 30 segundos - Desnaturação	
50-60°C – 30 segundos - Anelamento dos <i>primers</i>	
72°C – 1 minuto/kb - Elongação	
72°C – 5 minutos - Elongação final	

A temperatura de anelamento e tempo de extensão foram ajustados conforme o par de *primers* utilizado e o tamanho do fragmento amplificado. Após a *PCR*, foram aplicados 3 μ L das amostras em gel de agarose 0,8% corado com brometo de etídio ou *GelRed*TM (*Biotium*), e submetidos à migração em Tampão TAE 1X a 75V. A visualização deu-se em transluminador ultravioleta (UV). Os produtos da *PCR* com resultado positivo (ou a excisão da banda do fragmento de interesse) foram purificados com *NucleoSpin Gel and PCR Clean-up Kit* (*MACHEREY-NAGEL*) ou com ilustra *GFX*TM *PCR DNA and Gel Band Purification Kit* (*GE Healthcare*). Subsequentemente, os produtos da *PCR* marcados e purificados foram quantificados em *Qubit Fluorometric 1.0* (*Invitrogen*) ou *NanoDrop 1000 Spectrophotometer* (*Thermo Cientific*).

Foram utilizados dois tipos de sequências genômicas como sondas: sequências gênicas e sequências de regiões intergênicas.

O protocolo para *HIS* utilizado foi o de Biémont *et al.* (2004), com menores modificações, conforme segue abaixo.

- Pré-hibridação

É o tratamento do material citológico que visa a impedir desidratação deste material para a não dissolução da sonda, bem como a interação da sonda com alguma proteína básica. Consiste em:

- Imersão em 2x SSC a 70°C por 30 minutos
- Enxague em 2x SSC a TA
- Imersão em etanol 70 GL por 5 minutos
- Imersão em etanol 95 GL por 10 minutos
- Secagem em TA por pelo menos 20 minutos

- Hibridação

Ocorre a desnaturação da sonda e do material citológico, seguido da hibridação dos ácidos nucleicos por complementaridade. A desnaturação neste protocolo é química, onde as proteínas associadas ao DNA são corroídas e as bases nitrogenadas ficam livres para o encontro com a sonda. Os passos são:

- Imersão em NaOH 0,07 N por 2 minutos e 10 segundos
- Enxague em 2x SSC
- Imersão em etanol 70 GL por 5 minutos
- Imersão em etanol 95 GL por 10 minutos
- Secagem em TA por pelo menos 20 minutos

A preparação da sonda que se vai aplicar no preparado cromossômico ocorre da seguinte forma:

Acrescentar de 100-150 ng do produto da sequência obtida por meio da marcação pela técnica da *PCR*, em um volume final de 30 µL completados com H₂O ultrapura

- Desnaturar a sonda a 100 °C por 10 minutos
- Imediatamente, banho de gelo por 3 minutos
- Dar um spin

Acrescentar 50 µL de solução de hibridação (10 µL 20x SSC + 10 µL de dextran sulfato 50% + 30 µL de formamida – para sondas homólogas).

- Dar um vórtex para homogeneizar

Aplicar todo o conteúdo acima no preparado cromossômico, e cobrir com lamínula

Colocar as lâminas em câmara úmida na posição horizontal e deixar a hibridação ocorrendo por 16 horas (máximo 20 horas) a 37°C para sondas homólogas, e a 25°C para sondas heterólogas.

- Etapa Pós-hibridação

Nesta etapa são feitas as lavagens de estringência para garantir a máxima quantidade de nucleotídeos marcados hibridados com seu sítio complementar específico, o que irá evitar marcações inespecíficas. Consiste em:

Imersão em 2x SSC a 37 °C por 10 minutos

Repetir o passo acima

Enxágue em 2x SSC a TA

Enxágue em 0,1% Triton-X-100 em 1x PBS por 5 segundos

Repetir o passo acima

Enxágue em 1x PBS

- Etapa de Detecção

Para a detecção dos sinais de hibridação utilizou-se o *kit* de sistema indireto de detecção enzimática ABC *Elite Vectastain* (*Vector Laboratories, Inc. Burlingame, CA, USA*), baseado no reconhecimento e na ligação irreversível da biotina (vitamina H) à glicoproteína Avidina. A solução é preparada com:

210 µL de 1x PBS

3,5 µL de Reativo A (Avidina DH)

3,5 µL de Reativo B (Peroxidase biotinilada)

Vortex para homogeneizar

Repousar no mínimo 15 minutos a TA. A preparação da solução de detecção deve iniciar juntamente com as primeiras lavagens de estringência.

Aplicar 20 µL em cima do preparado cromossômico e cobrir com lamínula.

Colocar as lâminas em câmara úmida na posição horizontal por 30 minutos a 37°C.

- Etapa de Revelação

Para a revelação dos sítios de hibridação positivos detectados pela reação enzimática, utilizou-se *DAB Substrate (Roche)*. Para tal, mistura-se 105 µL de *DAB (3,3'-Diaminobenzidina)* + 945 µL de tampão. A reação que irá revelar os sítios de hibridação consiste na oxidação do *DAB* por meio da peroxidação dos íons peróxido presentes no buffer, pela enzima peroxidase neste momento já complexada à sonda conforme descrito no item Detecção.

Aplicar 200 µL da solução *DAB* + Tampão sobre o preparado cromossômico e cobrir com lamínula.

Colocar as lâminas em câmara úmida na posição horizontal por 45 minutos

Enxague em 0,1% Triton-X-100 em 1x PBS por 5 segundos

Enxague em 1x PBS

A **Figura II.1** mostra uma esquematização do sistema de revelação e detecção utilizados na presente Tese.

Após a Etapa da Detecção, os cromossomos politênicos são corados imergindo as lâminas por 5 minutos em solução de corante Giemsa 5%, conforme abaixo especificado.

Solução de Giemsa 5%

2,5 µL de corante Giemsa (*New Prov*[®])

50 µL de tampão fosfato (Veja Apêndice I, pág. 94)

Imersão por 5 minutos

Enxaguar abundantemente com água corrente

Secar

Para a montagem permanente das lâminas foi utilizado *EUKITT*[®] (*Panreac*) ou *Eukitt*[®] *quick-hardening mounting médium (SIGMA-ALDRICH)*. As melhores marcações obtidas foram fotografadas em fotomicroscópio com contraste de fase, analisadas e processadas com o programa Adobe Photoshop[®] CS6.

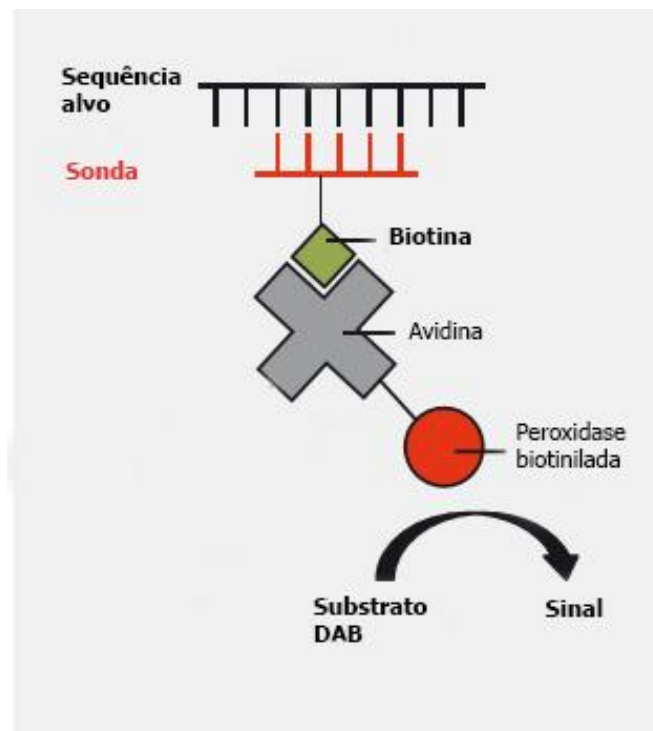


Figura II.1: Representação esquemática do sistema indireto de detecção e revelação utilizados na hibridação *in situ* não fluorescente. As sondas marcadas com biotina-14-dUTP são reconhecidas pela proteína Avidina, a qual possui quatro sítios de ligação de alta afinidade para Biotina. Esta Avidina é complexada com a enzima Peroxidase, a qual é também biotinilada, de acordo com especificações do kit *ABC Elite Vectastain* (*Vector Laboratories, Inc. Burlingame, CA, USA*), não representado com fidelidade na figura. Esta enzima atuará no seu substrato, o Peróxido de Hidrogênio, o qual se encontra homogeneizado ao *DAB* (3,3'-*Diaminobenzidina*). Esta reação oxida o *DAB*, o que resultará em uma coloração marrom escuro-preta na região, visível sob microscopia óptica. Retirado e modificado de *Jena Bioscience*, disponível online:

https://www.jenabioscience.com/images/741d0cd7d0/bro_FluorescentProbes_WEB.pdf

II.3 - Extração de DNA genômico

O DNA genômico foi extraído por fenol-clorofórmio a partir de moscas adultas de acordo com o método aprimorado de Sassi *et al.* (2005), amplamente utilizado por nosso grupo de pesquisa. A metodologia empregada está descrita no Apêndice I (pág. 93).

II.4- Estabelecimento das sondas no genoma

As sequências utilizadas como sondas foram analisadas com base no genoma da linhagem Gd-H4-1 de *D. willistoni* (*Drosophila* 12 Genomes Consortium, 2007), *Release* 1.04, disponível no banco de dados Flybase - A Database of *Drosophila* Genes & Genomes (St. Pierre *et al.*, 2014). Estas sequências eram preferencialmente pertencentes às regiões exônicas, por estas serem mais conservadas; embora de acordo com a necessidade também foram feitas sondas em região intrônica e intergênica. Entretanto, evitou-se que nestas sequências contivessem algum domínio repetitivo proteico, bem como algum *TE* inserido.

A inferência de domínio repetitivo proteico foi feita por BLASTN nos bancos de dados Flybase (St. Pierre *et al.*, 2014; Configurações: *Database: Genome Assembly, Species: D. willistoni, Advanced Settings: padrão do banco de dados*) e no banco de dados NCBI (Geer *et al.*, 2010; Configurações: *Database: Others – Nucleotid Colletion: nr/nt, Organism: D. willistoni, Optime for: Somewhat similar sequences, Expect threshold: 7*). Deu-se preferência para a escolha de sequências com um único *hit* significativo (e-value 0,0).

Por sua vez, a inferência de *TEs* foi feita pela submissão da sequência de interesse para sonda no banco de dados *Repbase* (Jurka *et al.*, 1996), o qual faz buscas por meio do programa *Censor* (Kohany *et al.*, 2006; Configuração: *Sequence Source: All*). Sendo assim, a sequência era comparada com todos os *TEs* presentes no banco de dados, descritos para

diferentes organismos. Deu-se preferência para a escolha de sequências sem a presença de *TEs* descritos para o gênero *Drosophila*, uma vez que a estringência para nenhuma sequência que contivesse *TEs* de outros organismos tornaria inviável a determinação das sondas.

II.5 – Desenho de *primers*

Os *primers* para amplificação das sondas foram desenhados baseados no genoma da linhagem Gd-H4-1 de *D. willistoni* (*Drosophila* 12 Genomes Consortium, 2007), disponível no banco de dados do Flybase - A Database of *Drosophila* Genes & Genomes (St. Pierre *et al.* 2014), Release 1.04. Para tal, utilizou-se o programa de desenho manual *Primer Designer* v.1.01 (*Scientific and Educational Software*) com as seguintes especificações: tamanho 20 pb; 40% mínimo e 60% máximo de conteúdo GC; não aceitar repetições de três nucleotídeos iguais (três *runs*), temperatura de anelamento entre 50-60°C. De preferência o nucleotídeo 3' terminal do primer G ou C, e se evitou dímeros de primers acima de 3 pb.

II.6 – Técnica da Reação em Cadeia da Polimerase (*PCR*) e medição dos fragmentos

A reação de amplificação realizou-se em termociclador *Master Cycler PTC-100TM* (*MJ Research, Inc*) e *T-100TM* (*BIO-RAD*) utilizando-se um volume total de 20 µL contendo 25 ng de DNA genômico, 1x buffer de reação, 0,4 µM de cada *primer* (senso e antisenso), 2,5 µM de dNTPs, 1,0 mM de MgCl₂ e 1,5 unidade de *Taq DNA polimerase, recombinante* (*Invitrogen*).

As condições do ciclo da *PCR* seguiram as sugestões dadas pelo fabricante:

95°C – 3 minutos - Desnaturação inicial	}	30 a 36 vezes
95°C – 30 segundos - Desnaturação		
50-60°C – 30 segundos - Anelamento dos <i>primers</i>		
72°C – 1 minuto/kb - Elongação		
72°C – 5 minutos - Elongação final (10 minutos para as sequências que seriam clonadas)		

A temperatura de anelamento e tempo de extensão foram ajustados conforme o par de *primers* utilizados e o tamanho do fragmento amplificado. Após a *PCR*, as amostras foram aplicadas em gel de agarose 0,8% corado com brometo de etídio ou *GelRed™* (*Biotium*), e submetidas à migração em Tampão TAE 1X a 75V. A visualização deu-se em transluminador ultravioleta (UV). Os fragmentos obtidos foram comparados com o marcador de peso molecular *1 Kb Plus DNA Ladder™* (*Invitrogen*) com relação ao seu tamanho em pares de base esperado.

II.7 – Clonagem

As sequências dos genes *Dwil\GK21049* e *Dwil\GK21115*, tiveram suas bandas excisadas do gel de agarose e purificadas com o Kit *GFX™ PCR and Gel Band Purification (GE)*, para ligação ao vetor plasmidial *pGEM -T Easy Vector System (Vector Systems I of Promega, Madison, WI, USA)*. Subsequentemente, estas sequências foram transformadas por choque térmico na linhagem *TOP 10* de *Escherichia coli*. Estas sequências serão utilizadas como sondas para a varredura dos pontos de quebra da inversão IIL-H no genoma da linhagem SG12.00 do Uruguai.

Na **Tabela II.2** encontram-se os *primers* específicos utilizados para a amplificação dos fragmentos dos genes *Dwil\GK21049* e *Dwil\GK21115* para a clonagem.

Tabela II.2: *Primers* e região genômica dos genes clonados *Dwi\GK21049* e *Dwi\GK21115*. Número do *scaffold* corresponde ao material disponível no banco de dados *FlyBase* (St. Pierre *et al.*, 2014) e corresponde aos quatro últimos números, os quais todos começam com *scf2_110000000*.

Gene	Nome do <i>Primer</i>	Sequência do <i>Primer</i> 5' - 3'	Número do <i>scaffold</i>	Região genômica da sequência
Dwi\GK210149	74L	GCCAACTCGTGATATTCTTG	4851	513.104 –
	74R	TGGCAAGGAACCTCCAGAGA		513.540
Dwi\GK21115	75L	TTGCATTCGTGGATAAAT	4851	517.004 –
	75R	ACACGTTACGAGGTACTION		517.315

Após a clonagem, a extração dos plasmídeos com a sequência de interesse foi feita por protocolo de lise alcalina conforme Sambroock & Russel (2004), com pequenas modificações (detalhes Veja Apêndice I, pág. 98).

II.8 - Sequenciamento

As sequências de interesse foram sequenciadas automaticamente com os *primers* universais M13 (senso e antisenso), ou com seus *primers* específicos, pela empresa de sequenciamento *Macrogen Inc.* (Seoul, Korea).

Para tal, utilizou-se 10 µL do produto da *PCR* destes *primers*, os quais para o envio foram purificados com *Exonuclease I (USB)* e *Shrimp Alkaline Phosphatase (USB)*. Ambos foram acrescentados à amostra, a qual foi submetida a 37°C por 30 minutos, seguido de inativação das enzimas a 80°C por 15 minutos em termociclador.

II.9 - Análise das sequências

As análises dos cromatogramas, edição e montagem foram feitas nos programas do *Staden Package* (Staden, 1996). A comprovação da sequência foi feita por BLASTN nos bancos de dados *Flybase* (St. Pierre *et al.*, 2014) e *NCBI* (Geer *et al.*, 2010), utilizando como sequência de busca (*query*) a sequência consenso obtida.

II.10 – Determinação da diferença do padrão de bandas do braço cromossômico IIL da linhagem sequenciada de *Drosophila willistoni* com relação ao fotomapa da espécie

Esta análise consistiu na comparação do padrão de bandas entre o fotomapa de *D. willistoni* (Schaeffer *et al.*, 2008) e a linhagem sequenciada Gd-H4-1, a fim de estabelecer, de maneira clara e definitiva, um correto alinhamento entre o padrão citológico e o padrão genômico desta última linhagem. As preparações citológicas foram feitas, analisadas e fotodocumentadas conforme a técnica de citologia clássica (**Tópico II.2.1** deste capítulo).

II.11 – Cruzamentos entre as linhagens Gd-H4-1 e SG12.00 de *Drosophila willistoni*

Foram feitos cruzamentos recíprocos entre as linhagens Gd-H4-1 (linhagem sequenciada) e SG12.00. Para tal, foram capturados dos estoques, a cada 3 h, fêmeas e machos virgens. Um macho e uma fêmea foram postos para cruzarem em tubos de cultivo de 50 mL, contendo meio de cultura padrão (Marques *et al.*, 1966). Após quatro dias, os casais foram passados para um

tubo com meio de cultivo novo, onde permaneceram por uma semana para ovoposição das fêmeas. A partir daí, as larvas emergentes foram alimentadas diariamente, até a obtenção de larvas em terceiro estágio de desenvolvimento ideais para o uso. As preparações citológicas foram feitas, analisadas e fotodocumentadas conforme a técnica de citologia clássica (**Tópico II.2.1** deste capítulo).

II.13 - Mapeamento de marcadores genéticos no cromossomo II de *Drosophila willistoni*

Os braços cromossômicos IIL e IIR de *D. willistoni* correspondem à montagem de oito *scaffolds* em cada um, no genoma publicado da espécie. Para esta montagem, fizeram-se análises da conservação dos blocos sintênicos, em comparação com as outras espécies também sequenciadas. Para comprovar a montagem e orientação dos *scaffolds* nesses braços cromossômicos, foi estabelecido um marcador genético, para ser utilizado como sonda, para cada um dos *scaffolds* na linhagem sequenciada Gd-H41, com exceção dos *scaffolds* 4577 (IIL) e 4558 (IIR), para os quais foram estabelecidos dois marcadores.

Adicionalmente, foi estabelecido um marcador genético para o braço XL e o cromossomo III, e dois marcadores para o braço XR, a fim de que se pudesse obter uma breve visualização acerca da montagem dos *scaffolds* nestes cromossomos.

As **Tabelas II.3** e **II.4** apresentam os marcadores utilizados nesta análise e seus respectivos *primers*. A determinação das sondas e o desenho dos primers foram feitos de acordo com os **Tópicos II.4** e **II.5** deste capítulo, respectivamente.

Já a determinação física destes marcadores nos cromossomos politênicos da linhagem Gd-H4-1 foi estabelecida pela técnica da *HIS* não fluorescente, conforme descrito no **Tópico II.2.2** deste capítulo, e da comparação com o fotomapa atual da espécie disponível (Schaeffer *et al.*,

2008). Esta determinação também levou em consideração a diferença do arranjo cromossômico do braço IIL da linhagem sequenciada, com relação ao arranjo do fotomapa da espécie (Rohde & Valente, 2012).

Todas as sondas utilizadas nesta análise foram sequenciadas diretamente do seu produto da *PCR*, de acordo com **Tópico II.8** deste capítulo. e as sequências analisadas e comprovadas, de acordo com **Tópico II.9** deste capítulo.

II.14 – Delimitação dos pontos de quebra da inversão IIL-H do cromossomo II de *Drosophila willistoni*

Para a delimitação dos pontos de quebra da inversão IIL-H fez-se necessário a análise conjunta do ponto de quebra distal desta inversão (ponto de quebra próximo ao telômero) e do ponto de quebra proximal (ponto de quebra próximo ao centrômero).

A análise do ponto de quebra distal, o qual está localizado citologicamente na linhagem SG12.00 na seção 53A do braço IIL (Rohde & Valente 2012), iniciou-se com as sondas 14, 15 e 16 (**Tabela II.3**), já abordadas no **Tópico II.13** deste capítulo, uma vez que elas circundavam esta região citológica. Posteriormente, se estabeleceram outros marcadores genéticos para serem utilizados como sonda (conforme **Tópicos II.4, II.5 e II.6** deste capítulo) para assim, por meio de caminhada cromossômica por *HIS* não fluorescente (**Tópico II.2.2** deste capítulo), determinar os genes flanqueadores deste ponto de quebra.

As sondas utilizadas na caminhada cromossômica do ponto distal estão descritas, juntamente com seus respectivos *primers*, na **Tabela II.5**. Estas sondas foram analisadas e estabelecidas com base no genoma da linhagem Gd-H4-1, e posteriormente por *HIS* nos cromossomos politênicos da linhagem SG12.00 do Uruguai, a qual apresenta a inversão IIL-H fixada.

Com a caminhada cromossômica, conseguiu-se esgotar a região genômica compreendida entre os genes *DwiIGK21048* e *DwiIGK21116*

(**Tabelas II.5 e II.6**). Analisando o genoma, constatou-se que esta região consistia de três regiões intergênicas, com tamanho não superior a 1,5 kb. A partir daí passou-se a utilizar uma abordagem para a determinação do ponto de quebra distal em ambas as linhagens pela técnica da *PCR* (de acordo com **Tópico II.6**). O esperado era a ausência de amplificação na linhagem SG12.00, uma vez que a região onde ocorresse o ponto de quebra não se obtenha amplificação, devido a região estar corrompida e ao contrário. Como precaução, inseriu-se uma região intergênica anterior e uma posterior, resultando em cinco intervalos intergênicos denominados de A a E (**Figura II.2**).

Os genes envolvidos na análise pela técnica da *PCR* das regiões intergênicas, e seus respectivos primers estão descritos na **Tabela II.6**.

A análise do ponto de quebra proximal teve seu início de maneira mais aleatória, uma vez que este ponto está localizado citologicamente na seção 55B da linhagem SG12.00 (Rohde & Valente, 2012). Sendo este ponto de quebra pertencente ao *scaffold* scf2_1100000004521, do qual só dispúnhamos do marcador físico da sonda denominada 13 (**Tabela II.3**), a qual se localiza muito distante da região citológica de interesse. De maneira exploratória, em uma primeira etapa de análise, estabeleceram-se marcadores genéticos para serem hibridados por *HIS* não fluorescente (**Tópico II.2.2** deste capítulo) a cada 2 Mb, a partir dos 6 Mb deste *scaffold*. A caminhada cromossômica do ponto de quebra proximal da inversão IIL-H compreendeu ao todo quatro etapas de planejamento e estabelecimento de sondas gênicas e intergênicas, juntamente com a tentativa de delimitação do ponto de quebra pela técnica da *PCR* (**Tópico II.6** deste capítulo).

A **Tabela II.7** apresenta os marcadores estabelecidos para o uso como sondas da caminhada cromossômica do ponto de quebra proximal, bem como as regiões intergênicas que foram analisadas pela técnica da *PCR*, juntamente com seus respectivos *primers*.

A região dos pontos de quebra distal e proximal da inversão IIL-H na linhagem Gd-H4-1, foram analisadas quanto à sintonia com *D. melanogaster* no banco de dados *Flybase* (St. Pierre *et al.*, 2014), e quanto à presença de *TEs*

no banco de dados *RepBase* (Jurka *et al.*, 1996), por meio da ferramenta de busca *Censor* (Kohany *et al.*, 2006).

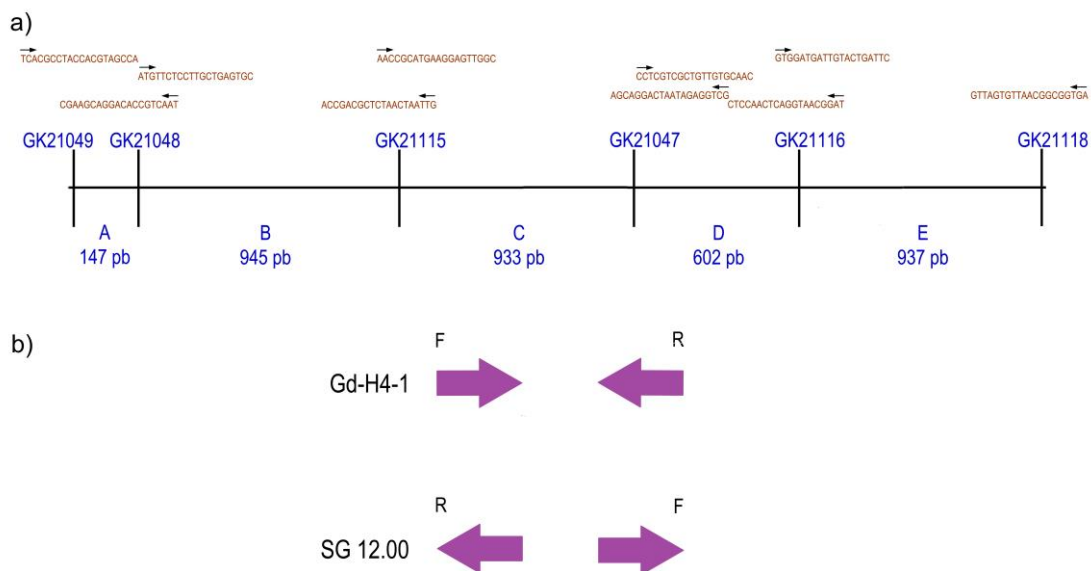


Figura II.2: a) Delimitação gênica do ponto de quebra distal da inversão IIL-H, pela técnica da *PCR*. Posicionamento dos *primers* nas regiões intergênicas denominadas de A-E, de acordo com o genoma da linhagem Gd-H4-1. b) Posicionamento esperado dos primers da região intergênica contendo o ponto de quebra distal na linhagem sequenciada Gd-H4-1, a qual tem a sequência intergênica íntegra; e na linhagem SG12.00, a qual apresenta a inversão IIL-H fixada e essa região corrompida pela inversão.

Tabela II.3: Marcadores genéticos utilizados para a redefinição dos *scaffolds* do cromossomo II de *Drosophila willistoni*, listados do centrômero ao telômero. Número do *scaffold* de acordo com o material disponível no banco de dados *FlyBase* (St. Pierre *et al.*, 2014) correspondendo aos quatro últimos números, os quais todos começam com scf2_110000000.

Braço cromossômico IIL					
<i>D. willistoni</i> Gene	Nome do Primer	Nome da sonda	<i>D. melanogaster</i> Gene ortólogo	Número do scaffold	Primers F e R (5'-3')
<i>Dwil</i>\GK18743	9L 9R	9	<i>Dmel</i> \Lamp1	4884(1)	GTCAAGCAGTAGCAGCACCA GCCACGCGAAGTTGATCGAC
<i>Dwil</i>\GK24519	10L 10R	10	<i>Dmel</i> \CG9515	4585(2)	ATTCAATTCACAGCACAACC GGACTCAATGCGGAACTATG
<i>Dwil</i>\GK18260	11L 11R	11	<i>Dmel</i> \Yuri	4577(3)	GTGAAGAGCCTACACACAGC CTTCTGAGATGATCCACGAC
<i>Dwil</i>\Adh	Dwil_AdhL Dwil_AdhR	<i>Adh</i>	<i>Dmel</i> \Adh	4577(3)	CATGGAACGTGTTAAGTGCC AGTTCACAGCAATGGTACGC
<i>Dwil</i>\GK23840	12L 12R	12	<i>Dmel</i> \CG17549	4516(4)	GGAGGATATGCTGGTGGTTA GTGCTGACTTGCTCCAAC TG
<i>Dwil</i>\GK15054	13L 13R	13	<i>Dmel</i> \CG7371	4521(5)	ATCCTGAGCCTGAGTTCCAC CGCCAAGAGAATTGTCATCG
<i>Dwil</i>\GK12721	14L 14R	14	<i>Dmel</i> \CG13127	4752(6)	GCAGCTCGATGAACTCTATG TCTCCGAAGACTGTGTACTC
<i>Dwil</i>\GK18432	15L 15R	15	<i>Dmel</i> \fusl	4945(7)	GCATCAGCCTCATATCCATC GTCAACACTCTCGGCTCCAG
<i>Dwil</i>\GK21099	16L 16R	16	<i>Dmel</i> \Rab3-GAP	4851(8)	CTGGAGCAGTCAAGGCGAGA ATCCAAGCATCCTAAGCGTG

Tabela II.3: Continuação.

Braço cromossômico IIR					
<i>D. willistoni</i> Gene	Nome do Primer	Nome da sonda	<i>D. melanogaster</i> Gene ortólogo	Número do scaffold	Primers F e R (5'-3')
<i>Dwil</i> \GK23049	Dwil_bwL Dwil_bwR	<i>bw</i>	<i>Dmel</i> \bw	4954(8)	CAGTAGTAACCACTCCGATG GCGGACACATTGTCTACCAG
<i>Dwil</i> \GK22144	7L 7R	7	-	4558(7)	CATTCGACGATCTCAGCAAC TCACTTCGGACTACTCCAGC
<i>Dwil</i> \GK22138	Dwil_pxL Dwil_pxR	<i>px</i>	<i>Dmel</i> \px	4558(7)	GCTGCATTAGATCCTCATAG GGCAGCCAACAGTCCATACA
<i>Dwil</i> \vlc	6L 6R	6	<i>Dmel</i> \vlc	4382(6)	CAACGCCACTATCTGTGAAG TCTCATTGCACTCACCTACA
<i>Dwil</i> \GK17912	20L 20R	20	<i>Dmel</i> \FLASH	4822(5)	ACGATGATCTGGATGAGTTC TACAACATACCTAGTTCCGC
<i>Dwil</i> \GK20645	32L 32R	32	<i>Dmel</i> \CG2269	4510(4)	CTGATGGACACCACAGAGTG ACTCGTACAACATGGCGGAC
<i>Dwil</i> \GK15808	3L 3R	3	<i>Dmel</i> \Ir60a	4514(3)	AACGAGGCAGTCACCGATAC CATATCGGACGCTCTTGAAC
<i>Dwil</i> \Adam	2L 2R	2	<i>Dmel</i> \Adam	4513(2)	GGTGAGGATGACGATGAGGA TCCGAATGTAAGAGCTCCAC
<i>Dwil</i> \GK19495	1L 1R	1	<i>Dmel</i> \Hsf	4512(1)	GGCTACCGTCATAAGATCAG AGAACATACGTGGACGTCAG

Tabela II.4: Marcadores genéticos utilizados como sondas para os cromossomos X e III de *Drosophila willistoni*. Número do *scaffold* de acordo com o material disponível no banco de dados *FlyBase* (St. Pierre *et al.*, 2014) correspondendo aos quatro últimos números, os quais todos começam com scf2_110000000.

<i>D. willistoni</i> Gene	Nome do Primer	Nome da sonda	<i>D. melanogaster</i> Gene ortólogo	Número do <i>scaffold</i> / Cromossomo	Primers F e R (5'-3')
<i>Dwi</i>\GK16707	68L	68	<i>Dme</i> \unc	4963/XL	ACTCAGTCTTCGACGGAAGC
	68R				AGTTGTATCGGATTCTACCA
<i>Dwi</i>\GK17758	5L	5	<i>Dme</i> \ida	4822/XR	GCTGCATTAGATCCTCATAG
	5R				GGCAGCCAACAGTCCATACA
<i>Dwi</i>\GK16749	30L	30	<i>Dme</i> \CG13313	4511/XR	GCTATCAGTCACCGTGTAGA
	30R				GGCAGTTGCTCCACCATCAC
<i>Dwi</i>\GK22422	31L	31	<i>Dme</i> \CG31204	4921/III	GAGTCAATGCGTCCATACCA
	31R				GGATAATCCTCACGAGACTG

Tabela II.5: Marcadores genéticos utilizados como sondas para a delimitação do ponto de quebra distal da inversão IIL-H de *Drosophila willistoni* via caminhada cromossômica. Número do *scaffold* de acordo com o material disponível no banco de dados *FlyBase* (St. Pierre *et al.*, 2014) e correspondendo aos quatro últimos números, os quais todos começam com scf2_110000000.

<i>D. willistoni</i> Gene	Nome do <i>Primer</i>	Nome da sonda	<i>D. melanogaster</i> Gene ortólogo	Número do <i>scaffold</i>	<i>Primers F e R (5'-3')</i>
<i>Dwi</i>IGK12721	14L	14	<i>Dmel</i> CG13127	4752	GCAGCTCGATGAACTCTATG
	14R				TCTCCGAAGACTGTGTACTC
<i>Dwi</i>IGK18432	15L	15	<i>Dmel</i> fusl	4995	GCATCAGCCTCATATCCATC
	15R				GTCAACACTCTCGGCTCCAG
<i>Dwi</i>IGK18461	24L	24	<i>Dmel</i> CG3285	4995	CTACAAGATCGCATTGGACG
	24R				TGAGCGTAAGGAAGACCGTA
<i>Dwi</i>IGK18345	25L	25	<i>Dmel</i> CG42389	4995	CTCTTG TAGCTTCGCTAGGC
	25R				CGGATCCGTA CTCAAC
<i>Dwi</i>IGK21099	16L	16	<i>Dmel</i> Rab3-GAP	4851	CTGGAGCAGTCAAGGCGAGA
	16R				ATCCAAGCATCCTAAGCGTG
<i>Dwi</i>IGK21105	21L	21	<i>Dmel</i> Lam	4851	CTAGCATTGCTGAAGGTGAG
	21R				CCGAGAATCTGAAGACCACA
<i>Dwi</i>IGK21048	22L	22	<i>Dmel</i> CG43345	4851	GCCGTGTCGATGGTATTAAG
	22R				GTGGCAACGACTGGTGAAGA
<i>Dwi</i>IGK21115	72L	72	<i>Dmel</i> Acon	4851	AGATGCAGTGGATGTTATGG
	72R				CCTTGTCGTAGTCAGCAGGA
<i>Dwi</i>IGK21116	34L	34	<i>Dmel</i> glu	4851	AGATGCTTCGGCACTTGATG
	34R				CGGTTAGCGGTTCACTCTTA
<i>Dwi</i>IGK21118	35L	35	<i>Dmel</i> Dap160	4851	CACGCGGTGTTCTCGTTCAG
	35R				CTCTTGAGTTAGCTGCAATA

Tabela II.6: Regiões intergênicas e *primers* do *scaffold* scf2_1100000004851 utilizados para delimitação do ponto de quebra distal da inversão IIL-H de *Drosophila willistoni*, pela técnica da *PCR*.

Região intergênica em <i>D. willistoni</i>	Nome do <i>Primer</i>	Nome do par de primers	<i>Primers F e R (5'-3')</i>
<i>Dwil</i> \GK21049 e <i>Dwil</i> \GK21048	41L 41R	41	TCACGCCTACCACGTAGCCA TAACTGCCACAGGACGAAGC
<i>Dwil</i> \GK21048 e <i>Dwil</i> \GK21115	42L 42R	42	ATGTTCTCCTTGCTGAGTGC GTTAATCAATCTCGCAGCCA
<i>Dwil</i> \GK21115 e <i>Dwil</i> \GK21047	43L 43R	43	AACCGCATGAAGGAGTTGGC GCTGGAGATAATCAGGACGA
<i>Dwil</i> \GK21047 e <i>Dwil</i> \GK21116	44L 44R	44	CCTCGTCGCTGTTGTGCAAC ATGGCAATGGACTCAACCTC
<i>Dwil</i> \GK21116 e <i>Dwil</i> \GK21118	45L 45R	45	GTGGATGATTGTACTIONGATTTC AGTGGCGGCAATTGTGATTG

Tabela II.7: Marcadores genéticos utilizados como sondas para a delimitação do ponto de quebra proximal da inversão IIL-H de *Drosophila willistoni* via caminhada cromossômica, e pela técnica da *PCR*. Sondas de regiões gênicas e intergênicas do *scaffold* scf2_1100000004521. Interg= Região intergênica. (-) = Sequência utilizada somente como sonda; (+) = Sequência utilizada como sonda e para a técnica da *PCR*; (*) = Sequência utilizada somente para a técnica da *PCR*.

<i>D. willistoni</i> Gene ou Região intergênica	<i>D. melanogaster</i> Gene ortólogo	Nome do <i>Primer</i>	Nome da Sonda	<i>PCR</i>	<i>Primers F e R (5'-3')</i>
<i>Dwil</i> \GK14795	<i>Dmel</i> \Osm-6	17L 17R	17	-	GCCGTAATCCTAGCACACTC AAGCATGTTGGTCCTCTTGG
<i>Dwil</i> \GK14718	<i>Dmel</i> \CG9135	36L 36R	36	-	AGAGCTGTATGCGAGTAACC ACGTGATCGGAAGAATGTCA
<i>Dwil</i> \GK15415	<i>Dmel</i> \CG8965	37L 37R	37	-	TCCATGTGCGGCAGTCCAGA GCATAACGAGCATTGATGTG
<i>Dwil</i> \GK15416	<i>Dmel</i> \CG14005	46L 46R	46	-	TTGCGATTACGCTACTATAC GAACAGTCTCTTGTCCTTGT
Interg <i>Dwil</i> \GK15416- <i>Dwil</i> \GK15418	-	62iAL 62iAR	62iA	+	TAACGGCAAGGATATTGACA AGATTAATAAGGCGCGACTC

Tabela II.7: continuação.

Interg		62iBL				
<i>Dwi</i> \GK15416-	-	62iBR	62iB	+	CATTGATTCATCCGTATAGC	
<i>Dwi</i> \GK15418					CTACAACTATCCAAGAGCAA	
Interg		50L				
<i>Dwi</i> \GK15416-	-	50R	50	+	CACAATCTAATTGCTTACGG	
<i>Dwi</i> \GK15418					TGACAAGTCCTGATTGATTC	
Interg		50bL				
<i>Dwi</i> \GK15416-	-	50bR	50b	+	GATCCTGAGGAACAGAATGG	
<i>Dwi</i> \GK15418					TATGTTCAATGCAAGTGCGA	
Interg		51L				
<i>Dwi</i> \GK15416-	-	51R	51	-	ATGACAGCCTCCAACAGTGA	
<i>Dwi</i> \GK15418					GCTGCTCGTTGGTAATTATG	
Interg		76L				
<i>Dwi</i> \GK15416-	-	76R	76	*	TCGCAATACTTAGTAGACTC	
<i>Dwi</i> \GK15418					AGGAGTGAGAGCGAGATGGA	
Interg		77L				
<i>Dwi</i> \GK15416-	-	77R	77	*	TACCATGAGTCCGATTGATA	
<i>Dwi</i> \GK15418					AATTGACCTTGTCGAATGGA	

Tabela II.7: continuação

Interg		78L				CCTAACAAGTGAGTTATCAG
<i>Dwi</i> \GK15416-	-	78R	78	*		CAGAAGCTTCTAATGCCATG
<i>Dwi</i> \GK15418						
Interg		71bL				CGATAAGCCTGACATAATTC
<i>Dwi</i> \GK15416-	-	71bR	71b	+		GTACTGTCATTTCGCATCGCT
<i>Dwi</i> \GK15418						
Interg		79L				CATTGACAGAAGCATCCACG
<i>Dwi</i> \GK15416-	-	79R	79	*		AATGACATCGATAGCATAGC
<i>Dwi</i> \GK15418						
Interg		80L				ACTGATGAGCTGCTTCTGAG
<i>Dwi</i> \GK15416-	-	80R	80	*		TAAGGCGAATCTTCTATACC
<i>Dwi</i> \GK15418						
Interg		70L				TACGTCCGCAAGTTCTTAGC
<i>Dwi</i> \GK15416-	-	70R	70	+		AAGTGAAGCCGAGGAAGAGA
<i>Dwi</i> \GK15418						
Interg		69L				ACATTAATTCTTGCTGCCAG
<i>Dwi</i> \GK15416-	-	69R	69	+		GTGGTTACAGTATTGCCGTA
<i>Dwi</i> \GK15418						

Tabela II.7: continuação.

<i>Dwi</i> \GK15418	<i>Dme</i> \CG31913	52bL 52bR	52b	-	CAGATATCCTTATGTGGTGC CGGTCTGAATGTATGGCAAT
Interg <i>Dwi</i> \GK15418- <i>Dwi</i> \GK15419	-	65L 65R	65	-	TGATCGTGATCAAGAGCATA GAGAAGCCTCAATTAGATGG
<i>Dwi</i> \GK15419	<i>Dme</i> \CG9171	47L 47R	47	-	ACCTTGGATTCCATACGATA ATGACAACTGGCACAGACTC
<i>Dwi</i> \GK19065	<i>Dme</i> \CG14006	48L 48R	48	-	ATAAGGATATGGCTCAGAAG GCCACAGTGATATTGATATG
<i>Dwi</i> \GK14687	<i>Dme</i> \CG18266	39L 39R	39	-	TAGGCTCGATATTGGAGACC GAAGATGATTCGGAATGGCA
<i>Dwi</i> \GK19024	<i>Dme</i> \CG15414	40L 40R	40	-	AATGCAGAGCATCTGGTATG ATGACGCTTCAGAGATTGCC
<i>Dwi</i> \GK14671	<i>Dme</i> \CG3213	28L 28R	28	-	TCTCATTGGACACATAGGCA GGAACAGATGCTTCATCACC
<i>Dwi</i> \GK14536	<i>Dme</i> \CG5850	19L 19R	19	-	GCATTAACCACCTCGCTACC GGAATGGCGTCTATGGATAC

Apêndice I

Análises citogenéticas

- Soluções:

Solução fisiológica 0,9% NaCl:

0,9g de NaCl

100 mL de água destilada autoclavada

Orceína aceto-láctica 5%

5 g de orceína sintética

55 mL de ácido acético glacial

45 mL de ácido láctico

- Misturar todos os componentes por fervura em capela de fluxo laminar. Repousar a solução por 24 h e filtrar. Armazenar em temperatura ambiente em frasco protegido da luz.

SSC Denhardt

0,012 g polivinilpirrolidona

0,012 g ficoll 400

0,012 g albumina de soro bovino (Fração V)

90 mL 20x SSC

600 mL H₂O destilada

20x SSC - Solução Sódio citrato estoque

3 M NaCl, 300 mM Citrato de sódio

87,66 g NaCl

44,11 g Citrato de Sódio

500 mL H₂O ultrapura

Total: 50 mL – Autoclavar

2x SSC

200 ml 20x SSC

800 ml de H₂O ultrapura

Total: 1000 mL

Autoclavar

10x PBS - Tampão fosfato-salino estoque

1,3 M NaCl, 0,07 M Na₂HPO₄, 0,03 M NaH₂PO₄

18,993 g NaCl

2,48 g Na₂HPO₄

1,03 g NaH₂PO₄

250 mL H₂O ultrapura

PH 7,4

Autoclavar

1x PBS

100 mL 10xPBS

900 mL de água

Total: 1000 mL

Autoclavar

Dextran Sulfato 50% (w/v)

Em um tubo de 5 mL:

1 g de dextran sulfato

1,3 ml de H₂O destilada

Misturar por aquecimento e agitação por seis horas

Armazenar a 4°C

Tampão fosfato**40% Solução A****0,05 M NaH₂PO₄·H₂O**3,45 g NaH₂PO₄·H₂O

Completar para 500 mL

Autoclavar

60% Solução B**0,05M Na₂HPO₄·12H₂O**8,95 g Na₂HPO₄·12H₂O

Completar para 500 mL

Autoclavar

Tampão TAE (Tris-Acetato-EDTA)

Estoque 50x

242 g TRIS base (2-amino-2-hidroximetil-propano-1,3-diol)

57,1 mL ácido acético glacial

100 mL de solução de 0,5 M EDTA dissódico, pH 8.0

Água destilada até 1 litro

0,5 M EDTA, pH 8.0

18,612 g de EDTA (Etilenodiaminotetraacetato dissódico) em 80 mL de H₂O destilada.

Agitar vigorosamente

Ajustar para pH 8,0 com hidróxido de sódio (aproximadamente 20 g de NaOH)

Completar para 1000 mL

Esterilizar por autoclavagem.

- Tratamento das lâminas para Citologia clássica dos cromossomos politênicos

Limpeza por imersão em água com sabão líquido por 24 horas

Enxaguar abundantemente

Limpeza por imersão em etanol 92,8 GL.

Secar

Tratamento com cobertura de gelatina comercial incolor de cada lâmina

Secar a temperatura ambiente (TA)

Armazenar a 4°C para o uso.

- Tratamento das lâminas para Hibridação *in situ* não fluorescente (HIS)

Limpeza por imersão em água com sabão líquido por 24 horas

Enxaguar abundantemente

Limpeza por imersão em etanol 92,8 GL

Secar

Imersão em solução SSC Denhardt 1% a 65 °C por 2 a 3 horas

Imersão em solução de ácido acético:água na razão 3:1 (v/v) por 20 minutos

Imergir por 2 segundos, 10 vezes em etanol 95 GL

Secar

Armazenar a 4°C

Análises moleculares

- Extração de DNA genômico:

Em um microtubo de 2 mL:

Macerar aproximadamente 100 adultos em nitrogênio líquido

Adicionar 750 µL de solução de lise (Tris-HCl 0,1 M; EDTA 0,1 M; SDS 1% e NaCl 0,06 M) e homogeneizar por inversão manual

Incubar a 65 °C por 1 hora

Adicionar 1V de fenol (pH 8,0) e agitar lentamente por 10 minutos

Centrifugar a 13.000 rpm por 10 minutos

Transferir a fase aquosa para outro microtubo e acrescentar 1V de fenol:clorofórmio na razão 1:1 (v/v)

Agitar lentamente por 10 minutos

Centrifugar a 13.000 rpm por 10 minutos

Transferir o sobrenadante para outro microtubo e acrescentar 1V de clorofórmio.

Agitar lentamente por 10 minutos

Transferir o sobrenadante para outro microtubo e acrescentar 1V de isopropanol gelado, ou 2V de etanol absoluto gelado

Deixar a -20 °C por 16 horas

No outro dia, centrifugar a 13.000 rpm por 10 minutos

Desprezar a fase aquosa

Lavar o pellet com etanol 70 GL a 13.000 rpm por 2 minutos

Secar o pellet a 37°C

Ressuspender em 50-100 µL de H₂O ultrapura

Migrar 1 µL da extração de DNA genômico em gel de agarose 0,8% a 100V para verificação da qualidade da extração.

Alternativamente, para a verificação da qualidade da extração pode-se clivar este DNA com qualquer enzima de restrição. Extrações de boa qualidade, sem nenhum contaminante, apresentam resultado satisfatório para clivagem.

Clonagem

- Reação de ligação plasmidial

O vetor *pGEM -T Easy* em seu sítio *MCS* contém sequências que são reconhecidas pelas enzimas de restrição *NotI* e *EcoRI* em ambos os lados. É um vetor linearizado com um nucleotídeo dTTP 3' em ambos os terminais. Este último permite a clonagem de fragmentos obtidos diretamente pela técnica da *PCR*, sem a necessidade de um sítio de reconhecimento enzimático complementar. Para isto, necessita de uma *Taq polimerase* que origine uma porção 5'-dATP terminal complementar. Subsequente à esta região,

encontram-se sítios de complementaridade com os *primers* universais M13F e M13R-puc.

Tabela Apêndice I: Especificações da reação de ligação plasmidial do vetor *pGEM -T Easy*

Componete da Reação	Reação Padrão	Controle Positivo
<i>2X Rapid Ligation Buffer, T4 DNA Ligase</i>	5µL	5µL
<i>pGEM® -T Easy Vector (50ng)</i>	1µL	1µL
Produto da <i>PCR</i>	XµL*	-
Insero de DNA controle	-	2µL
<i>T4 DNA Ligase (3U Weiss/µl)</i>	1µl	1µL
H ₂ O completar até volume final	10µL	10µL

Xµl* = Neste trabalho utilizou-se 3 µL do produto da *PCR* sem completar com H₂O até o volume final.

- Meio de cultura bacteriana LB (*Lysogeny Broth*)

Meio LB líquido

2,5 g LB

100 mL H₂O

Autoclavar

Meio LB sólido

2,5 g LB

1,5 g ágar bacteriológico

100 mL H₂O

Autoclavar

- Bactérias já transformadas são cultivadas em meio líquido e sólido contendo o antibiótico ampicilina a 50 mg/mL. Aplicar 200 µL a cada 100 mL de meio.

- Preparação de células competentes:

Protocolo para 10 reações de ligação

Inocular a linhagem *TOP 10* de *E. coli* em 3 mL de meio LB líquido, sem ampicilina, a 37 °C pelo período de uma noite.

No outro dia, verter o cultivo de 3 mL em 30 mL de meio LB líquido, sem ampicilina

Deixar crescer sob agitação constante a 37 °C por aproximadamente 3 horas, até se obter uma coloração turva

Verter o cultivo em tubo falcon de 15 mL e centrifugar a 5.000 rpm por 5 minutos

Desprezar o sobrenadante

Repetir o passo acima até esgotar o cultivo

Ressuspender o *pellet* em 5 mL de 0,1 M MgCl₂ frio e estéril e homogeneizar por vórtex

Centrifugar a 5.000 rpm por 5 minutos e desprezar o sobrenadante

Ressuspender o *pellet* em 5 mL de 0,1 M CaCl₂ frio e estéril e homogeneizar por vórtex

Inocular no gelo por 30 minutos

Centrifugar a 5.000 rpm por 5 minutos e desprezar o sobrenadante

Ressuspender o *pellet* em 1 mL de 0,1 M CaCl₂ frio e estéril e homogeneizar por vórtex

- Transformação bacteriana:

O gene *LacZ*, um componente estrutural do Operon *Lac*, codifica a β -galactosidase, uma enzima intracelular que degrada o dissacarídeo lactose em glicose e galactose. O vetor *pGEM -T Easy* utiliza desta propriedade para oferecer diferenciação entre as bactérias transformadas, das bactérias não transformadas (ao mesmo tempo que também serve para a seleção de vetores

ligados ao fragmento de interesse, ou não). A inserção do vetor portando o fragmento de interesse complementa o gene da β -galactosidase presente na bactéria, um processo chamado α -complementação. Por sua vez, os vetores com as sequências de interesse ligadas têm seu gene *LacZ* corrompido, não permitindo a posterior complementação deste gene na bactéria hospedeira, o qual se encontra com uma porção deletada. A visualização disto se dá na coloração das colônias transformadas resultantes. Para isto, adiciona-se ao meio LB sólido *IPTG* (100 mM) e *X-gal* (50 mg/mL).

O *X-gal* é um agente cromogênico análogo à lactose. Ele é degradado pela β -galactosidase em 5-bromo-4-cloro-indoxyl, o qual espontaneamente oxida, resultando em uma coloração azul nas colônias originárias de bactérias não transformadas, e com o gene *LacZ* ativo. Por sua vez, as colônias originárias das bactérias transformadas com a sequência de interesse, uma vez que apresentam o gene *LacZ* corrompido, não irão processar o *X-gal*, resultando em colônias de coloração branca.

O *IPTG* (*Isopropyl β -D-1-thiogalactopyranoside*) é um análogo da galactose, o qual induz a expressão do gene *LacZ*, reforçando o processamento do *X-gal* pelo gene. A este mecanismo de seleção dá-se o nome de *Blue-white screen* (Sambroock & Russel, 2004).

Primeiro dia:

Acrescentar 100 μ L da bactéria ressuspensa em 0,1 M CaCl_2 à cada reação de ligação

Inocular por 30 minutos em gelo

Imediatamente inocular por 5 minutos a 42°C

Após, inocular por 5 minutos em gelo

Adicionar 400 μ L de meio LB líquido sem ampicilina em cada reação, e deixar crescer por 1 hora a 37°C com agitação

Após, plaquear com alça de Drigalsky 200 μ L de cada reação de transformação em placa de petri, contendo aproximadamente 25 mL de meio LB sólido com ampicilina e *IPTG+X-gal* (100 μ L de *IPTG* + 20 μ L de *X-gal*, que devem ser plaqueados à chama anteriormente).

Incubar as placas pelo período de uma noite em estufa a 37 °C

Segundo dia:

Repicar de cinco a 10 colônias que cresceram em outra placa de petri com divisões para controle.

Deixar as placas pelo período de uma noite em estufa a 37 °C

Terceiro dia:

Fazer *PCR* das colônias repicadas para comprovar se são positivas

Com uma ponteira de 10 µL autoclavada, tocar na colônia, e coletar junto um pouco de meio

Pipetar com esta ponteira em um tubo de 0,6 mL contendo 50 µL de H₂O ultrapura

Submeter a 98 °C por 10 minutos no termociclador, para explosão das colônias

Dar um spin

Utilizar 1 µL de cada colônia explodida na reação da *PCR*, utilizando *primers* universais, ou os *primers* específicos da sequência de interesse

Fazer extração do DNA plasmidial de alta qualidade das colônias comprovadas pela *PCR*

Extração de DNA plasmidial (miniprep)

Soluções (preparadas na hora)

- Solução de lise alcalina I

Para 5 mL:

400 µL de 0,1 M TRIS pH 8.0

100 µL de 0,5 M EDTA pH 8.0

4,5 mL de H₂O ultrapura

- Solução de lise alcalina II**0,1 M NaOH, 1% (w/v) SDS****NaOH 10 M – solução estoque (armazenar em pote plástico)**

40 g de NaOH

Completar o volume até 100 mL de H₂O**NaOH 1 M**

100 µL de NaOH 10 M

900 µL de H₂O ultrapura**SDS 10% - solução estoque**

10 mg de SDS

Completar para 100 mL com H₂O ultrapura

Homogeneizar bem

Para 5mL:

1 mL de 1M NaOH

500 µL de SDS 10%

3,5 mL de H₂O ultrapura**- Solução de lise alcalina III**

3 M potássio, 5 M acetato

5 M acetato de potássio - 60 mL (29,44 g de acetato de potássio em 60 mL)

Ácido acético glacial - 11,5 mL

H₂O - 28,5 mL

Inocular uma colônia da bactéria transformada em até 2 mL de meio LB líquido com ampicilina pelo período de uma noite a 37 °C

No outro dia centrifugar a 13.000 rpm por 3 minutos

Retirar o sobrenadante de meio LB líquido

Ressuspender o *pellet* de bactérias com 100 µL de Solução de lise alcalina e dar um vigoroso vórtex

Adicionar 200 μL de Solução de lise alcalina II e homogeneizar por inversão cinco vezes (observação: esta etapa não pode extrapolar 5 minutos)

Adicionar 150 μL de Solução de lise alcalina III gelada e homogeneizar por inversão. Armazenar em gelo por cinco minutos

Centrifugar a 13.000 rpm por 5 minutos

Passar o sobrenadante para outro tubo

Adicionar 1V de isopropanol, ou 2V de etanol absoluto gelado

Armazenar a -20°C por pelo menos 16 horas

Centrifugar a 13.000 rpm por 5 minutos

Secar o *pellet* com o tubo aberto em estufa a 37°C

Ressuspender o pellet em 50-100 μL de H_2O ultrapura

Adicionar 1-1,5 μL de RNAse A a 50 mg/mL

Incubar em banho-maria a 37°C por 1 hora

Migrar 1 μL da extração plasmidial em gel de agarose 0,8% a 75V para comprovação e verificação da qualidade da extração.

CAPÍTULO III

III - Resultados e Discussão

III.1 – Estudo dos pontos de quebra de inversões de *Drosophila willistoni*.

I. Comparação do arranjo cromossômico do braço IIL entre a linhagem sequenciada Gd-H4-1 e o fotomapa da espécie.

A análise do padrão de bandas do braço IIL da linhagem de *D. willistoni* sequenciada Gd-H4-1 (*Drosophila* 12 Genomes Consortium, 2007), comparada com o padrão de bandas do fotomapa estabelecido para a espécie (Schaeffer *et al.*, 2008) apontou a presença de duas inversões fixadas na linhagem Gd-H4-1: Inversão IIL-A e Inversão IIL-F. De acordo com o fotomapa, a inversão IIL-A abrange as seções 38B-39C. Já a inversão IIL-F engloba as seções 50C-53A (**Figura III.1.1**) (Rohde & Valente, 2012).

Drosophila willistoni teve sua primeira linhagem padrão estabelecida por Dobzhansky (1950). Neste estudo, o autor forneceu um mapa cromossômico dos cromossomos politênicos da espécie, desenhados com câmara clara. Esta linhagem foi obtida a partir de uma única fêmea capturada perto da cidade de Belém, Pará (Brasil). Assim como foram estabelecidas uma linhagem padrão para *D. melanogaster* (Bridges, 1935), e *D. pseudoobscura* (Dobzhansky & Sturtevant, 1938), uma linhagem protótipo para *D. willistoni* seria usada para fins de estudos populacionais, e da história evolutiva de suas inversões cromossômicas. O primeiro mapa cromossômico de *D. willistoni*, com base no padrão de bandas da linhagem de Belém, foi dividido em 100 seções arbitrárias (1-16 no braço XL, 17-36 no braço XR, 37-55 no braço II-L, 56-77 no braço II-R e 78-100 no cromossomo III), e nesse mesmo trabalho, características inerentes de cada cromossomo foram descritas (Dobzhansky, 1950).

O estudo subsequente de Da Cunha *et al.* (1950) analisou o polimorfismo cromossômico de *D. willistoni* em 20 localidades ao longo da distribuição geográfica da espécie na América do Sul, mais duas linhagens da América Central (Costa Rica e México). Este estudo analisou a F1 de fêmeas capturadas na natureza, contrastando os polimorfismos encontrados com o mapa cromossômico estabelecido no estudo anterior (Dobzhansky, 1950), e procurou estabelecer os limites dessas inversões de acordo com as seções do mapa disponível. Conforme esta análise, as inversões heterozigotas IIL-A e IIL-F apresentaram ampla ocorrência ao longo da distribuição geográfica da espécie. A linhagem de Porto Rico, região geográfica mais próxima da Ilha de Guadalupe no Caribe (a qual é o local de origem da linhagem Gd-H4-1), apresentou essas inversões em heterozigose nas frequências de 10% e 40% respectivamente.

a)



b)

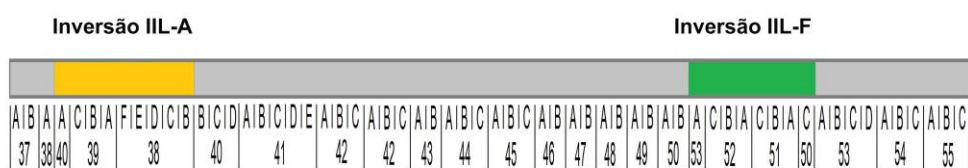


Figura III.1.1: Análise do arranjo cromossômico da linhagem sequenciada Gd-H4-1 de *Drosophila willistoni*. **a)** Fotomapa do braço IIL de *Drosophila willistoni* (Schaeffer *et al.*, 2008). **b)** Esquematização do arranjo cromossômico do braço IIL da linhagem sequenciada Gd-H4-1, mostrando a reorganização das seções e subseções com relação ao fotomapa. Em destaque as seções envolvidas com as inversões IIL-A (amarelo) e IIL-F (verde).

A análise por cruzamentos entre as linhagens amostradas com a linhagem padrão de Belém apontou que as populações de *D. willistoni* têm arranjos fixados distintos, principalmente no cromossomo X, que se relacionam com as distintas regiões geográficas. De acordo com os dados de polimorfismo obtidos, os autores inferem que *D. willistoni* possui maior polimorfismo para inversões heterozigotas na região central do Brasil, com considerável decréscimo destas frequências em direção às periferias da distribuição. Isso se enquadra no conceito de centro de origem de Vavilov (1926), no qual as áreas com maior variação genética de um determinado organismo representam a antiguidade da ocupação deste lugar; e geralmente constituem o centro de origem destes (Da Cunha *et al.*, 1950).

Uma extensão da análise de Da Cunha *et al.* (1950), feita por Da Cunha *et al.* (1954), aumentou a amostragem do polimorfismo de *D. willistoni* para outras regiões do sul do Brasil, Uruguai e Argentina, somando estes resultados aos previamente estabelecidos. Este estudo classificou as inversões polimórficas da espécie em três categorias de distribuição: raras/endêmicas, restritas, e de ampla distribuição. As inversões IIL-A e IIL-F foram classificadas como sendo polimorfismos de ampla distribuição; a primeira ocorrendo desde a Flórida até o Sul do Brasil, e a segunda inversão ocorrendo desde Cuba (América Central) até a Argentina.

A classificação das inversões IIL-A e IIL-F como pertencentes à categoria de ampla distribuição geográfica teve no decorrer histórico da caracterização do polimorfismo cromossômico de *D. willistoni* grande embasamento, oriundo principalmente de análises de polimorfismos cromossômicos das regiões marginais da distribuição da espécie, tais como os de distintas ilhas da América Central, México e sul dos Estados Unidos (Dobzhansky, 1957; Townsend, 1952; 1958; Ayala *et al.*, 1971), e do extremo sul da América Latina (Valente & Araújo, 1986; Valente *et al.*, 2001; 2003).

A ampla distribuição geográfica da inversão IIL-F pode ser atribuída à maior adaptabilidade de seus portadores a diferentes componentes do valor adaptativo, como escolha de vários levedos (Da Cunha, 1957), diferentes

substratos (Valente & Araújo, 1985), entre outros. O mesmo se pode inferir sobre a inversão IIL-A (revisão em Rohde & Valente, 2012).

O estudo de Regner *et al.* (1996) inseriu *D. willistoni* no instigante contexto acerca da origem de suas inversões, uma vez que foi pioneiro no estabelecimento da associação do sítios de ocorrência do elemento *P*, com diferentes pontos de quebra de inversões presentes nesta espécie. Como parte deste estudo, estabeleceu-se a construção do primeiro fotomapa de *D. willistoni*, escolhendo como padrão o arranjo cromossômico da linhagem WIP-4, originária do estado da Bahia (Brasil); uma vez que a linhagem padrão estabelecida por Dobzansky (1950), já não era mantida em nenhum outro centro de pesquisa. A linhagem WIP-4, mantida desde a década de 1960 no Laboratório de *Drosophila* da UFRGS, mostrou-se com o arranjo cromossômico similar à linhagem de Belém (Pará), em comparações feitas com o mapa desta linhagem. Também mostrou baixo nível de inversões segregantes, bem como material citológico de ótima qualidade para as análises citogenéticas, trazendo com isto muitas vantagens, uma vez que os cromossomos politênicos de *D. willistoni* são reconhecidamente um material de difícil manejo e análise.

O aprimoramento do fotomapa de *D. willistoni*, feito por Regner *et al.* (1996), ocorreu em virtude de uma ampla revisão do polimorfismo cromossômico da espécie desenvolvido no grupo de pesquisa do Laboratório de *Drosophila* da UFRGS. Esse fotomapa foi primeiramente apresentado subsequente ao sequenciamento do genoma de *D. willistoni* (*Drosophila* 12 Genomes Consortium, 2007), onde se estabeleceu a ordem e a orientação dos *scaffolds* deste genoma, com os respectivos cromossomos politênicos (Schaeffer *et al.*, 2008). O fotomapa aprimorado da espécie em questão corresponde aos arranjos cromossômicos com maior frequência nas populações de *D. willistoni*, ao longo da sua distribuição geográfica; bem como acrescentou o estabelecimento de subseções, objetivando proporcionar uma maior precisão na determinação física dos limites das inversões cromossômicas e estudos associados (Rohde & Valente, 2012).

Entretanto, o arranjo cromossômico do fotomapa aprimorado de *D. willistoni* difere da linhagem padrão tradicional e atual da espécie, WIP-4, no

seu braço cromossômico IIL, uma vez que essa linhagem apresenta as inversões IIL-I e IIL-F fixadas com relação ao fotomapa. Por sua vez, como já exposto acima, o padrão de bandas do fotomapa não corresponde ao padrão de bandas da linhagem sequenciada Gd-H4-1, no que concerne ao braço cromossômico IIL (o qual é o foco deste estudo), e também pela ocorrência de uma inversão no braço cromossômico XL (Inversão XL-C). Já os braços cromossômicos XR e IIR, e o cromossomo III apresentam o mesmo padrão citológico. Sumarizando, a fórmula básica do braço cromossômico IIL na linhagem WIP-4 é IILFI, e da linhagem sequenciada é IILAF.

Assim, é oportuno comentar que devido à aleatoriedade da escolha das linhagens-padrão (Sperlich & Pfriem, 1986) e à dificuldade de mantê-las por muitas gerações em bancos de estoques sob condições de cultivo variáveis, recomenda-se cautela nas generalizações. É recomendável também, que cada linhagem de *Drosophila* a ser sequenciada tenha sua constituição cromossômica cuidadosamente avaliada, e seus arranjos determinados. Isto mostra que a moderna citogenômica depende de um prévio, sólido e concomitante conhecimento da citogenética.

A presente análise passará a considerar o arranjo cromossômico da linhagem Gd-H4-1 como padrão para o estudo dos pontos de quebra das inversões em *D. willistoni*, de maneira que esta estratégia seja reproduzível em outros estudos desta natureza. Para isto, fez-se necessário, primeiramente, estabelecer um correto alinhamento entre o padrão citológico presente no fotomapa da espécie (Schaeffer *et al.*, 2008; Rohde & Valente, 2012) com o padrão ocorrente na linhagem sequenciada Gd-H4-1; especificamente neste estudo, referente ao braço cromossômico IIL (**Figura III.1.1**).

Sendo assim, qualquer análise dos pontos de quebra de inversões se dará, primeiramente, por meio do estabelecimento dos genes flanqueadores dos pontos de quebra distal e proximal. Entretanto, esse estabelecimento necessitará ser determinado na linhagem sequenciada, e em outra linhagem de *D. willistoni* que contenha em homozigose a inversão em foco. As inversões ocorrentes nessa outra linhagem podem ser inferidas por meio de cruzamentos com a linhagem Gd-H4-1, e as alças de inversões heterozigotas serão

consideradas como diferenças ocorrentes no arranjo cromossômico da linhagem que venha a ser analisada, com relação à linhagem sequenciada Gd-H4-1, a qual será o padrão.

Esta estratégia de análise se baseia na metodologia inferida por Sturtevant & Dobzhansky (1936) para as análises das tríades de inversões sobrepostas, com vistas à elucidação das relações filogenéticas entre essas inversões. Porém, devido à época relacionada, as análises das tríades de inversões sobrepostas levavam em consideração somente as diferenças citológicas ocorrentes entre a linhagem padrão e a outra linhagem em análise, e assim a direção evolutiva da origem destas inversões não podia ser estabelecida. Já no estudo dos pontos de quebras de inversões, a análise molecular da região envolvida na inversão pode vir a elucidar a idade da inversão, e no caso das inversões sobrepostas, pode vir a elucidar a idade e a direção evolutiva destas.

Análises dos pontos de quebras de inversões com base nesta estratégia vêm sendo feitas nos estudos envolvendo os pontos de quebra de inversões em *D. buzzatii* (Cáceres *et al.*;1999; Casals *et al.*, 2003; Delprat *et al.*, 2009; Calvete *et al.*, 2012) e *D. subobscura* (Papaceit *et al.*, 2012; Puerma *et al.*, 2014).

III.2 – Estudo dos pontos de quebra de inversões de *Drosophila willistoni*.

II. Determinação dos diferentes arranjos cromossômicos do braço III entre as linhagens Gd-H4-1 e SG12.00 (do Uruguai) por meio de cruzamentos.

O estudo dos pontos de quebra de inversões em *Drosophila* consiste na delimitação dos pontos de quebra da inversão de interesse em uma linhagem padrão, e em outra linhagem, da mesma espécie, com a inversão em análise fixada; mesmo que na natureza sua ocorrência seja predominantemente segregante. Isto porque, o uso de linhagens portando a inversão a ser delimitada em heterozigose dificulta o mapeamento físico por meio de caminhada cromossômica, bem como a análise molecular das regiões dos pontos de quebra. Como uma inversão se origina a partir de duas quebras na fita dupla, e a subsequente religação da sequência entre essas quebras na ordem invertida, seus pontos de quebra são classificados em: ponto de quebra distal (relacionado ao telômero) e ponto de quebra proximal (relacionado ao centrômero) (Powell, 1997).

Para a inferência do possível mecanismo biológico envolvido na gênese de uma inversão é necessário ter acesso, em nível molecular, à sequência das regiões dos pontos de quebra distal e proximal delimitadas. Uma vez que estas sequências fornecerão as informações necessárias para as análises de mecanismo de origem, idade da inversão e direção evolutiva; será feita uma comparação destas regiões em ambas as linhagens. Esse acesso, quando envolve uma sequência de geralmente até 4 kb, pode ser obtido pela técnica da *PCR*, clonagem e sequenciamento. Entretanto, na maioria das vezes, as regiões dos pontos de quebra envolvem sequências de tamanho e características inacessíveis pela *PCR*, e os sistemas de amplificação de sequências maiores (acima de 10 kb) geralmente obtém pouco sucesso. Sendo assim, o acesso aos pontos de quebra se obtém, com maior precisão, ou via a construção de bibliotecas genômicas, ou de genomas montados disponíveis (Calvete *et al.*, 2012; Corbett-Detig, 2012).

Na presente análise, afora a linhagem Gd-H4-1, a linhagem de *D. willistoni* escolhida para o estudo dos pontos de quebra da inversão IIL-H (a qual será o foco deste estudo) foi a SG12.00. Esta linhagem se estabeleceu a partir de uma fêmea que apresentava a mutação cardinal, a qual emergiu de uma pequena população denominada SG2.00, coletada em Montevideo, Uruguai (Soler & Goñi, 2012). Análises citológicas da linhagem SG12.00 (Goñi & Valente, 2012) comparadas ao arranjo cromossômico do fotomapa de *D. willistoni* (Schaeffer *et al.*, 2008) apontaram que essa linhagem portava a inversão IIL-H fixada, bem como as inversões III-J, III-B e III-A segregantes. A disponibilidade de uma linhagem de *D. willistoni* com a inversão IIL-H fixada, e com boa viabilidade, fornece um excelente material para as análises dos pontos de quebra de inversões que ocorrem naturalmente.

Uma metodologia amplamente utilizada nos estudos dos arranjos cromossômicos fixados em *Drosophila* consiste em cruzar, geralmente machos de uma população (ou linhagem) que se tenha interesse, com fêmeas de uma linhagem da mesma espécie (ou de outra espécie cujo cruzamento resulte em híbridos interespecíficos), para a qual se conheça o padrão citológico, na maioria das vezes sendo esta a linhagem referência da espécie (linhagem padrão). As alças de inversões que se originam nos diferentes cromossomos da F1 destes cruzamentos mostram, justamente, as diferenças (inversões, bem como outras alterações cromossômicas estruturais) do padrão cromossômico ocorrentes na linhagem que se está analisando (Krimbas & Powell, 1992).

Dada a caracterização e o estabelecimento da linhagem sequenciada Gd-H4-1 como a padrão para os estudos dos pontos de quebra de inversões em *D. willistoni* (**Tópico III.1** deste capítulo), as diferenças do arranjo cromossômico do braço IIL da linhagem SG12.00, respectivo àquela linhagem padrão, foram estabelecidas por meio de cruzamentos recíprocos entre machos e fêmeas virgens de ambas as linhagens (♀ SG12.00 x ♂ Gd-H4-1 e ♀ Gd-H4-1 x ♂ SG12.00).

A análise de dois indivíduos de cada cruzamento mostrou que a linhagem SG12.00 apresenta as inversões IIL-A, IIL-F e IIL-H fixadas com relação à linhagem padrão Gd-H4-1 (**Figura III.2.1** e **Figura III.2.2**).

Conforme foi apresentado no **Tópico III.1** deste capítulo, a inversão IIL-F em estado heterozigoto apresenta ampla distribuição geográfica entre as populações de *D. willistoni*. Diferentes estudos têm apontado sua associação com variáveis adaptativas (Valente & Araújo, 1985; 1986). Esta inversão compreende as seções 50C e 53A (Figura III.1.1). Por sua vez, a inversão polimórfica IIL-H é mais restrita às populações situadas em latitudes meridionais da América do Sul (Rohde & Valente, 2012), abrangendo as seções 53A e 55C. Subentende-se que estas inversões sejam monofiléticas em suas origens.

Classicamente, essas duas inversões são conhecidas por compartilharem citologicamente um dos pontos de quebra, no caso o ponto de quebra distal da inversão IIL-F e o ponto de quebra proximal da inversão IIL-H.

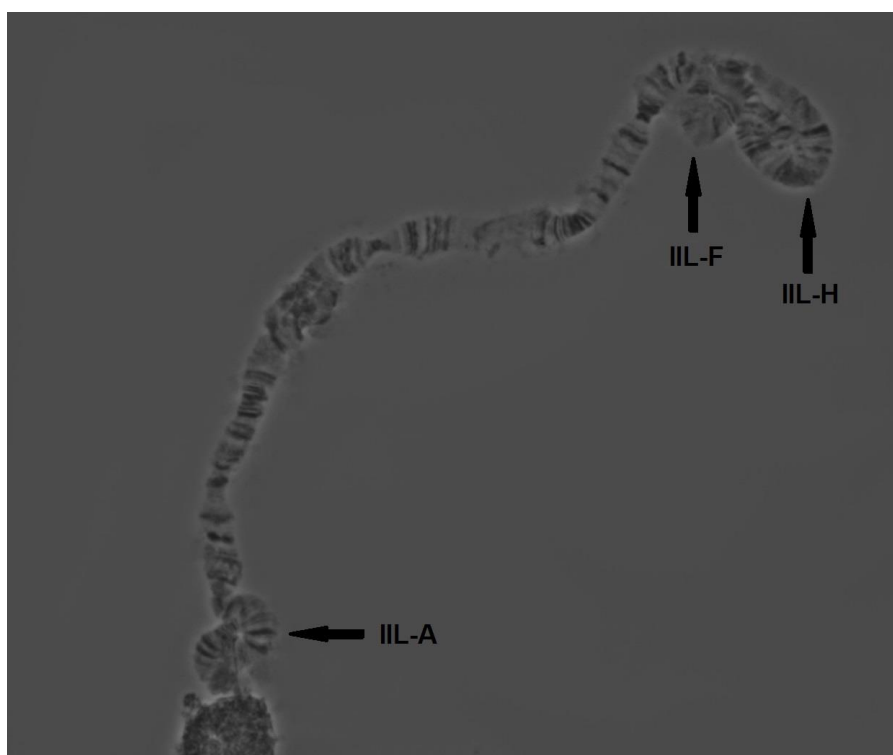


Figura III.2.1: Arranjo cromossômico do braço IIL resultante do cruzamento entre as linhagens ♀ Gd-H4-1 x ♂ SG12.00 de *Drosophila willistoni*. Setas indicam os arranjos fixados da linhagem SG12.00: Inversões IIL-A, IIL-F e IIL-H.

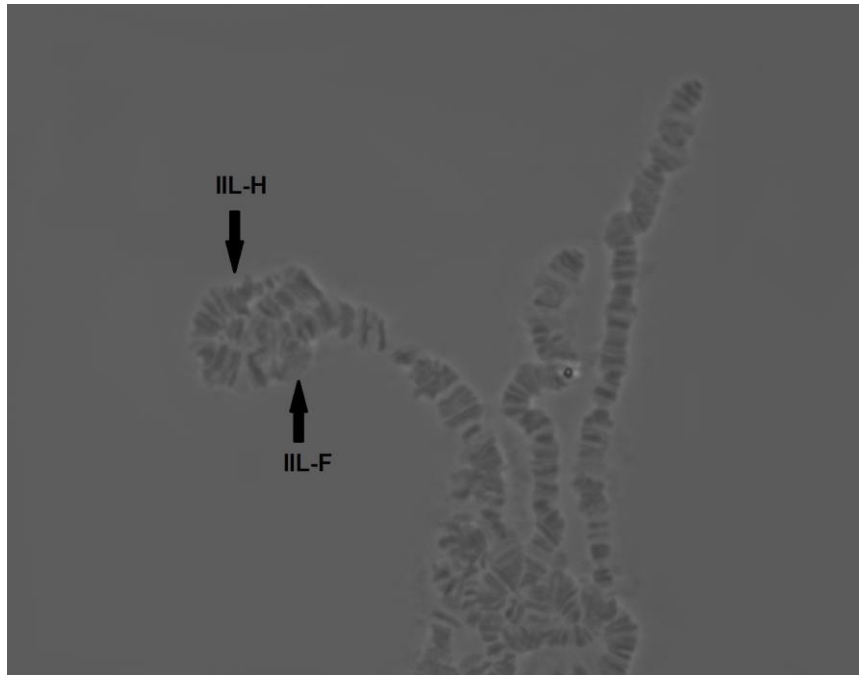


Figura III.2.2: Arranjo cromossômico do braço IIL resultantes do cruzamento entre as linhagens ♀ SG12.00 x ♂ Gd-H4-1 *Drosophila willistoni*. Setas indicam os arranjos fixados da linhagem SG12.00: Inversões IIL-F e IIL-H.

Os resultados citológicos dos cruzamentos entre a linhagem sequenciada e a linhagem SG12.00 apresentados nesta análise, conforme já constatado por Valente & Araújo (1986) e Rohde & Valente (2012), corroboram o reuso em nível citológico do ponto de quebra situado na seção 53A por ambas as inversões IIL-F e IIL-H (**Figura III.1.1**), característica que pode implicar reuso em nível molecular (Pevzner & Tesler, 2003). A **Figura III.2.3** mostra uma representação esquemática bidimensional do duplo heterozigoto IIL-FH, de acordo com as representações clássicas de polimorfismos cromossômicos (Dobzhansky & Pavlovsky, 1955; revisão em Sperlich & Pfriem, 1986). Entretanto, esta representação é apenas uma aproximação de uma situação biológica envolvendo estas duas inversões, sendo o comportamento do ponto de quebra compartilhado pelas duas inversões na seção 53A, ainda não totalmente esclarecido em níveis gráficos.

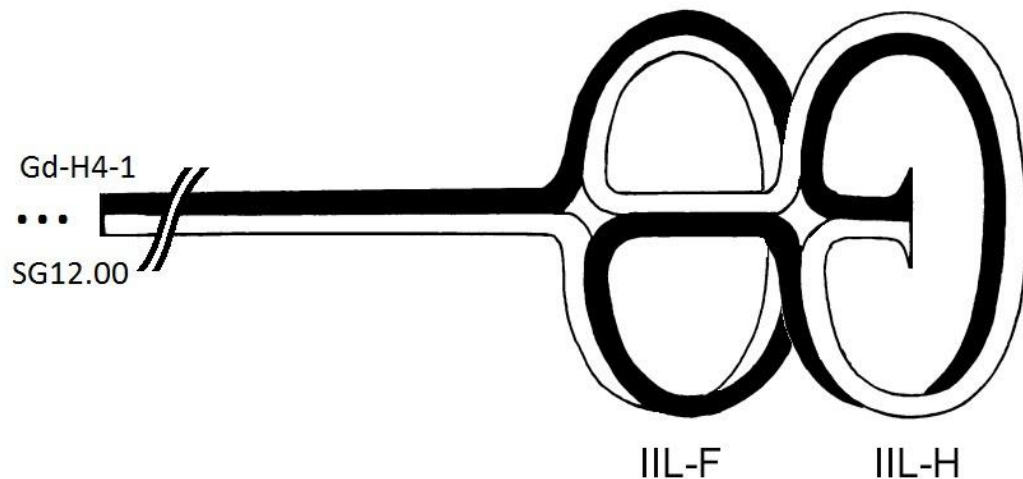


Figura III.2.3: Representação esquemática das alças de inversões da dupla heterozigose (Inversões IIL-F e IIL-H) decorrentes dos cruzamentos recíprocos entre a linhagem sequenciada Gd-H4-1 e a linhagem SG12.00 de *Drosophila willistoni*. Retirado e modificado de Dobzhansky & Pavlovsky (1955).

Casos de reuso em nível citológico foram corroborados em nível molecular nas análises das inversões sobrepostas $2m$ e $2n$ de *D. buzzatii* (Calvet *et al.*, 2012) e E_1 e E_2 em *D. subobscura* (Puerma *et al.*, 2014). Na presente análise, objetivamos delimitar os pontos de quebras da inversão IIL-H, porém, o conhecimento de que esta inversão divide citologicamente seu ponto de quebra proximal com a inversão IIL-F na linhagem SG12.00, nos antecipa a necessidade de uma maior cautela na análise dessa região. Entretanto, esta peculiaridade da linhagem SG12.00 disponível para nosso estudo, também pode vir a nos proporcionar conhecimentos adicionais acerca do reuso dos pontos de quebras de inversões em *Drosophila*.

III.3 – Estudo dos pontos de quebra de inversões de *Drosophila willistoni*.

III. Redefinição da ordem e da orientação dos *scaffolds* estabelecidos na montagem do genoma do cromossomo II de *Drosophila willistoni*.

Observação: Este estudo foi publicado na revista *G3: Genes, Genomes and Genetics*, e o artigo publicado encontra-se no **Capítulo V** da presente Tese.

O genoma de *D. willistoni*, sequenciado pelo Consórcio *Drosophila* 12 Genomes Consortium (2007), teve sua montagem de seus *scaffolds* alinhada à ordem citológica dos cromossomos politênicos correspondentes, com base na ortologia dos genes conservados em sintonia nas porções terminais dos *scaffolds*; e também de acordo com os marcadores genéticos e físicos que se tinha estabelecido para a espécie. Estes *scaffolds* foram numerados de 1 (associado ao centrômero) até n (por exemplo *scaffold* 8 no braço cromossômico IIL, associado ao telômero) (Schaeffer *et al.*, 2008)

Com isso, obteve-se a junção de cinco *scaffolds* para o braço cromossômico XL, de acordo com os marcadores genéticos dos genes *N* (*Notch*), *ct* (*cut*), *sn* (*singed*), *f* (*forked*), *sc* (*scute*), *y* (*yellow*), *lz* (*lozenge*) e *w* (*white*) (Spassky & Dozhansky, 1950), e a orientação baseada no marcador físico do gene *br* (*broad*). A montagem do braço cromossômico XR resultou na junção de oito *scaffolds*, com base nos marcadores genéticos dos genes *Adk-2* (*Adenylate kinase 2*), *Est-7* (*Esterase P*), *Idh* (*Isocitrate dehydrogenase*), *Pgm-1* (*phosphoglucose mutase*), *sh* (*short*), *ru* (*rughoide*) (Lancefield & Metz, 1922; Ferry *et al.*, 1923; Spassky & Dozhansky, 1950), e nos marcadores físicos dos genes *E74* (*Ecdysone-induced protein 74EF*), *E75* (*Ecdysone-induced protein 75B*), *Sod* (*Superoxide dismutase*), *Hsp83* (*Heat shock protein 83*) e *Hsp27* (*Heat shock protein 27*), que assim como o marcador *br* no braço XL, foram estabelecidos por nosso grupo de pesquisa (Bonorino *et al.*, 1993; Rohde *et al.*, 1994;1995; Rieger, 1999).

O cromossomo II teve o genoma do braço IIL, constituído por oito *scaffolds*, montado com base nos marcadores genéticos dos genes *px* (*plexus*) e *bw* (*brown*) (Spassky & Dozhansky, 1950); já o braço IIR, também composto por oito *scaffolds*, foi montado com base nos marcadores genéticos *hk* (*Hoock*), *Gpdh* (*Glycerol 3 phosphate dehydrogenase*) e *Adh* (*Alcohol dehydrogenase*), e também com base no marcador físico do gene *Adh* (Rohde *et al.*, 1995) e *Cl* (*Clippeed*) (Schaeffer *et al.*, 2008).

Já o cromossomo III resultou na montagem genômica de três *scaffolds*, de acordo com os marcadores físicos do loco *Xdh* (*Xanthine dehydrogenase*), e dos genes *ank* (*ankyrin*), *ct^D* (*cubitus interruptus Dominant*) e *ey* (*eyeless*) (Papaceit & Juan, 1998).

É extremamente notória a carência de marcadores gênicos em *D. willistoni*, sendo este quesito ainda mais saliente para o cromossomo II desta espécie (**Figura III.3.1**). Isso tem implicação direta na interpretação das análises genômicas. Em vista disso, a fim de comprovar a ordem e a orientação dos *scaffolds* nesse cromossomo, e dar uma correta base inicial à caminhada cromossômica para a delimitação dos pontos de quebra da inversão IIL-H, foram estabelecidas sondas, mais especificamente, uma sonda para cada um dos *scaffolds* do cromossomo II de *D. willistoni*, a fim de determinar a localização física por *HIS* destas sequências nos cromossomos politênicos da linhagem Gd-H4-1.

Essas sondas foram sequenciadas diretamente dos seus produtos da *PCR*, pela empresa de sequenciamento Macrogen Inc. (Seoul, Korea). As sequências consensos obtidas pela análise dos cromatogramas foram submetidas à BLASTN nos bancos de dados *Flybase* (St. Pierre *et al.*, 2014) e *NCBI* (Geer *et al.*, 2010), comprovando em 100% dos casos que essas sondas correspondiam às regiões gênicas escolhidas para cada *scaffold*.

De maneira extremamente surpreendente, todos os marcadores físicos estabelecidos para os *scaffolds* do braço IIL hibridaram no braço cromossômico IIR, ou seja, exatamente o contrário do esperado. Sendo assim, fez-se necessário estabelecer sondas para cada um dos *scaffolds* do braço IIR, e de maneira já esperada, e mesmo assim surpreendente; todas as sondas

hibridaram no braço IIL. Soma-se a isto o fato de que as sondas do braço IIR hidridaram em uma ordem reversa com relação à ordem proposta dos *scaffolds*, estando a sonda do *scaffold* 8 relacionando-se ao centrômero, e a sonda do *scaffold* 1 relacionando-se ao telômero. Já as sondas que hibridaram no braço III estavam de acordo com a ordem estabelecida (Schaeffer *et al.*, 2008) do centrômero (1) ao telômero (8).

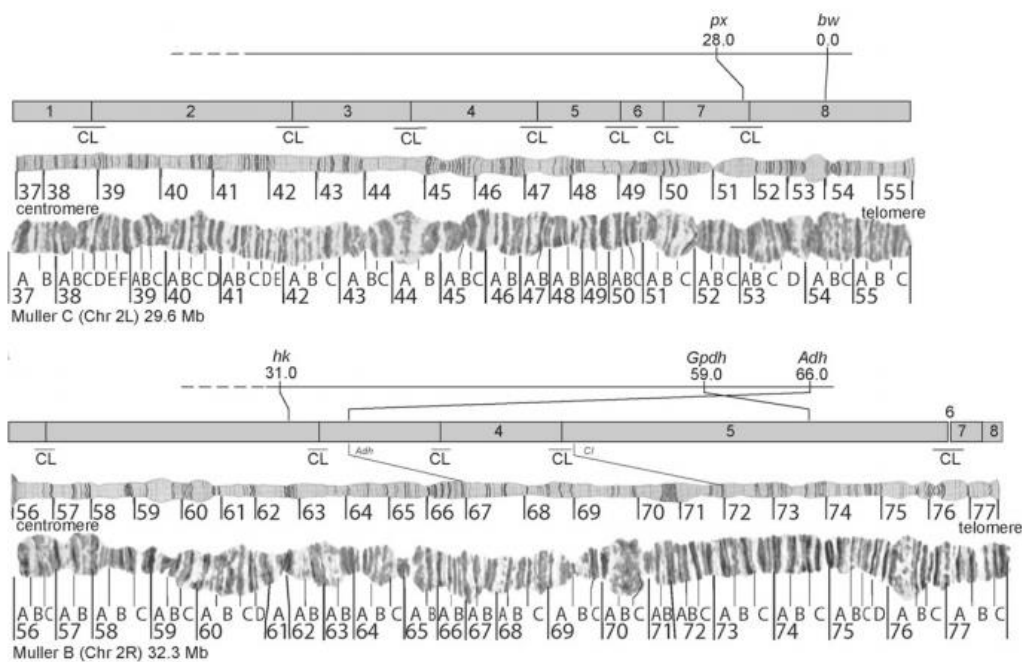


Figura III.3.1: Esquematização da associação dos *scaffolds* do genoma do cromossomo II de *Drosophila willistoni*, com os respectivos cromossomos politênicos, com base nos marcadores disponíveis. No topo da figura o braço IIL, na base da figura o braço IIR. Note que os *scaffolds* estão ordenados de 1 (centrômero) até 8 (telômero). Retirado de Schaeffer *et al.* (2008).

Subsequentemente, foram estabelecidas sondas para os genes *Dwil/GK23049 (bw)*, *Dwil/GK22138 (px)* e *Adh*, devido ao fato de esses terem sido praticamente os únicos marcadores genéticos (afora o gene *C*), que serviram de base para a montagem do genoma do cromossomo II. Assim sendo, elas serviriam, neste contexto, para o reforço e controle da conclusão final dos resultados. Estas hibridações resultaram nos genes *bw* e *px* mapeando fisicamente na base do braço cromossômico IIR, nas subseções 56C (**Figura III.3.2**) e 59B (**Figura III.3.3**) respectivamente; sendo que esses marcadores gênicos foram considerados estando associados ao telômero no braço III L (**Figura III.3.1**). Por sua vez, a sonda do gene *Adh* (**Figura III.4**) mapeou fisicamente na subseção 44A do braço cromossômico III L.

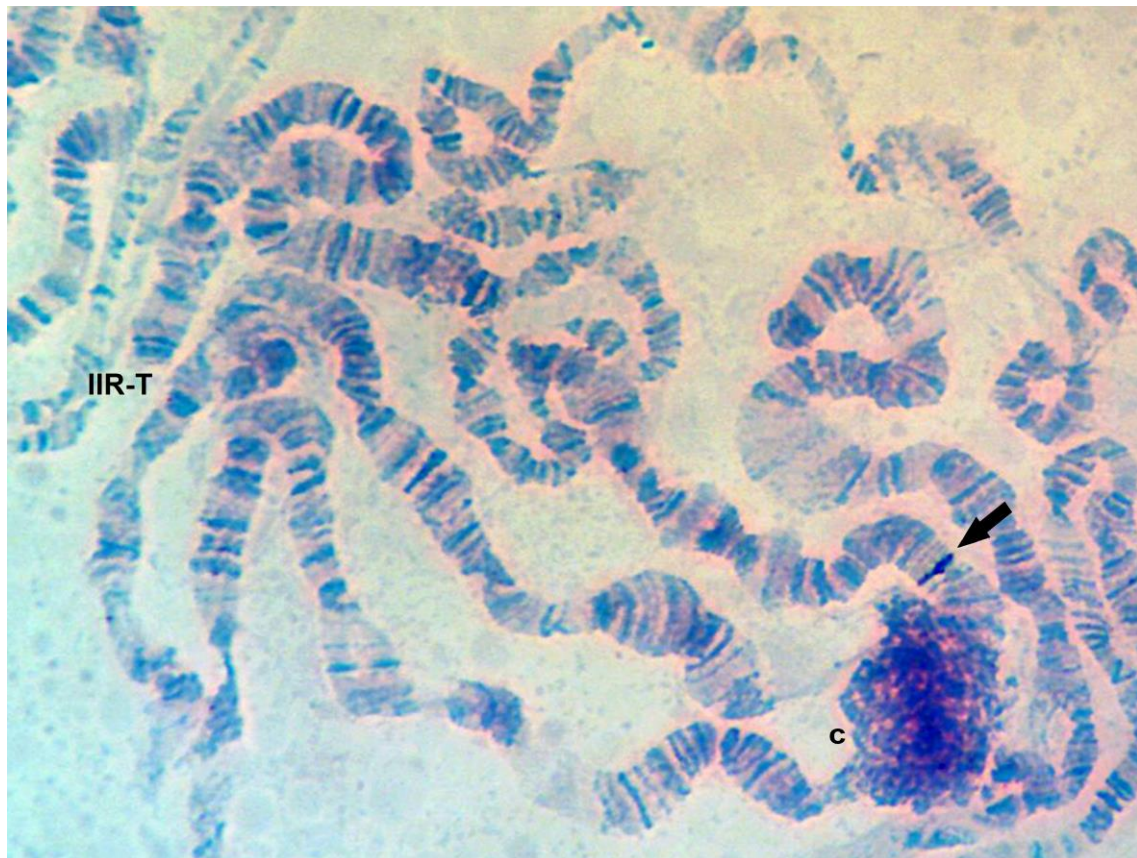


Figura III.3.2: Localização *in situ* da sonda do gene *bw* (*scaffold* 4954) em *Drosophila willistoni*. Seta indica o sítio de hibridação na subseção 56C do braço cromossômico IIR. **IIR-T:** Região telomérica do braço IIR. **C:** cromocentro.

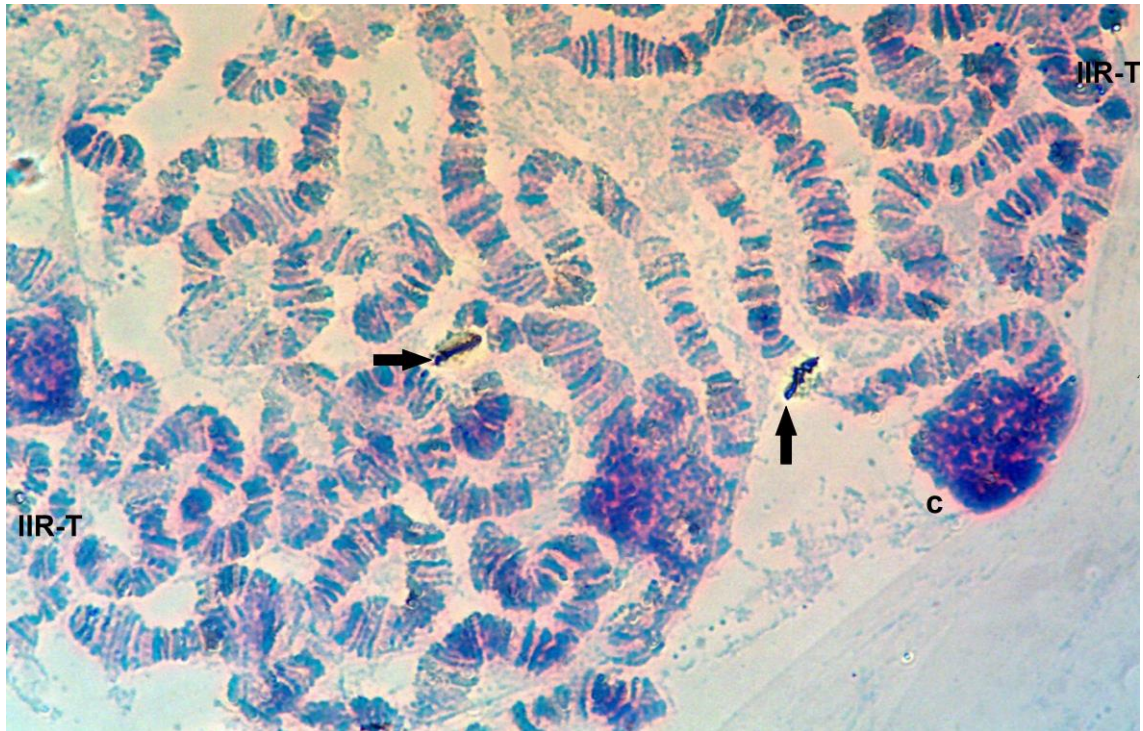


Figura III.3.3: Localização *in situ* da sonda do gene *px* (*scaffold* 4558) em *Drosophila willistoni*. Setas indicam o sítio de hibridação na subseção 59B do braço cromossômico IIR. **IIR-T:** Região telomérica do braço IIR. **C:** cromocentro.

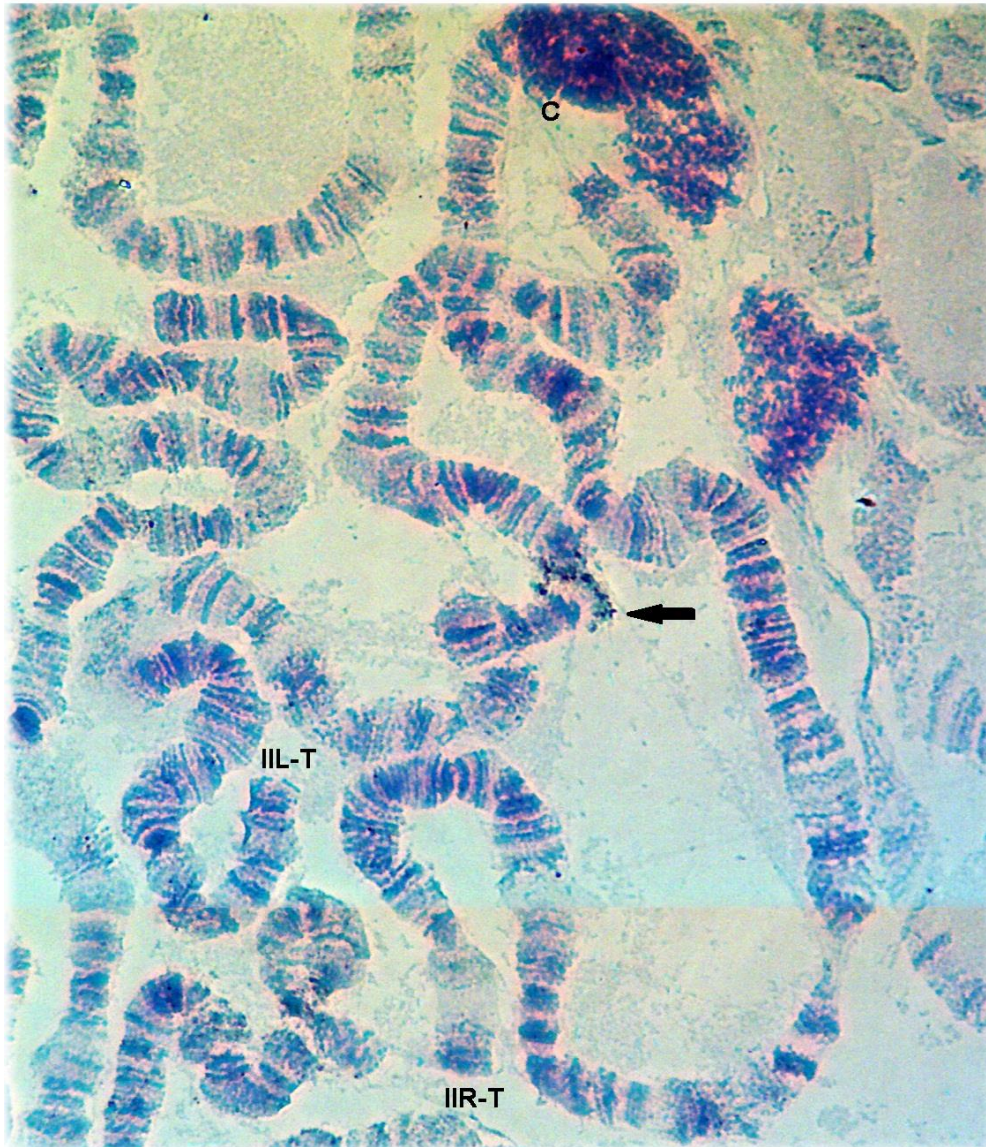


Figura III.3.4: Localização *in situ* da sonda do gene *Adh* (*scaffold* 4577) em *Drosophila willistoni*. A figura mostra os braços IIL e IIR intactos para análise. Seta indica sítio de hibridação na subseção 44A do braço cromossômico IIL. **IIL-T:** Região telomérica do braço IIL. **IIR-T:** Região telomérica do braço IIR. **C:** Cromocentro.

A **Tabela III.3.1** apresenta as sondas utilizadas como marcadores físicos nessa análise, bem como a posição genômica dos genes e a posição citológica obtida por *HIS*. Demais dados sobre as sondas estão disponíveis na **Tabela II.4** do **Capítulo II** desta Tese. A ordenação e a orientação dos *scaffolds* já estão de acordo com os resultados obtidos nesta análise. Note que os *scaffolds* 4577 no braço IIL, e 4558 no braço IIR tiveram duas sondas mapeadas.

Sendo assim, tinha-se perante os resultados um quadro bastante contraditório do que se vinha tradicionalmente utilizando em diferentes análises, com relação ao conteúdo genômico do cromossomo II de *D. willistoni*.

De acordo com a designação dos Elementos de Muller (Muller, 1940), o cromossomo II de *D. willistoni* tem seus braços cromossômicos IIL e IIR correspondentes aos Elementos C e B (**Figura III.3.1**).

A designação do conteúdo genômico do cromossomo II de *D. willistoni*, teve seu começo nos estudos de Ferry *et al.* (1923), Sturtevant & Novitsky (1941) e Spassky & Dozhansky (1950). Estas análises foram feitas por meio da construção de mapas de ligação, estabelecidos por cruzamentos de linhagens com marcadores fenotípicos (mutações fenotípicas). Os primeiros dois estudos relatam as dificuldades de associar as mutações ocorridas em *D. willistoni* com as mutações de *D. melanogaster*, entretanto, inferem que o cromossomo II de *D. willistoni* é composto pelos Elementos de Muller B e C, como na outra espécie (Ferry *et al.*, 1923; Sturtevant & Novitsky, 1941). Um importante ponto a se considerar é que, o alto grau de polimorfismo para inversões paracêntricas ocorrente em *D. willistoni* prejudica o estabelecimento de um mapa de ligação padrão para a espécie, e neste caso o mapa de ligação corresponderia como padrão para as linhagens utilizadas (Ehrman & Powell, 1982).

O estudo de Spassky & Dozhansky (1950) estabeleceu um mapa de ligação para os cromossomos de *D. willistoni*, por meio de cruzamentos de mutantes obtidos espontaneamente da natureza (coletados em Belém, Pará), ou induzidos por radiação; com a linhagem padrão estabelecida para a espécie, também de Belém, Pará, Brasil (Dozhansky, 1950).

Tabela III.3.1: Redefinição dos *scaffolds* do cromossomo II de *Drosophila willistoni*. Posição genômica dos genes escolhidos para o estabelecimento de sondas, e posições citológicas obtidas no presente estudo. Posição do gene e número do *scaffold* de acordo com o banco de dados *Flybase* (St. Pierre *et al.*, 2014), sendo que este último corresponde aos quatro últimos números, os quais todos começam com scf2_110000000. Orientação do *scaffold* de acordo com Schaeffer *et al.* (2008). Entre parênteses, número do *scaffold* correspondente à ordenação de acordo com a **Figura III.3.6**. Genes listados do centrômero ao telômero.

Braço cromossômico IIL						Braço cromossômico IIR					
Sonda	<i>D. willistoni</i> gene	Número do <i>scaffold</i>	Posição genômica do gene	Orientação do <i>scaffold</i>	Posição citológica	Sonda	<i>D. willistoni</i> Gene	Número do <i>scaffold</i>	Posição genômica do gene	Orientação do <i>scaffold</i>	Posição citológica
9	<i>DwilGK18743</i>	4884(1)	878.652..883.516	+	37B	bw	<i>DwilGK23049</i>	4954(8)	2.790.903..2.793.354	+	56C
10	<i>DwilGK24519</i>	4585(2)	282.915..284.232	+	38A	7	<i>DwilGK22144</i>	4558(7)	73.772..80.277	+	58C
11	<i>DwilGK18260</i>	4577(3)	1.906.481..1.918.823	-	45A	px	<i>DwilGK22138</i>	4558(7)	211.240..284.571	+	59B
Adh	<i>DwilAdh</i>	4577(3)	2.980.068..2.982.042	-	44A	6	<i>Dwilvlc</i>	4382(6)	1.240.131..1.244.696	-	62A
12	<i>DwilGK23840</i>	4516(4)	60.623..62.773	+	46A	20	<i>DwilGK17912</i>	4822(5)	150.635..154.598	+	64B
13	<i>DwilGK15054</i>	4521(5)	716.961..719.512	+	47A	32	<i>DwilGK20645</i>	4510(4)	3.566.941..3.573.564	-	68C
14	<i>DwilGK12721</i>	4752(6)	52.463..54.862	+	55A	3	<i>DwilGK15808</i>	4514(3)	205.020..207.611	-	71A
15	<i>DwilGK18432</i>	4945(7)	417.864..419.442	+	55A	2	<i>DwilAdam</i>	4513(2)	22.848..24.439	+	72A
16	<i>DwilGK21099</i>	4851(8)	226.612..230.981	-	55C	1	<i>DwilGK19495</i>	4512(1)	1.798.077..1.801.918	+	77A

Os autores ressaltam que esses mutantes oriundos da natureza se caracterizavam pela baixa prevalência de inversões cromossômicas. Os resultados obtidos nessa análise para o cromossomo II inferiam que o gene *bw* localizava-se em uma região genômica com baixa recombinação, podendo esta ser centrômero ou telômero; e que o gene *px* estava na posição 28 cM, juntamente com o gene *abb* (*abbreviated*), ou seja, posterior ao gene *bw*. O mesmo estudo ainda infere a posição genômica relativa de outros quatro genes marcadores para o cromossomo II (**Figura III.3.5**). Assim como os estudos anteriores, essa análise apontou que o conteúdo genômico do cromossomo II de *D. willistoni* era correspondente ao cromossomo II de *D. melanogaster* (Elementos de Muller B e C); porém, sem estabelecer uma associação entre o braço cromossômico e o Elemento de Muller correspondente.

Essa associação foi proposta primeiramente por Lakovaara & Saura (1972). Neste estudo foi estabelecido um mapa de ligação para os cromossomos de *D. willistoni*, a partir de mutantes oriundos das Ilhas Leewart, os quais foram apontados por serem livres de inversões (Ayala *et al.*, 1971). Da análise de quatro *locos* enzimáticos do cromossomo II (*Est-5*, *Gpdh*, *Mdh-2*, e *Adh*), os autores inferiram que esses se encontravam no braço cromossômico IIR de *D. willistoni*, e que este correspondia ao braço 2L de *D. melanogaster*, Elemento de Muller B.

É cabível aqui ressaltar que esses três estudos distintos se utilizaram de diferentes linhagens de *D. willistoni* para a análise, bem como essas linhagens não correspondem à Gd-H4-1 de Guadalupe, a qual foi sequenciada.

O gene *Adh* salientou-se como um marcador físico essencial na montagem do genoma do cromossomo II de *D. willistoni*. Este gene nas décadas de 1980/90 foi extensivamente estudado em diversas espécies do gênero *Drosophila*. Ele caracteriza-se por ser diferencialmente regulado nas distintas fases do desenvolvimento de diversos organismos, participando da metabolização de álcoois oriundos de alguns processos celulares (Goldberg, 1980). Em *D. willistoni* ele foi molecularmente caracterizado por Anderson *et al.* (1993). Nessa espécie, este gene apresenta duas peculiaridades com relação ao de outras espécies do gênero, que é a ausência do terceiro íntron (o que se

mostrou também para outras espécies do grupo *willistoni*), e diferente padrão do uso de códons.

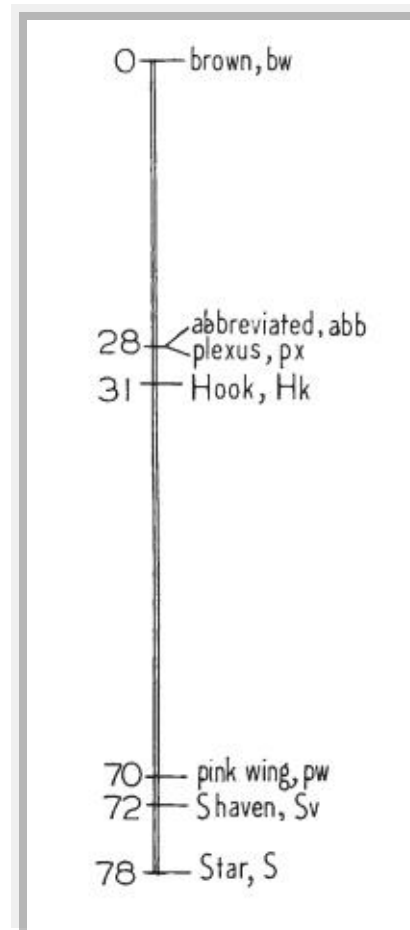


Figura III.3.5: Mapa de ligação do cromossomo II de *Drosophila willistoni* inferido por Spassky & Dozhansky (1950). Note que o gene *bw* está posicionado em uma região de recombinação nula. Retirado do referido trabalho.

Posteriormente, Rohde *et al.* (1995) o hibridaram *in situ*, utilizando como sonda o clone SAC-PAT (Moses, 1986), em sete espécies do grupo *willistoni*. Esta análise inferiu que o gene *Adh* mapeava fisicamente no braço IIR das espécies deste grupo, incluindo a *D. willistoni*, sendo esta inferência foi utilizada para a associação dos *scaffolds* aos cromossomos politêncos neste braço cromossômico por Schaeffer *et al.* (2008). Entretanto, o estudo salienta

que a determinação da localização física principal desse gene foi estabelecida por quantificação dos sítios positivos de hibridação, uma vez que a mesma apresentava múltiplos sinais. Sendo assim, em *D. willistoni* a inferência foi baseada na análise de 28 braços IIR, e destes, 18% apresentaram consistente sinal na seção 67 deste braço cromossômico. Seguindo o pressuposto de Lakovaara & Saura (1972), estes resultados foram interpretados como mais uma evidência de que os braços cromossômicos IIL e IIR eram equivalentes aos Elementos de Muller C e B, respectivamente (Vera L. S. Valente, comunicação pessoal).

A sonda estabelecida e hibridada no presente estudo apresentou somente um sítio positivo de hibridação para o gene *Adh*, mapeando na seção 44A (**Figura III.3.3**). Cabe salientar aqui que todas as demais sondas utilizadas nesse estudo se mostraram com apenas um sítio cromossômico positivamente hibridado. Somando a isso o estabelecimento de uma sonda específica, a inferência apresentada aqui para este marcador físico é bastante precisa.

Outro ponto interessante da presente análise refere-se ao gene *bw*. Como mencionado acima, esse gene foi posicionado em uma região de baixa recombinação por Spassky & Dozhansky (1950). Para a associação dos *scaffolds* com relação aos cromossomos politênicos, ele foi interpretado como posicionado na região telomérica do braço IIL, seguido pelo gene *px* (Schaeffer *et al.*, 2008). Nossos resultados mostram que este gene encontra-se na região centromérica do braço cromossômico IIR, de fato uma região de baixa recombinação meiótica; seguido adiante pelo gene *px* em *D. willistoni*.

Os resultados obtidos nessa análise alteram a atribuição dos Elementos de Muller do cromossomo II de *D. willistoni*, de tal modo que o braço cromossômico IIL é homólogo ao braço 2L de *D. melanogaster* (Elemento de Muller B) e o braço cromossômico IIR é homólogo ao braço 2R de *D. melanogaster* (Elemento de Muller C). Esse resultado é contrário ao que tem sido tradicionalmente aceito para esta espécie. A escassez de marcadores genéticos, e a interpretação ambígua dos estudos clássicos baseados em cruzamentos desta espécie altamente polimórfica têm certamente sido determinantes para a manutenção deste erro por tanto tempo.

A **Figura III.3.6** mostra a redefinição dos *scaffolds* no cromossomo II de *D. willistoni* resultante da presente análise, de acordo com os marcadores físicos descritos na **Tabela III.3.1**.

Com relação à orientação dos *scaffolds* estabelecidos de acordo com os marcadores genéticos disponíveis (Schaeffer *et al.*, 2008), a presente análise mostra que a orientação da junção dos *scaffolds* do braço cromossômico IIR está com a ordem invertida. Esta inferência é reforçada pela determinação física dos dois genes nos *scaffolds* 4577 (braço IIL) e 4558 (braço IIR). A ordem dos genes no *scaffold* 4577 mostrou-se conforme o esperado, de acordo com suas posições genômicas. Por sua vez, no *scaffold* 4558, o qual é inferido como estando em ordem negativa (*minus*), o esperado seria a sonda do gene *Dwil\GK22138 (px)* se posicionar em um sítio cromossômico anterior à outra sonda gênica *Dwil\GK22144* utilizada. Entretanto, conforme visto na **Tabela III.1** e na **Figura III.3.6**, a determinação física dessas sondas por *HIS* resultou no gene *Dwil\GK22144* posicionando-se em um sítio cromossômico anterior ao do gene *Dwil\GK22138 (px)*. Outro fator que reforça esta inferência é que estes *scaffolds* tiveram sua orientação inferida anteriormente, considerando que esses pertenciam ao braço cromossômico IIL, o qual teria os genes *bw* e *px* posicionados na região telomérica (Schaeffer *et al.*, 2008); quando em realidade eles estão posicionados na região centromérica do braço IIR (ou seja, em ordem oposta).

No presente estudo também foram estabelecidas três sondas para o cromossomo X, e uma sonda para o cromossomo III, a fim de averiguar brevemente a associação dos *scaffolds* nesses cromossomos (**Tabela III.3.2**).

Estas sondas foram igualmente sequenciadas diretamente dos seus produtos da *PCR* e confirmadas por BLASTN nos bancos de dados *Flybase* (St. Pierre *et al.*, 2014) e *NCBI* (Geer *et al.*, 2010).

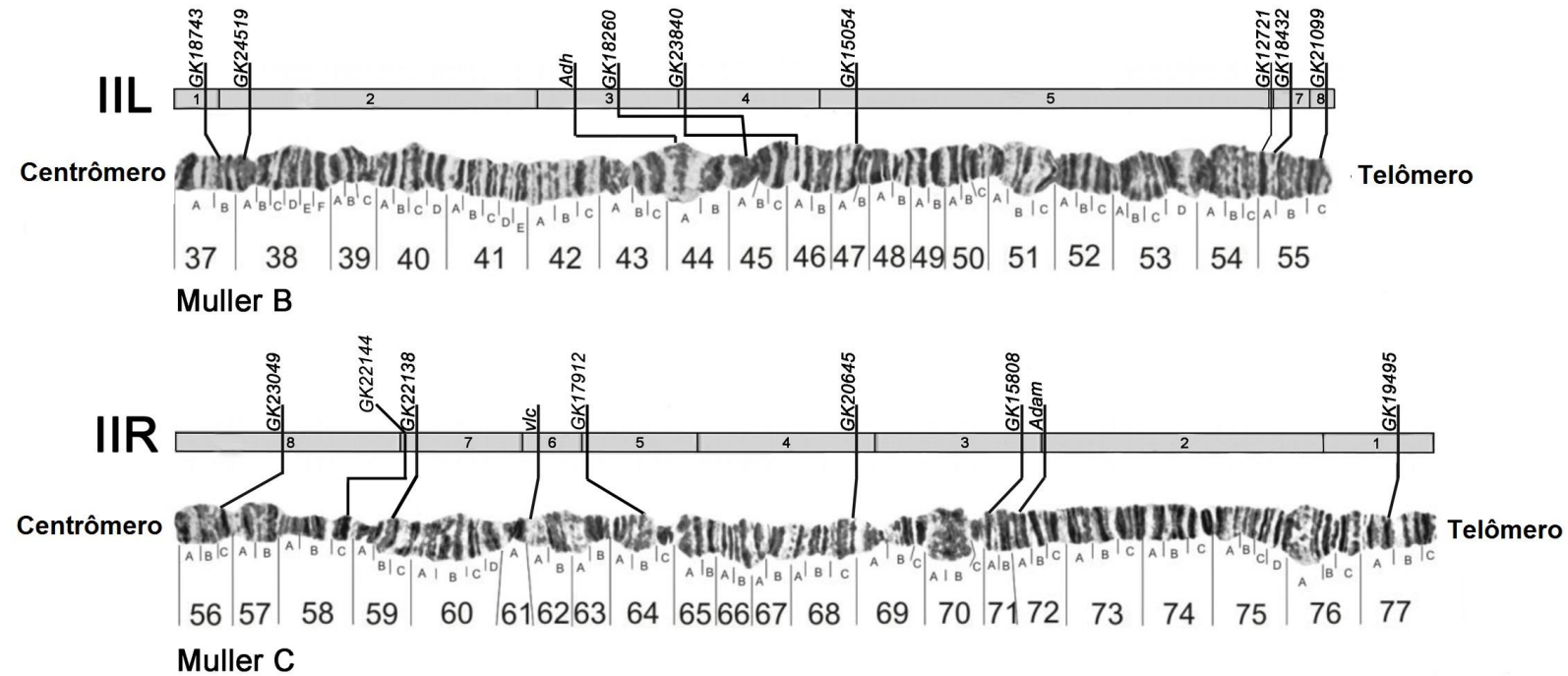


Figura III.3.6: Redefinição dos *scaffolds* e marcadores gênicos do cromossomo II de *Drosophila willistoni* (Schaeffer *et al.*, 2008), de acordo com este estudo (**Tabela II.1**). Números 1-8 no braço cromossômico IIL correspondem aos *scaffolds* 4884, 4585, 4577, 4516, 4521, 4752, 4945 e 4851 na **Tabela III.3.1**. Números 1-8 no braço cromossômico IIR correspondem aos *scaffolds* 4512, 4513, 4514, 4510, 4822, 4382, 4558 e 4954 na **Tabela III.3.1**.

Tabela III.3.2: Posição genômica dos genes utilizados para o estabelecimento de sondas, e posições citológicas obtidas na análise da designação dos *scaffolds* dos cromossomos X e III de *Drosophila willistoni*, no presente estudo. Posição do gene e número do *scaffold* de acordo com o banco de dados *Flybase* (St. Pierre *et al.*, 2014), sendo que este último corresponde aos quatro últimos números, os quais todos começam com scf2_110000000.

Sonda	<i>D. willistoni</i> gene	Posição genômica do gene	Número do <i>scaffold</i>	Posição citológica/ Cromossomo
68	<i>Dwi\GK16707</i>	432.088..435.746	4963	1C/XL
5	<i>Dwi\GK17758</i>	3.033.141..3.035.483	4822	27C/XR
30	<i>Dwi\GK16749</i>	7.841.949..7.843.999	4511	34B/XR
31	<i>Dwi\GK22422</i>	3.260.674..3.262.239	4921	99D/III

A sonda 68 estabelecida para o braço cromossômico XL, de acordo com a montagem do genoma posiciona-se no *scaffold* associado à região telomérica desse braço. Entretanto, nosso mapeamento físico resultou na sua posição citológica na seção 1C do braço cromossômico XL, sendo assim, na região mais centromérica desse braço. Esse resultado aponta um desacordo com a posição estabelecida inicialmente para esse gene (Schaeffer *et al.*, 2008). Isto pode ser devido a um caso de quimerismo na montagem do genoma, ou uma indicação de que a orientação da montagem dos *scaffolds* nesse braço está invertida. Nossos dados não são suficientes para embasar nenhuma hipótese acerca da montagem do genoma do braço XL de *D. willistoni*. Entretanto, note que em *D. melanogaster* o ortólogo deste gene (*Dme\unc*) mapeia citologicamente na seção 19F1 do cromossomo X desta espécie, uma região próxima ao centrômero.

O estudo de Schaeffer *et al.* (2008) infere que o *scaffold* 4822 de *D. willistoni* apresenta uma terça parte pertencente ao braço cromossômico XR, e as duas partes restantes ancoradas ao braço cromossômico IIR (correção feita

neste estudo). Em vista disso, a sonda 5 foi designada para o gene *Dwil\GK17758*, pertencente à região desse *scaffold* ancorada no braço cromossômico XR. O resultado obtido mapeou fisicamente esse gene na seção 27C desse braço cromossômico, mostrando uma concordância com a observação feita por Schaeffer *et al.* (2008) acerca da montagem do *scaffold* 4822. Além disso, outra sonda foi designada para o braço XR (sonda 30), associada ao *scaffold* mais telomérico desse braço. O mapeamento físico apontou a posição citológica desta sonda na seção 27C do braço XR, estando de acordo com a posição telomérica desse gene. O braço cromossômico XR teve seu genoma montado com base na conservação da sintonia com as outras espécies, bem como de acordo com seis marcadores genéticos obtidos pelo estabelecimento de mapas de ligação, e mais cinco marcadores físicos mapeados por *HIS* (genes *E74*, *E75*, *Sod*, *Hsp83*, e *Hsp27*). Visivelmente, um número consideravelmente maior de marcadores quando comparados aos braços cromossômicos XL e II. Os resultados inferidos por *HIS* no presente estudo mostraram estar em coerência com a associação dos *scaffolds* feita Schaeffer *et al.* (2008), a esse braço cromossômico.

O cromossomo III de *D. willistoni* é constituído pelos Elementos de Muller E + F fusionados. Esta hipótese foi levantada já por Sturtevant e Novitski (1941), devido à construção de um mapa de ligação de *D. willistoni*, a fim de estabelecer as homologias com os elementos de Muller. Por *HIS*, Papacit e Juan (1998) mapearam fisicamente em *D. willistoni* os genes *cubitus interruptus (ci)*, *ankyrin (Ank)* e *eyeless (ey)*, que são marcadores genéticos do Elemento de Muller F de *D. melanogaster*. Com isso, eles corroboraram por meio de técnicas moleculares, pela primeira vez, a fusão desses elementos nessa espécie. Esse evento, atualmente, foi confirmado para todo subgrupo *willistoni* (Powell *et al.*, 2011; Pita *et al.*, 2014.), com um tempo de ocorrência estimado em 15 Mya. Tendo em vista a boa determinação dos marcadores físicos disponíveis para a base do cromossomo III, o presente estudo mapeou por *HIS* uma sonda determinada para o gene *Dwil\GK22422* (sonda 31) do *scaffold* 4921 (**Tabela III.2**), o mais telomérico desse cromossomo. Sua localização física se mostrou na subseção 99D, estando de acordo com a localização desse gene no referido *scaffold*. A montagem do genoma do

cromossomo III de *D. willistoni* é considerada como a melhor suportada com relação aos seus marcadores físicos, comparado aos demais cromossomos (Schaeffer *et al.*, 2008).

De maneira geral os marcadores físicos utilizados no presente estudo, para os cromossomos X e III de *D. willistoni*, sugerem um quadro de maior sucesso na montagem e associação dos *scaffolds* com relação aos seus respectivos cromossomos, sem a ocorrência da troca de conteúdo genético (pelo menos em grandes proporções), como no caso dos Elementos de Muller B e C. No entanto, a linhagem sequenciada Gd-H4-1 ainda necessita de uma melhor resolução para esses cromossomos, estabelecendo marcadores físicos por *HIS*, como apresentado nesta análise para o cromossomo II.

No total, foram mapeados fisicamente 22 marcadores gênicos nos cromossomos da linhagem sequenciada de *D. willistoni*. Considerando que os demais marcadores desta espécie foram estabelecidos anteriormente em distintas linhagens de diferentes regiões geográficas, a inferência da posição física desses marcadores resulta em posições citológicas diferentes na linhagem sequenciada. Sendo assim, é extremamente importante para a montagem de um genoma o correto alinhamento do seu conteúdo gênico e citológico. Para *D. willistoni*, este padrão é apresentado na presente análise pela primeira vez.

Este estudo contribui para uma caracterização mais precisa do genoma de *D. willistoni*, proporcionando uma melhor base para a continuação de estudos genéticos e genômicos a ele associados. Como *D. willistoni* é um modelo instigante para estudos cromossômicos e evolutivos, em particular, os resultados aqui descritos irão ajudar na análise da evolução cromossômica interespecífica e evolução do genômica do grupo (como em Ruiz *et al.*, 1997; Ranz *et al.*, 2001; Papaceit *et al.*, 2006; com outras espécies) e também na caracterização dos pontos de quebra de inversões cromossômicas (como em Delprat *et al.*, 2009). De fato, esta análise contempla um passo anterior à delimitação dos pontos de quebra da inversão IIL-H de *D. willistoni*, conforme será apresentado no **Tópico III.4** desta Tese.

III.4 – Estudo dos pontos de quebra de inversões de *Drosophila willistoni*.

IV. Mapeamento físico de alta resolução e delimitação gênica dos pontos de quebra da inversão III-H.

Dentre as espécies do gênero *Drosophila*, a *D. melanogaster* e a *D. buzzatii* se destacam pelo pioneirismo nos estudos dos pontos de quebras de inversões. Por meio de distintas análises nestas espécies, atualmente se tem base para a ocorrência de dois principais mecanismos envolvidos na origem de inversões: o mecanismo de quebras escalonadas (Ranz *et al.*, 2007; Corbett-Detig *et al.*, 2012), o qual tem sido apontado como o responsável pela origem da maioria das inversões ocorrentes no subgrupo *melanogaster*, e o mecanismo de recombinação ectópias entre cópias invertidas de determinados TEs (Delprat *et al.*, 2009; Fonseca *et al.*, 2012), com estes estudos agregando muito ao conhecimento do papel modulador de genomas subsidiado por TEs (Cáceres *et al.*, 1999) (Veja **Capítulo I, Tópico I.4**, pág. 40).

Apesar de os primeiros estudos terem sido anteriores à era dos genomas (Wesley & Eanes, 1994; Cicera *et al.*, 1995; Andolfatto *et al.*, 1999; Cáceres *et al.*, 1999), utilizando principalmente bibliotecas genômicas. A partir do sequenciamento do genoma de *D. melanogaster* (Adams *et al.*, 2000) somado à alta resolução da caracterização do seu genoma, e os subsequentes genomas de outras linhagens desta espécie, atualmente este cenário permite a análise dos pontos de quebra dessa espécie por meio de algoritmos elaborados para este fim (Ranz *et al.*, 2007; Corbett-Detig *et al.*, 2012). Entretanto, esta não é a realidade para a maioria das espécies, muitas de elevada importância em diferentes focos de análises (*Drosophila* 12 Genomes Consortium, 2007). De qualquer forma, a delimitação dos pontos de quebra de uma inversão é uma etapa primordial para os estudos de sua gênese.

Drosophila willistoni é uma espécie que se destaca por ser uma das mais polimórficas para inversões segregantes em todo o gênero *Drosophila* (sendo apontada muitas vezes como a espécie mais polimórfica) (Kirmbas & Powell, 1992). A disponibilização de um genoma sequenciado para esta

espécie traz novas facilidades para as análises do seu vasto polimorfismo cromossômico para inversões paracêntricas, agora com foco nos fatores que originaram estas inversões. Todavia, o genoma de *D. willistoni* se caracteriza pela pobreza de resolução gênica (Schaeffer *et al.*, 2008), e este fator torna, principalmente, a delimitação de seus pontos de quebras mais laboriosa e necessitante de ajustes.

As inferências apontadas nos **Tópicos III.1, III.2 e III.3** estabeleceram as bases iniciais para uma correta delimitação dos pontos de quebra da inversão IIL-H de *D. willistoni*.

Neste tópico, a presente análise utilizou o genoma disponível da linhagem Gd-H4-1 no banco de dados *Flybase* (St. Pierre *et al.*, 2014) para o desenho de *primers* e o delineamento da caminhada cromossômica, a qual foi feita por *HIS* não fluorescente, majoritariamente nos cromossomos politênicos da linhagem SG12.00.

A linhagem SG12.00 fixou a inversão IIL-H (Goñi & Valente, 2012). Esta inversão abrange a ponta do braço IIL, entre as seções 53A e 55B (Rohde & Valente, 2012) (**Figura III.4.1**).

A fim de realizar uma análise padronizada, os pontos de quebra da inversão IIL-H foram denominados na linhagem sequenciada Gd-H4-1 de *AB* (ponto de quebra distal) e *CD* (ponto de quebra proximal), *AC* (ponto distal) e *BD* (ponto proximal) na linhagem SG12.00 que porta a inversão IIL-H fixada. Esta padronização segue as análises detalhadas dos pontos de quebra de Cáceres *et al.* (1999), Casals *et al.* (2003), Delprat *et al.* (2009), Calvete *et al.* (2012) (**Figura III.4.2**).

A montagem do genoma do braço cromossômico II de *D. willistoni* resultou no estabelecimento de oito *scaffolds*, sendo o *scaffold* 1 o mais centromérico, e o *scaffold* 8 o mais telomérico. O estudo descrito no **Tópico III.3** deste capítulo, estabeleceu os marcadores físicos iniciais para a esquematização da delimitação dos pontos de quebra distal (mais próximo ao telômero), e do ponto de quebra proximal (próximo ao centrômero) da inversão IIL-H.

Apesar de a caminhada cromossômica para a delimitação dos pontos de quebra distal e proximal da inversão IIL-H terem sido feitas em grande parte

simultaneamente, a apresentação no presente estudo será em detalhamento para cada ponto de quebra em isolado.

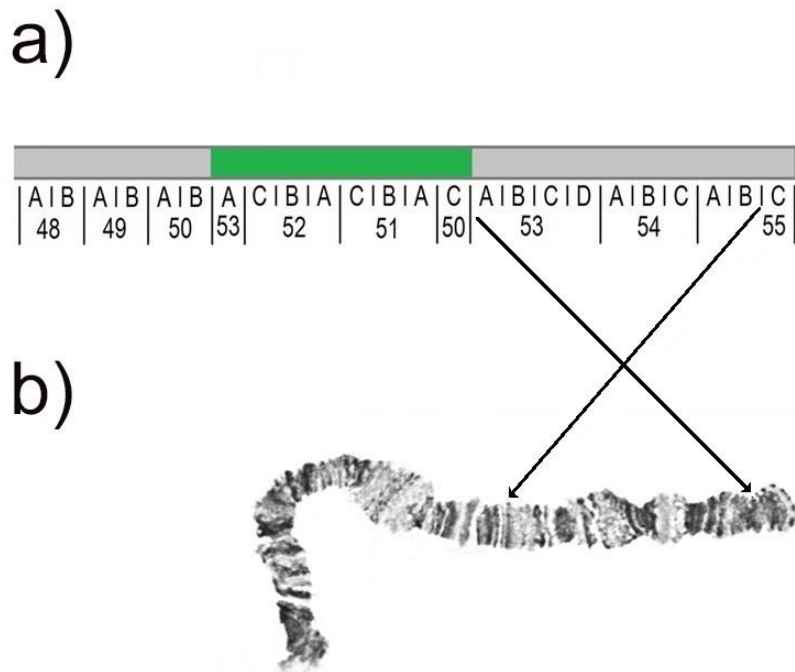


Figura III.4.1: Delimitação citológica da inversão IIL-H em *Drosophila willistoni*.

a) Representação esquemática do arranjo cromossômico da ponta do braço IIL da linhagem Gd-H4-1 (em verde inversão IIL-F destacada), com relação ao fotomapa da espécie (**Tópico III.1**). **b)** Ponta do braço IIL da linhagem SG12.00 (Goñi & Valente, 2012), com a inversão IIL-H fixada. Flechas em preto associando as seções 53A e 55B que delimitam a inversão IIL-H em ambas as imagens.

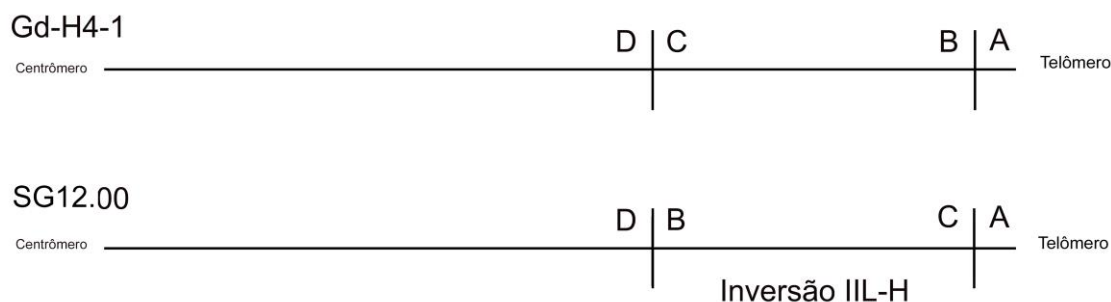


Figura III.4.2: Representação esquemática da denominação dos pontos de quebra distal e proximal da inversão IIL-H utilizada no presente estudo, nas linhagens Gd-H4-1 e SG12.00.

Delimitação do ponto de quebra distal da inversão IIL-H:

Como etapa inicial, para a caminhada cromossômica do ponto de quebra distal da inversão IIL-H, utilizou-se as sondas dos genes *DwilGK12721* (*scaffold 6*), *DwilGK18432* (*scaffold 7*) e *DwilGK21099* (*scaffold 8*) estabelecidas no estudo descrito no **Tópico III.3** deste capítulo. O mapeamento físico dessas sondas, nos cromossomos politênicos da linhagem SG12.00, apontou que o ponto de quebra distal estava ou no *scaffold 7*, ou no *scaffold 8*, descartando o *scaffold 6* (**Figura III.4.3**). O *scaffold 7* do braço IIL possui 1,3 Mb, e o *scaffold 8* possui 640 kb, sendo estes, juntamente com o *scaffold 6* (91 kb), os mais pequenos do genoma de *D. willistoni*. Além, disso, o *scaffold 8* possui orientação *minus* (-). A partir daí, foram estabelecidas mais duas sondas para o *scaffold 7* (genes *DwilGK18461* e *DwilGK18345*, o qual era o último gene deste *scaffold*), e mais quatro sondas para o *scaffold 8* (genes *DwilGK21105*, *DwilGK21048*, *DwilGK21116* e *DwilGK21118*) (**Tabela III.4.1**).

Os resultados desta caminhada cromossômica, por meio de hibridação *in situ* das sondas apresentadas na **Tabela III.4.1**, indicaram que o ponto de quebra distal da inversão IIL-H estava no intervalo compreendido entre os genes *DwilGK21048* e *DwilGK21116* (**Figura III.4.4**). A análise dessa região, no genoma da linhagem Gd-H4-1, mostrou a presença de mais três genes

entre eles. Assim, passou-se a utilizar uma abordagem para a determinação do ponto de quebra distal por meio da técnica da *PCR*, uma vez que se esperava que na região onde o ponto de quebra se estabeleceu, não ocorreria amplificação na *PCR* da linhagem SG12.00, porque esta estaria corrompida e ao contrário (**Figura II.2b, Capítulo II**). Como as regiões intergênicas entre estes genes compreendiam sequências entre 147 pb a 937 pb, *primers* foram estabelecidos para a amplificação destas regiões (**Tabela III.4.2**). Como precaução, analisou-se a região intergênica anterior ao gene *Dwil\GK21048* e também a região posterior ao gene *Dwil\GK21116*, totalizando cinco regiões intergênicas na análise (**Figura III.4.5**).

O resultado da análise das regiões intergênicas pela *PCR* tanto na linhagem Gd-H4-1, quanto na linhagem SG12.00 mostrou que o ponto de quebra distal da inversão IIL-H ocorreu no intervalo entre os genes *Dwil\GK21048* e *Dwil\GK21115* (**Figura III.4.5**).

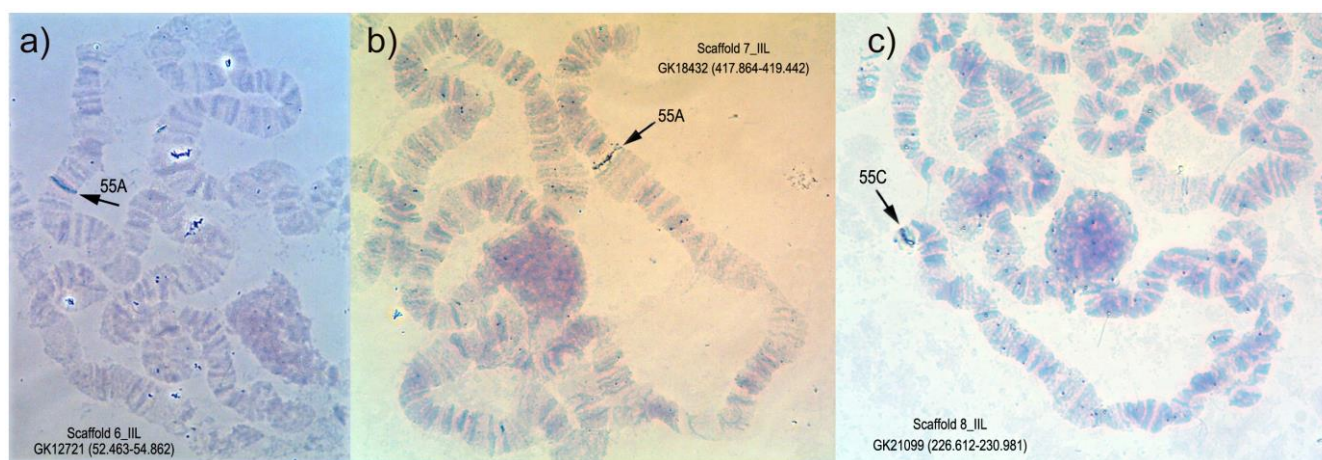


Figura III.4.3: Resultados dos mapeamentos físicos obtidos na etapa inicial da caminhada cromossômica do ponto de quebra distal da inversão IIL-H em *Drosophila willistoni*. **a)** gene *Dwil\GK12721* (*scaffold* 6), **b)** gene *Dwil\GK18432* (*scaffold* 7), **c)** *Dwil\GK21099* (*scaffold* 8). Setas indicam sítio de hibridação no braço IIL da linhagem SG12.00.

Posteriormente a hibridação *in situ* da sonda do gene *Dwil\GK21048* na linhagem Gd-H4-1, e a elaboração de uma sonda para o gene *Dwil\GK21115* (sonda 72), a qual foi hibridada nas linhagens Gd-H4-1 e SG12.00, finalizaram a delimitação do ponto de quebra distal da inversão IIL-H em *D. willistoni*. Conforme esperado, na linhagem sequenciada Gd-H4-1, os genes *Dwil\GK21048* e *Dwil\GK21115* mapearam fisicamente na seção 55B (**Figura III.4.6a**, **Figura III.4.6c**). Já na linhagem com a inversão SG12.00, o gene *Dwil\GK21048* mapeou fisicamente na seção 55B (**Figura III.4.6b**), mostrando estar fora da inversão uma vez que permanecia na ponta do braço IIL. O gene *Dwil\GK21115* também mapeia na seção 55B, entretanto no ponto de quebra proximal; indicando que a quebra distal, além de ter ocorrido entre esses dois genes, também particionou a seção 55B (**Figura III.4.6d**).

Tabela III.4.1: Marcadores gênicos utilizados para a delimitação do ponto de quebra distal da inversão IIL-H, por caminhada cromossômica, na linhagem SG12.00 de *Drosophila willistoni*. Posição do gene e número do *scaffold* de acordo com o banco de dados *Flybase* (St. Pierre *et al.*, 2014), sendo que este último corresponde aos quatro últimos números, os quais todos começam com scf2_110000000. Entre parênteses, número do *scaffold* correspondente à ordenação de acordo com a **Figura III.3.6 (Tópico III.3)**. Sonda 72 foi estabelecida *a posteriori* dos resultados obtidos pela técnica da *PCR* (**Tabela III.4.2**)

Sonda	<i>D. willistoni</i> Gene	Número do <i>scaffold</i>	Posição do gene no <i>scaffold</i>	Posição citológica
14	<i>Dwil</i> \GK12721	4752(6)	52.463..54.862	55A
15	<i>Dwil</i> \GK18432	4945(7)	417.864..419.442	55A
24	<i>Dwil</i> \GK18461	4945(7)	777.780..779.503	55B
25	<i>Dwil</i> \GK18345	4945(7)	918.142..1.035.355	55B
16	<i>Dwil</i> \GK21099	4851(8)	226.612..230.981	55C
21	<i>Dwil</i> \GK21105	4851(8)	449.106..451.173	53C
22	<i>Dwil</i> \GK21048	4851(8)	513.672..515.215	55B
72	<i>Dwil</i> \GK21115	4851(8)	516.051..520.321	55B
34	<i>Dwil</i> \GK21116	4851(8)	523.243..527.719	55B
35	<i>Dwil</i> \GK21118	4851(8)	528.655..533.613	55B

Tabela III.4.2: Intervalos gênicos (denominados de A-E) utilizados para a delimitação do ponto de quebra distal da inversão IIL-H, pela técnica da *PCR*, em ambas as linhagens de *Drosophila willistoni*, de acordo com a **Figura II.2 (Capítulo II)**. Posição no *scaffold* e número do *scaffold* de acordo com o banco de dados *Flybase* (St. Pierre *et al.*, 2014), sendo que este último corresponde aos quatro últimos números, os quais todos começam com scf2_110000000. Entre parênteses, número do *scaffold* correspondente à ordenação de acordo com a **Figura III.3.6 (Tópico III.3)**. Inter: Intergênica. +: Amplificação positiva pela técnica da *PCR*.

Região Intergênica	Número do <i>scaffold</i>	Posição no <i>scaffold</i>	PCR Gd-H4-1/SG12.00
A = Inter			
<i>Dwil</i> \GK21049- <i>Dwil</i> \GK21048	4851(8)	513.304.. 514.299	+/+
B = Inter			
<i>Dwil</i> \GK21048- <i>Dwil</i> \GK21115	4851(8)	514.748.. 516.180	+/-
C = Inter			
<i>Dwil</i> \GK21115- <i>Dwil</i> \GK21047	4851(8)	519.899.. 520.995	+/+
D = Inter			
<i>Dwil</i> \GK21047- <i>Dwil</i> \GK21116	4851(8)	522.013.. 523.934	+/+
E = Inter			
<i>Dwil</i> \GK21047- <i>Dwil</i> \GK21118	4851(8)	527.319.. 528.800	+/+

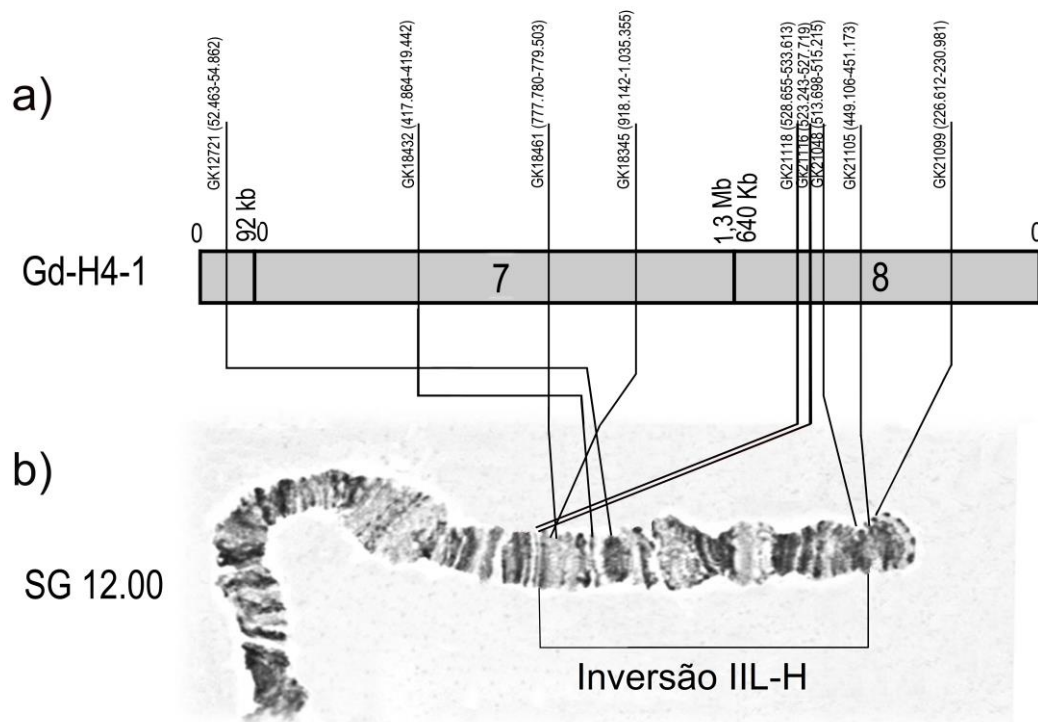


Figura III.4.4: Esquematização dos resultados obtidos por caminhada cromossômica, para a delimitação do ponto de quebra distal da inversão IIL-H, na linhagem SG12.00 de *Drosophila willistoni*. **a)** Representação esquemática dos *scaffolds* 6, 7 e 8 no genoma da linhagem Gd-H4-1, com as respectivas sondas utilizadas. **b)** Ponta do braço IIL da linhagem SG12.00 (Goñi & Valente, 2012), com a inversão IIL-H fixada delimitada abaixo. Linhas em preto conectando **a)** e **b)** fazem a correspondência da posição genômica das sondas dos genes *Dwil*\GK12721, *Dwil*\GK18432, *Dwil*\GK18461, *Dwil*\GK18345, *Dwil*\GK21099, *Dwil*\GK21105, *Dwil*\GK21088, *Dwil*\GK21116 e *Dwil*\GK21118 com suas posições citológicas.

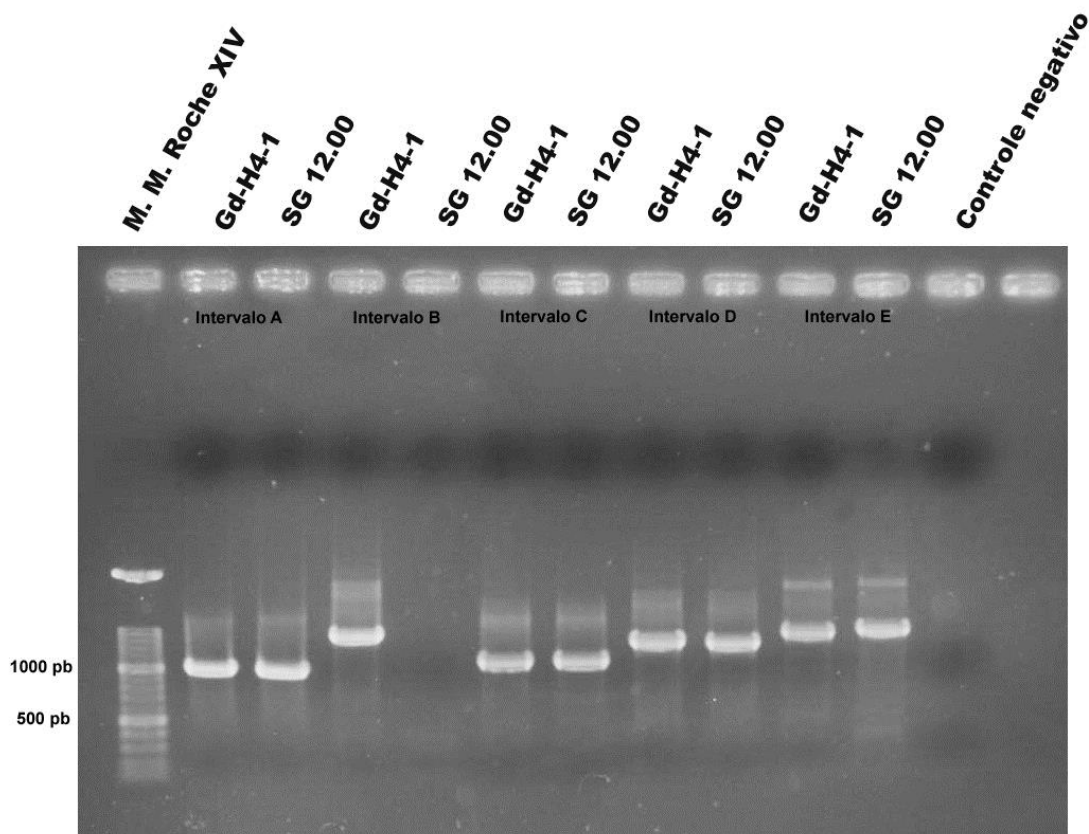


Figura III.4.5: Visualização em gel de agarose 0,7% das cinco regiões intergênicas obtidas pela técnica da *PCR*, para a delimitação do ponto de quebra distal da inversão IIL-H de *Drosophila willistoni*. As *PCRs* foram feitas na linhagem sequenciada Gd-H4-1 e na linhagem SG12.00, a qual possui a inversão fixada. Intervalo A: entre genes *Dwil\GK21049* e *Dwil\GK21048*. Intervalo B: entre genes *Dwil\GK21048* e *Dwil\GK21115*. Intervalo C: entre genes *Dwil\GK21115* e *Dwil\GK21047*. Intervalo D: entre genes *Dwil\GK21047* e *Dwil\GK21116*. Intervalo E: entre genes *Dwil\GK21116* e *Dwil\GK21118*. A análise mostra que o ponto de quebra distal da inversão IIL-H encontra-se no intervalo B.

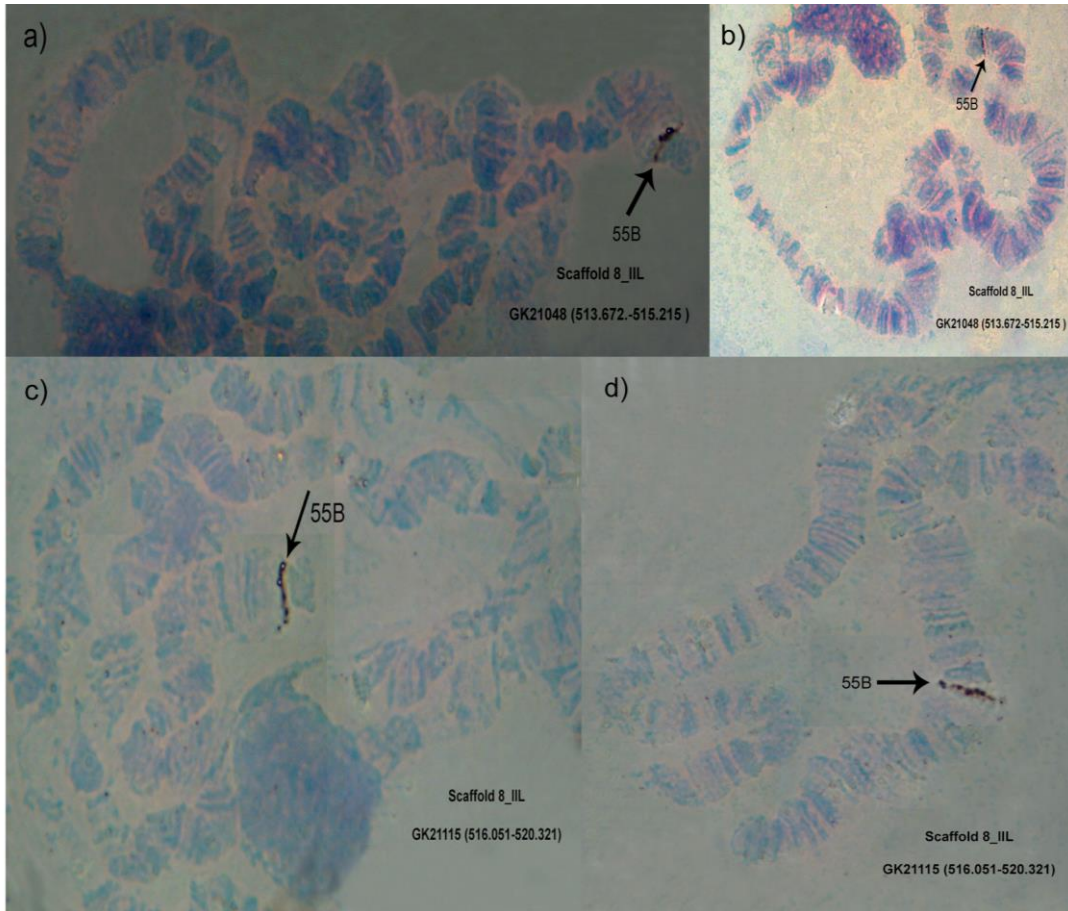


Figura III.4.6: Mapeamento físico dos genes que delimitam o ponto de quebra distal da inversão IIL-H em *Drosophila willistoni*. **a)** e **c)** genes *Dwil*\GK21048 e *Dwil*\GK21115 na linhagem Gd-H4-1 respectivamente. **b)** e **d)** genes *Dwil*\GK21048 e *Dwil*\GK21115 na linhagem SG12.00. Setas indicam sítio de hibridação no braço cromossômico IIL na seção 55B, mostrando que esta seção foi particionada pela inversão.

Os genes *Dwil*\GK21048 e *Dwil*\GK21115 foram sequenciados diretamente dos seus produtos da *PCR* em ambas as linhagens Gd-H4-1 e SG12.00, pela empresa de sequenciamento MacroGen Inc. (Seoul, Korea). As seqüências consensos obtidas pela análise dos cromatogramas foram submetidas à BLASTN nos bancos de dados Flybase (St. Pierre *et al.*, 2014) e *NCBI* (Geer *et al.*, 2010), comprovando em 100% dos casos que as sondas tratavam-se das regiões gênicas escolhidas para as sondas.

A região intergênica do ponto de quebra distal da inversão IIL-H, delimitada pelos genes *Dwil\GK21048* e *Dwil\GK21115*, possui no genoma 836 pb. Estes dois genes são ortólogos, respectivamente aos genes *CG43345* e *Acon* em *D. melanogaster*, os quais mantêm a paridade nesta espécie, entretanto em orientação oposta com relação à *D. willistoni*. Considerando a anotação completa dos genes no genoma de *D. melanogaster*, a região intergênica entre *CG43345* e *Acon* possui 11 pb.

O genoma de *D. willistoni* carece de informações acerca das regiões promotoras de seus genes, e considerando a anotação dos genes ortólogos de *Dwil\GK21048* e *Dwil\GK21115* em *D. melanogaster*, pode ser que a região intergênica entre estes genes seja de menor tamanho; e seguindo esta suposição, uma quebra nesta região poderia vir a romper a região promotora de um destes genes em *D. willistoni*.

A análise de buscas por similaridades desta região intergênica (836 pb) com *TEs*, por meio do banco de dados *Repbase* (Jurka *et al.*, 1996), apontou a ocorrência de uma similaridade de 93,33% com o *TE Chapaev-2_CE* (29 pb) de *Caenorhabditis elegans* (Kapitonov & Jurka, 2007), e outra de 85,42% com o *TE GYPSY14-I_AG* (48 pb) de *Anopheles gambiae* (Tubio *et al.*, 2003). Como esta busca se baseou em todos os *TEs* anotados neste banco de dados, a região se mostrou pobre para a presença deste tipo de sequência.

Delimitação do ponto de quebra proximal:

A caminhada cromossômica para a delimitação do ponto de quebra proximal da inversão IIL-H mostrou-se muito mais laboriosa e complexa. Conforme visto, o ponto de quebra distal da inversão IIL-H encontra-se no *scaffold* 8 do braço IIL; por sua vez, o ponto de quebra proximal encontra-se no *scaffold* 5 do mesmo braço cromossômico. O *scaffold* 5 é o de maior tamanho do cromossomo II de *D. willistoni*, com 12,5 Mb. O marcador físico que já se tinha estabelecido, na abordagem descrita no **Tópico III.3**, encontra-se no início deste *scaffold*, na posição genômica 716.968 - 719.505. O mapeamento físico desse marcador foi na seção 47A do braço IIL, bastante afastada do ponto de quebra proximal da inversão IIL-H na seção 53B. Sendo assim, para a determinação deste ponto de quebra, foram necessárias quatro etapas distintas de planejamento e análises.

A primeira etapa consistiu na determinação de sondas a partir da segunda metade do *scaffold* 5, aproximadamente a cada 2,5 Mb. A hibridação *in situ* destas sondas nos cromossomos politênicos da linhagem SG12.00 mostrou que as sondas mais extremas dos genes *DwilGK14795* (região entre 6.444.384 - 6.445.729) e *DwilGK14536* (região entre 12.267.294 - 12.269.546) localizavam-se nas seções 51B e 54C, respectivamente. Por sua vez, a sonda estabelecida para o gene *DwilGK14671* (região entre 8.996.332 - 8.998.389) localizava-se na seção 53C (**Figura III.4.7**).

Esta análise mostrou que o ponto de quebra proximal da inversão IIL-H estava entre os 7 Mb e 8 Mb do *scaffold* 5, uma vez que a sonda do gene *DwilGK14795* se posicionava fora da inversão, a sonda do gene *DwilGK14536* se posicionava dentro da inversão e a sonda *DwilGK14671* se posicionava dentro da inversão, porém mais próxima ao ponto de quebra distal.

Dando seguimento à análise, na segunda etapa foram estabelecidas uma sonda para o término da região entre os 7 Mb e cinco sondas para a região entre os 8 Mb do *scaffold* 5 (**Tabela III.4.3**).

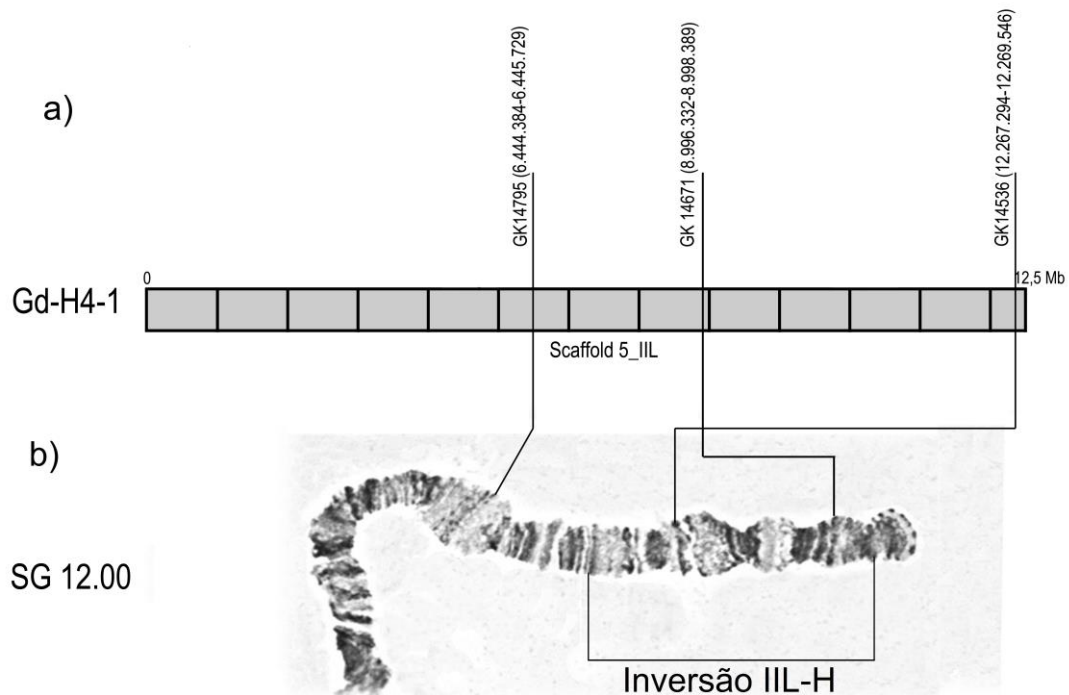


Figura III.4.7: Esquematização dos resultados obtidos na primeira etapa para a delimitação do ponto de quebra proximal da inversão IIL-H de *Drosophila willistoni*. **a)** Representação esquemática do *scaffold* 5 no genoma disponível da linhagem Gd-H4-1, com divisões representando 1 Mb cada, e as respectivas sondas utilizadas. **b)** Ponta do braço IIL da linhagem SG12.00 (Goñi & Valente, 2012), com a inversão IIL-H fixada delimitada abaixo. Linhas em preto conectando **a)** e **b)** fazem a correspondência da posição genômica das sondas dos genes *Dwil*\GK14795, *Dwil*\GK14671 e *Dwil*\GK14536 com suas posições citológicas.

Tabela III.4.3: Marcadores gênicos utilizados na primeira e segunda etapas da caminhada cromossômica, para o estabelecimento do ponto de quebra proximal da inversão IIL-H na linhagem SG12.00 de *Drosophila willistoni*. Posição do gene e número do *scaffold* de acordo com o banco de dados *Flybase* (St. Pierre *et al.*, 2014), sendo que este último corresponde aos quatro últimos números, os quais todos começam com scf2_110000000. Entre parênteses, número do *scaffold* correspondente à ordenação de acordo com a **Figura III.3.6 (Tópico III.3)**.

Sonda	<i>D. willistoni</i> gene	Número do <i>scaffold</i>	Posição do gene no <i>scaffold</i>	Posição citológica
17	<i>Dwil</i> \GK14795	4521(5)	6.444.384..6.445.729	51B
36	<i>Dwil</i> \GK14718	4521(5)	7.875.579..7.877.694	51A
37	<i>Dwil</i> \GK15415	4521(5)	8.178.091..8.179.762	50C
48	<i>Dwil</i> \GK19065	4521(5)	8.463.003..8.464.256	53B
39	<i>Dwil</i> \GK14687	4521(5)	8.686.947..8.688.509	53B
40	<i>Dwil</i> \GK19024	4521(5)	8.942.706..8.943.796	53C
28	<i>Dwil</i> \GK14671	4521(5)	8.996.332..8.998.389	53C
19	<i>Dwil</i> \GK14536	4521(5)	12.267.294..12.269.546	54C

De acordo com os resultados obtidos na segunda etapa da caminhada cromossômica, ao mesmo tempo em que ficou evidente que o ponto de quebra proximal da inversão IIL-H encontra-se no intervalo entre os genes *Dwil*\GK15415 e *Dwil*\GK19065, também se mostrou evidente que o intervalo entre os genes *Dwil*\GK14795 e *Dwil*\GK15415 estavam posicionados citologicamente, na linhagem SG12.00, ao contrário da ordem disponível no genoma da linhagem Gd-H4-1 (**Figura III.4.8**).

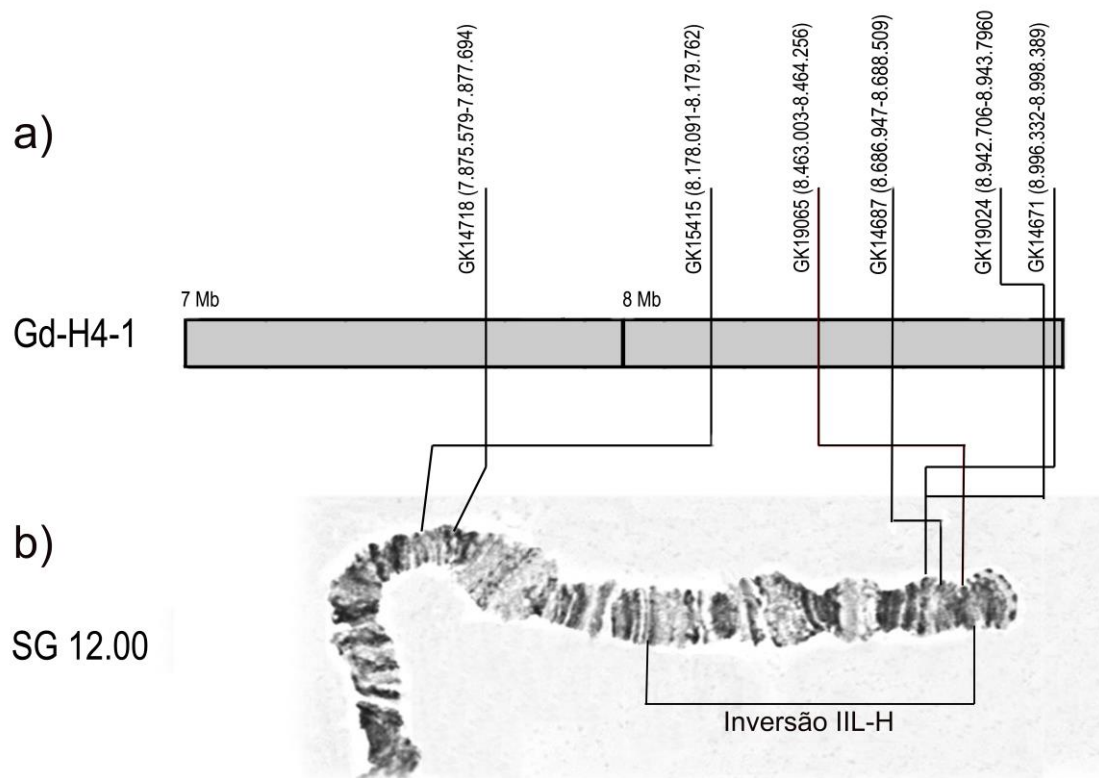


Figura III.4.8: Esquematização dos resultados obtidos na segunda etapa para a delimitação do ponto de quebra proximal da inversão IIL-H de *Drosophila willistoni*. **a)** Representação esquemática do *scaffold* 5 no genoma disponível da linhagem Gd-H4-1, com suas respectivas sondas utilizadas, dando-se um “zoom” na região entre os 7 Mb e 8 Mb deste *scaffold*. **b)** Ponta do braço IIL da linhagem SG12.00 (Goñi & Valente, 2012), com a inversão IIL-H fixada delimitada abaixo. Linhas em preto conectando **a)** e **b)** fazem a correspondência da posição genômica das sondas dos genes *Dwil*\GK14718, *Dwil*\GK15415, *Dwil*\GK19065, *Dwil*\GK14687, *Dwil*\GK19024 e *Dwil*\GK14671 com suas posições citológicas.

O motivo para esta diferença entre a posição genômica e citológica na região entre 6.444.384 e 8.179.762 (**Tabela III.4.2** e **Figura III.4.8**) é devido ao fato de que a linhagem SG12.00 possui a inversão IIL-F no braço cromossômico IIL fixada em relação à linhagem sequenciada Gd-H4-1, que apresenta o arranjo estabelecido padrão para a presente análise. A inversão IIL-F situa-se entre as seções 50C e 53A (**Figura III.4.1**), e citologicamente compartilha seu ponto de quebra distal com o ponto de quebra proximal da inversão IIL-H (Rohde & Valente, 2012).

Surge com isto um quadro de maior complexidade para a delimitação do ponto de quebra proximal da inversão IIL-H. Uma vez que o ponto de quebra distal dessa inversão está estabelecido e se mostra independente, o ponto de quebra proximal pode ter o reuso de uma região específica, envolvida com as duas inversões em nível molecular (Pevzner & Tesler, 2003).

A terceira etapa desta análise consistiu em explorar mais a região entre os genes *DwilGK15415* e *DwilGK19065*, uma região bastante extensa com 286 kb. Para isto, se estabeleceram nove sondas para serem hibridadas na linhagem SG12.00, sendo que oito eram sondas gênicas, e uma sonda era de região intergênica (entre os genes *DwilGK15418* e *DwilGK15419*). À parte, utilizou-se uma abordagem pela técnica da *PCR* para as regiões intergênicas nesta etapa. Entretanto, desconsiderando a região intergênica entre os genes *DwilGK15418* e *DwilGK15419* que era muito extensa, as demais variavam entre 625 pb a 6.915 pb. Neste caso, as *PCRs* realizadas não se mostraram conclusivas para estabelecer entre quais genes, definitivamente, teria ocorrido a quebra proximal da inversão IIL-H (dados não mostrados).

No decorrer da terceira etapa, o genoma de *D. willistoni* passou por uma atualização na tentativa de aprimorar a sua inferência gênica. Nesta atualização, os genes *DwilGK14694* e *DwilGK14693*, que correspondiam respectivamente às sondas 50 e 51, tiveram suas anotações retiradas, estabelecendo uma região intergênica entre os genes *DwilGK15416* e *DwilGK15418* com aproximadamente 66 kb. Assim, a **Tabela III.4.4** já apresenta estes dados corrigidos.

Devido à inversão IIL-F presente na linhagem SG12.00, o gene *DwilGK15416* mapeia fisicamente na seção 50C do braço IIL desta linhagem, a qual seria o limite do ponto de quebra distal dessa inversão no genoma (**Figura III.4.9**). A hibridação *in situ* da sonda do gene *DwilGK15416* na linhagem sequenciada Gd-H4-1, também mapeou fisicamente na seção 50C, e sendo este o último gene anotado que limita o ponto de quebra proximal da inversão IIL-H, e o ponto de quebra distal da inversão IIL-F, este resultado nos mostra que o ponto de quebra proximal da inversão IIL-H no genoma *D. willistoni* está delimitado pelos genes *DwilGK1516* e *DwilGK15418* (**Figura III.4.9**).

A sonda 52b é oriunda do gene *DwilGK15418*, o qual delimita o ponto de quebra proximal da inversão IIL-H no genoma. Na linhagem SG12.00 esta sonda mapeia na seção 53A, na ponta distal do cromossomo com esta inversão, ou seja, fechando o ponto de quebra AC da inversão IIL-H. Entretanto, no genoma disponível da linhagem Gd-H4-1, o intervalo entre este gene e o gene *DwilGK15416* é de 66 kb.

Todas as demais sondas a partir da sonda 50 mapearam fisicamente na seção 53A na linhagem SG12.00 (**Figura III.4.9**), porém na ponta distal, uma vez que a inversão IIL-H utiliza somente uma parte dessa seção citológica. Esta situação pode ser mais bem visualizada no mapeamento físico da sonda 52b (gene *DwilGK15418*), e da própria sonda 50 (anteriormente correspondente ao gene *DwilGK14694*) (**Figura III.4.10**).

A sonda 50 apresentou três sinais de hibridação no braço IIL da linhagem SG12.00, nas seções 53A, 52C e 45A, também apresentando sinais múltiplos no cromossomo X e III desta linhagem. Esta sonda está inserida em uma região no genoma de aproximadamente 13.716 pb (8.299.774 – 8.313.490), entre os genes *DwilGK15416* e *DwilGK15418*. Uma análise desta região para a presença de *TEs* no banco de dados *Repbase* (Jurka *et al.*, 1996) apontou a presença de dezenas de fragmentos pertencentes a distintos elementos móveis descritos para *Drosophila*. Isto implica em altas similaridades destas sequências com outras regiões do genoma em *D. willistoni*, tornando mais difícil o estabelecimento de sondas específicas.

Tabela III.4.4: Marcadores físicos utilizados na terceira etapa da caminhada cromossômica, para o estabelecimento do ponto de quebra proximal da inversão IIL-H na linhagem SG12.00 de *Drosophila willistoni*. Posição genômica e número do *scaffold* de acordo com o banco de dados *Flybase* (St. Pierre *et al.*, 2014), sendo que este último corresponde aos quatro últimos números, os quais todos começam com scf2_110000000. Entre parênteses, número do *scaffold* correspondente à ordenação de acordo com a **Figura III.3.6 (Tópico III.3)**. Inter: Intergênica. III: Sinal de hibridação *in situ* positivo no cromossomo III da linhagem SG12.00.

Sonda	<i>D. willistoni</i> Gene/Intergênica	Número do <i>scaffold</i>	Posição genômica no <i>scaffold</i>	Posição citológica
46	<i>Dwil</i> \GK15416	4521(5)	8.294.814..8.295.756	50C
50	Inter <i>Dwil</i> \GK15416- <i>Dwil</i> \GK15418	4521(5)	8.299.693..8.304.165	53A, 52C, 45A
71b	Inter <i>Dwil</i> \GK15416- <i>Dwil</i> \GK15418	4521(5)	8.320.822..8.322.074	III
70	Inter <i>Dwil</i> \GK15416- <i>Dwil</i> \GK15418	4521(5)	8.331.224..8.332.220	53A
69	Inter <i>Dwil</i> \GK15416- <i>Dwil</i> \GK15418	4521(5)	8.344.262..8.345.444	53A
51	Inter <i>Dwil</i> \GK15416- <i>Dwil</i> \GK15418	4521(5)	8.359.244..8.360.412	53A
52b	<i>Dwil</i> \GK15418	4521(5)	8.360.903..8.361.876	53A
63	Inter <i>Dwil</i> \GK15418- <i>Dwil</i> \GK15419	4521(5)	8.375.331..8.376.737	53A
47	<i>Dwil</i> \GK15419	4521(5)	8.387.445..8.390.092	53B

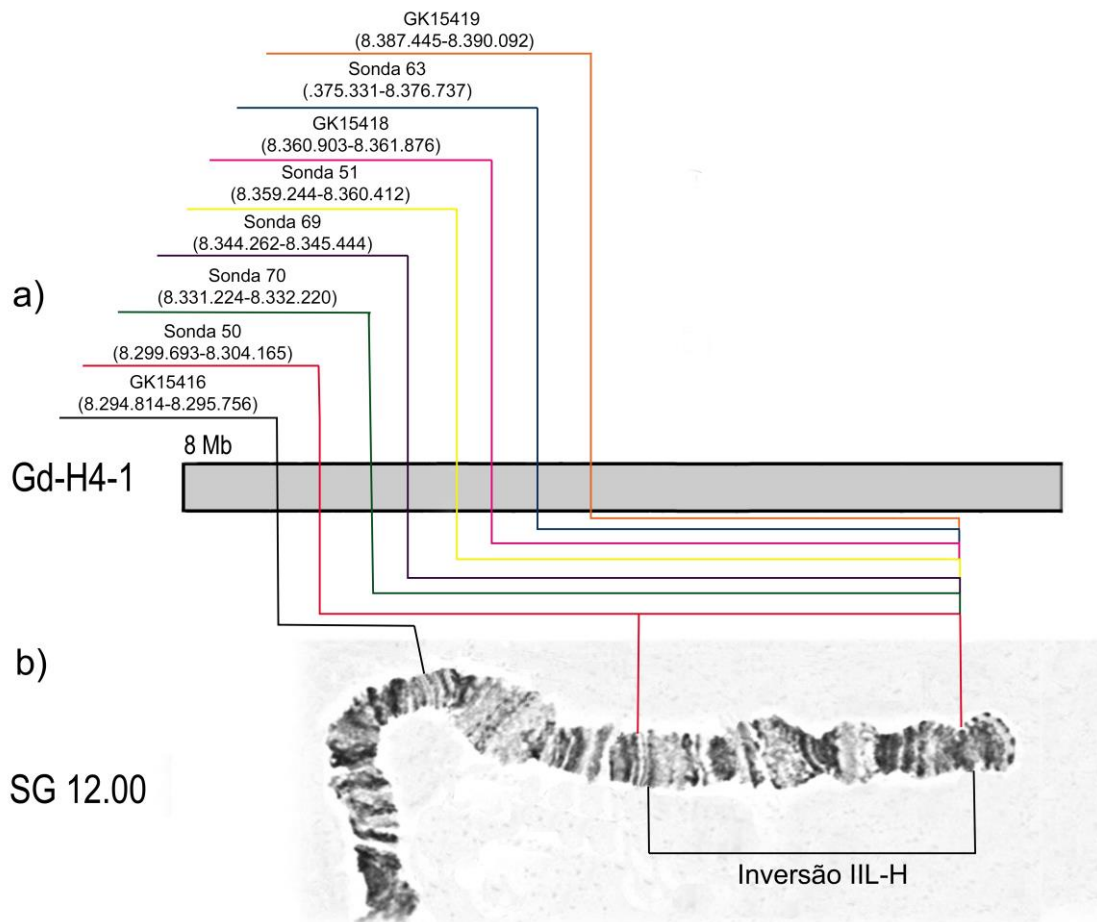


Figura III.4.9: Esquematisação dos resultados obtidos na terceira etapa para a delimitação do ponto de quebra proximal da inversão IIL-H em *Drosophila willistoni* **a)** Representação esquemática do *scaffold* 5 no genoma disponível da linhagem Gd-H4-1, com suas respectivas sondas utilizadas, dando-se um “zoom” na região dos 8 Mb deste *scaffold*. **b)** Ponta do braço IIL da linhagem SG12.00 (Goñi & Valente, 2012), com a inversão IIL-H fixada delimitada abaixo. Linhas conectando **a)** e **b)** fazem a correspondência da posição genômica das sondas do gene *Dwil*GK15416, Sonda 50, Sonda 70, Sonda 69, do gene *Dwil*GK15418, Sonda 63 e o gene *Dwil*GK15419 com as suas posições citológicas. Resultados não estão em escala real.

De fato, quando do estabelecimento da sonda 50 (que contém 968 pb) essa apresentava por BLASTN no banco de dados *Flybase* (St. Pierre *et al.*, 2014), três *hits* significativos (e-value 00) com sequências menores pertencentes aos genes *Dwil\GK13415* (cromossomo III), *Dwil\GK23815* (braço cromossômico IIL, *scaffold* 4) e *Dwil\GK22960* (braço cromossômico IIR). Já o banco de dados *NCBI* (Geer *et al.*, 2010) apontou *hits* significativos com os genes *Dwil\GK13415* e *Dwil\GK22960*. Entretanto, esses três genes também tiveram suas inferências retiradas na atual *release* do genoma de *D. willistoni* (R.1.04). Uma análise por BLASTX da sequência da sonda 50 no banco de dados *NCBI* (Geer *et al.*, 2010), mostrou identidade acima de 90% com prováveis proteínas deduzidas para estes genes, e acima de 80% também com outros genes, com o e-value entre 1e-74 a 9e-108.

No que diz respeito a sinais múltiplos, geralmente esses são devidos à presença de outras cópias de uma mesma família de genes, pseudogenes ou segmentos de DNA que partilham uma determinada sequência com similaridade com a sonda utilizada. O sinal de hibridação mais forte entre as diversas cópias múltiplas é considerado o sinal primário, consistindo na sequência estabelecida para a análise (Hanz *et al.*, 1997). Esta situação tem sido reportada e embasada para o gene *Act5C*, o qual dá origem a seis sinais de hibridação em *D. melanogaster* (Fyrberg *et al.*, 1980), em *D. repleta* e *D. buzzatii* (Hanz *et al.*, 1997), e outras diversas espécies do gênero *Drosophila* (Loukas & Kafatos, 1996). Essas seis cópias, de acordo com suas intensidades, representam outros membros (de menor tamanho) da família gênica da Actina (*Act42A*, *Act57A*, *Act79B*, *Act87E* e *Act88F*). Situação semelhante ocorre com o gene *Hsp70* (Holmgren *et al.* 1979; Hanz *et al.*, 1997).

A busca no banco de dados *Repbase* (Jurka *et al.*, 1996) para *TEs* na sonda 50, mostra uma região de 63 pb, a qual tem 76,12% de similaridade com o retrotransposon *Copia-19_GM-I* de *Glycine max* (Schmutz *et al.*, 2010). Essa similaridade, pelo seu tamanho, muito dificilmente seria visível por hibridação *in situ* não fluorescente. A análise para a ocorrência de quadros abertos de leitura (*ORFs*, do Inglês *Open Reading Frame*) no programa *Geneious* 6.1.5 (<http://www.geneious.com>, Kearse *et al.*, 2012) aponta quase duas dezenas

desses *ORFs* só na sequência utilizada como sonda. Muito provavelmente esta característica teve influência na primeira predição gênica de *D. willistoni*. Entretanto, o fato de uma sequência apresentar *ORFs* não implica que ela seja traduzida (Deonier *et al.*, 2005). Sendo assim, os múltiplos sinais da sonda 50 podem se encaixar em um caso de pseudogenes.

O sinal primário dessa sonda foi o da seção 53A, e isto delimitou o intervalo onde o ponto de quebra proximal ocorreu para aproximadamente 7.034 pb, entre o gene *DwilGK1516* e a sonda 50.

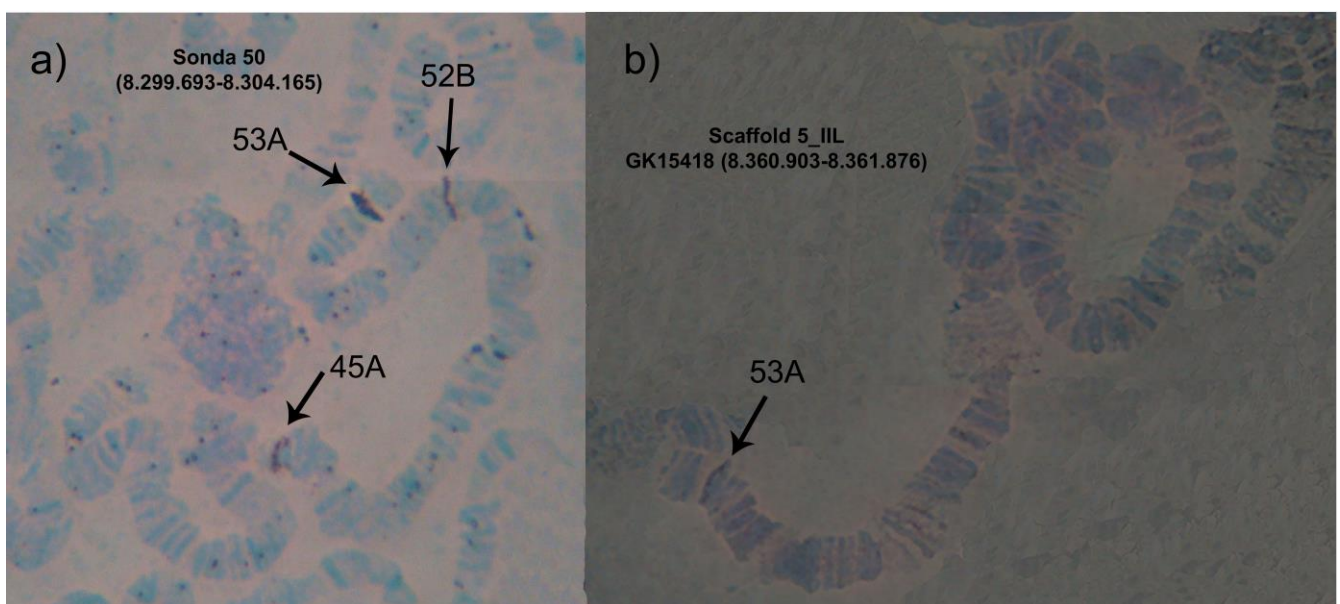


Figura III.4.10: Mapeamento físico das sondas 50 e 52b na linhagem SG12.00 de *Drosophila willistoni*. **a)** Setas indicam o sítio de hibridação da sonda 50 nas seções 53A, 52C e 45A do braço IIL. Essa sonda limitou o ponto de quebra proximal da inversão IIL-H a uma região de aproximadamente 7.034 pb. **b)** Seta indica o sítio de hibridação da sonda 52b na seção 53A do braço cromossômico IIL. A sonda 52b corresponde ao gene *DwilGK15418*, o qual é um dos genes que delimitam o ponto de quebra proximal da inversão IIL-H.

De acordo com o já exposto acima, os resultados obtidos na terceira etapa da delimitação do ponto de quebra proximal da inversão IIL-H delimitam esta inversão no genoma de *D. willistoni* entre os *DwilGK15416* e

Dwil\GK15418. Entretanto, a sonda 50 mostrou estar mais próxima do gene *Dwil*\GK15416, e assim, mais próxima também do local onde ocorreu quebra proximal dessa inversão. Em vista disto, a quarta e final etapa da delimitação do ponto de quebra proximal consistiu em explorar esta região de aproximadamente 7.034 pb entre a sonda 50 e o gene *Dwil*\GK15416. Para isto, estabeleceram-se mais três sondas intergênicas (62iA, 62iB e 50b), bem como foram feitas *PCRs* com os primers dessas sondas, na linhagem Gd-H4-1 e SG12.00 (**Tabela III.4.3**), a fim de estabelecer o local mais específico que delimita o ponto de quebra proximal da inversão IIL-H.

Tabela III.4.5: Marcadores físicos utilizados na quarta etapa da caminhada cromossômica, para o estabelecimento do ponto de quebra proximal da inversão IIL-H na linhagem SG12.00 de *Drosophila willistoni*. Posição genômica e número do *scaffold* de acordo com o banco de dados *Flybase* (St. Pierre *et al.*, 2014), sendo que este último corresponde aos quatro últimos números, os quais todos começam com scf2_110000000. Entre parênteses, número do *scaffold* correspondente à ordenação de acordo com a **Figura III.3.6 (Tópico III.3)**. Inter: Intergênica. III: Sinal de hibridação *in situ* positiva no cromossomo III da linhagem SG12.00. +: Amplificação positiva pela técnica da *PCR*.

<i>D. willistoni</i> gene	Número do <i>scaffold</i>	Posição genômica no <i>scaffold</i>	Posição citológica	Nome da Sonda/ Primer	<i>PCR</i> Gd-H4- 1/SG12.00
Inter <i>Dwil</i> \GK15416- <i>Dwil</i> \GK15418	4521(5)	8.296.060.. 8.297.138	50C	62iA	+/+
Inter <i>Dwil</i> \GK15416- <i>Dwil</i> \GK15418	4521(5)	8.297.347.. 8.298.558	50C, 53A	62iB	+/+
Inter <i>Dwil</i> \GK15416- <i>Dwil</i> \GK15418	4521(5)	8.300.329.. 8.299.716	III	50b	+/+

Os resultados da quarta etapa se mostraram muito exitosos na sua finalidade, uma vez que na linhagem SG12.00 a sonda intergênica 62iA, a qual se posiciona subsequentemente ao gene *Dwii\GK15416* no genoma, mapeou fisicamente na seção 50C. Sendo assim, essa sonda corrobora na linhagem SG12.00 sua posição adjacente à sonda 46. Por sua vez, a sonda 62iB, a qual é constituída de 1.212 pb, e se encontra subsequente à sonda 62iA no genoma, apresentou na linhagem SG12.00 dois sítios de hibridação no braço IIL, um na seção 50C (que nessa linhagem corresponde ao ponto de quebra proximal da inversão IIL-F), e outro na seção 53A (correspondente nessa linhagem ao ponto de quebra distal da inversão IIL-H) (**Figura III.4.11**) As sondas 46 (*Dwii\GK15416*) e 62iB foram também mapeadas na linhagem Gd-H4-1, apontando como seus sítios físicos a seção 50C, com a presença de somente um sinal de hibridação para ambas as sondas.

Os resultados obtidos acima citados, somados aos resultados da delimitação ponto de quebra distal da inversão IIL-H, são esquematizados e resumidos na **Figura III.4.12**. Podemos observar que, com relação ao ponto de quebra distal da inversão IIL-H, as regiões A e C são independentes e não sofrem interferência da inversão IIL-F. Já o quadro muda totalmente com relação ao ponto de quebra proximal da inversão IIL-H, uma vez que na linhagem SG12.00 os resultados apontam que a região compreendida pela sonda, denominada 62iB, é utilizada pelo ponto de quebra proximal da inversão IIL-H, e pelo ponto de quebra distal da inversão IIL-F.

A abordagem pela técnica da *PCR* não foi capaz de apontar a quebra na região 62iB na linhagem SG12.00, porque esta se encontra duplicada nessa linhagem.

As sondas 46 (*Dwii\GK15416*), 62iA, 62iB e 52b (*Dwii\GK15418*) foram sequenciadas diretamente dos seus produtos da *PCR*, em ambas as linhagens Gd-H4-1 e SG12.00, pela empresa de sequenciamento Macrogen Inc. (Seoul, Korea). As sequências consensos obtidas pela análise dos cromatogramas foram submetidas à BLASTN nos bancos de dados *Flybase* (St. Pierre *et al.*, 2014) e *NCBI* (Geer *et al.*, 2010), comprovando em 100% dos casos que as sequências tratavam-se das regiões gênicas escolhidas para as sondas. As

sequências consenso da sonda 62iB nas linhagens Gd-H4-1 e SG12.00 apresentaram 97% de identidade.

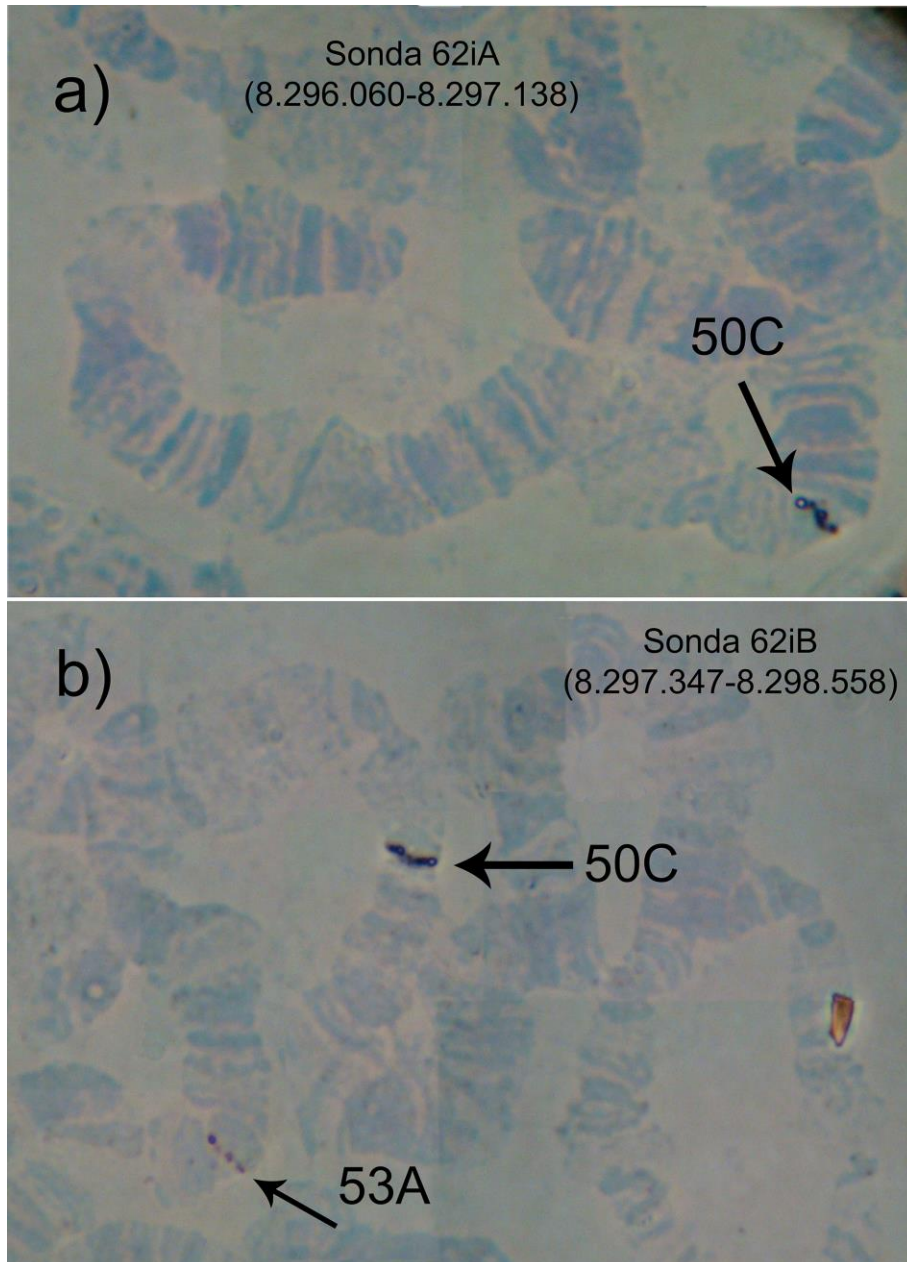


Figura III.4.11: Mapeamento físico das sondas intergênicas 62iA e 62iB na linhagem SG12.00 de *Drosophila willistoni*. **a)** Seta apontando o sítio de hibridação da sonda 62iA na seção 50C. **b)** Seta apontando os sítios de hibridações da sonda 62iB nas seções 50C e 53A.

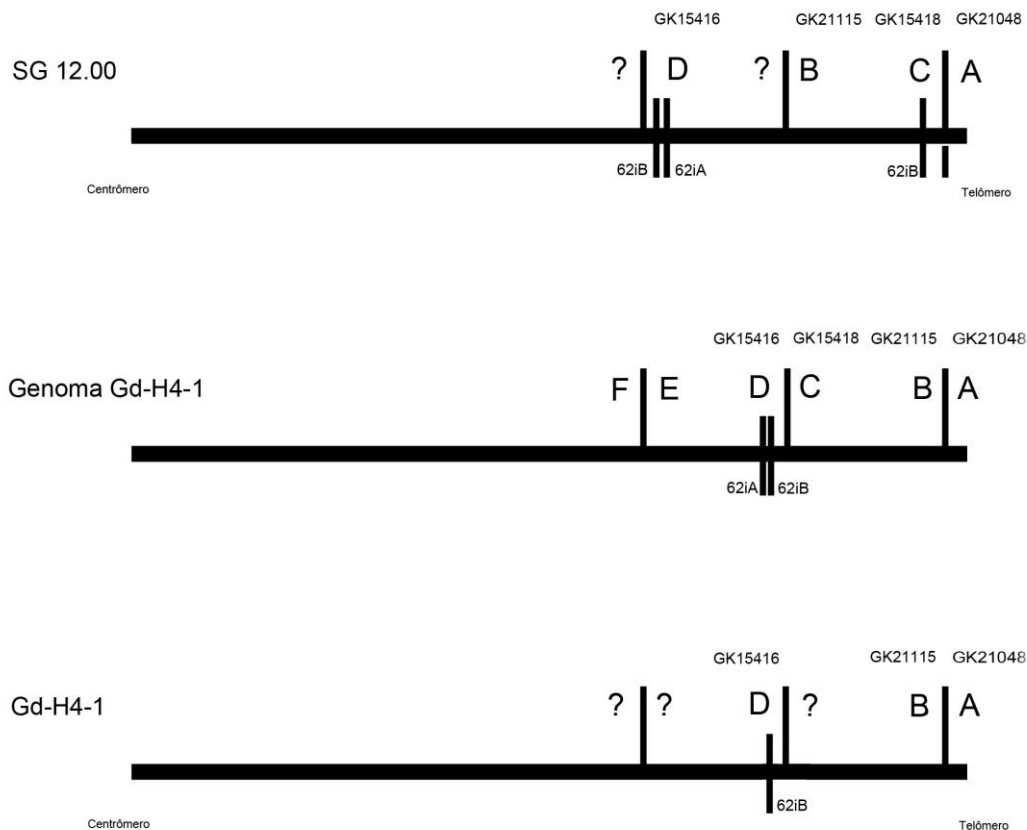


Figura III.4.12: Esquematização dos principais resultados obtidos na caminhada cromossômica, para a delimitação gênica dos pontos de quebra da inversão IIL-H de *Drosophila willistoni*. No topo da figura: esquematização dos resultados da caminhada cromossômica por hibridação *in situ* das sondas 46 (*Dwil*\GK15416), 62iA, 62iB, 52b (*Dwil*\GK15418), 22 (*Dwil*\GK21048) e 72 (*Dwil*\GK21115) na linhagem SG12.00. No centro da figura: esquematização da posição genômica das sondas 46 (*Dwil*\GK15416), 62iA, 62iB, 52b (*Dwil*\GK15418), 22 (*Dwil*\GK21048) e 72 (*Dwil*\GK21115) no genoma da linhagem Gd-H4-1. Na base da figura: esquematização dos resultados da caminhada cromossômica por hibridação *in situ* das sondas 46 (*Dwil*\GK15416), 62iB, 22 (*Dwil*\GK21048) e 72 (*Dwil*\GK21115) na linhagem Gd-H4-1. Ponto de quebra distal da inversão IIL-H correspondente à região AB na linhagem GD-H4-1 e AC na linhagem SG12.00. Ponto de quebra proximal da inversão IIL-H correspondente à região CD no genoma da linhagem Gd-H-1 e B? na linhagem SG12.00. Região EF correspondente ao ponto de quebra proximal da inversão IIL-F no genoma da linhagem Gd-H4-1.

Perante os resultados obtidos na análise da delimitação dos pontos de quebras distal e proximal da inversão IIL-H, dois possíveis cenários globais envolvendo a origem das inversões IIL-H e IIL-F na linhagem SG12.00 emergem.

No primeiro cenário, a inversão IIL-H teria ocorrido primeiro, duplicando a sequência da sonda 62iB nesse processo. Posteriormente, essa região (que no presente estudo compreende 1.212 pb, porém pode englobar mais sequências ao seu redor) seria reutilizada pelo ponto de quebra distal da inversão IIL-F, desta vez levando consigo a sequência da sonda 62iB (**Figura III.4.13a**) para o ponto de quebra proximal da inversão IIL-F.

No segundo cenário teria ocorrido o oposto do primeiro, onde a inversão IIL-F teria ocorrido anteriormente, duplicando a sequência da sonda 62iB. Posteriormente, essa região seria reutilizada pelo ponto de quebra proximal da inversão IIL-H, e desta vez levando consigo a sequência da sonda 62iB (**Figura III.4.13b**).

Um cenário envolvendo a origem de uma das inversões ocorrendo primeiramente na linhagem Gd-H4-1, não explicaria parcimoniosamente a duplicação da sonda 62iB nas seções 50C e 53A na linhagem SG12.00.

Caso as inversões IIL-H e IIL-F fossem inversões independentes em ambos os pontos de quebra, então a sonda 46 (*Dwi\GK15416*) mapearia fisicamente na seção 53A da linhagem SG12.00, ao lado da sonda 22 (*Dwi\GK21048*), delimitando assim o ponto de quebra distal da inversão IIL-H nessa linhagem. Por sua vez, caso estas inversões fossem sobrepostas em um pequeno segmento, considerando o primeiro cenário proposto acima como exemplo, uma parte da sequência da região B estaria mapeando no ponto de quebra proximal da inversão IIL-F na linhagem SG12.00 (**Figura III.4.12**), e vice-versa. Entretanto, o quadro que se encontra é a ocorrência de uma duplicação da sequência da sonda 62iB na linhagem SG12.00. As *PCRs* dessa sequência em ambas as linhagens renderam uma amplificação única. Caso essa sequência tivesse sido particionada em uma das duas linhagens, devido a algum ponto de quebra sobreposto, nesse caso não ocorreria amplificação, pois a sequência estaria corrompida.

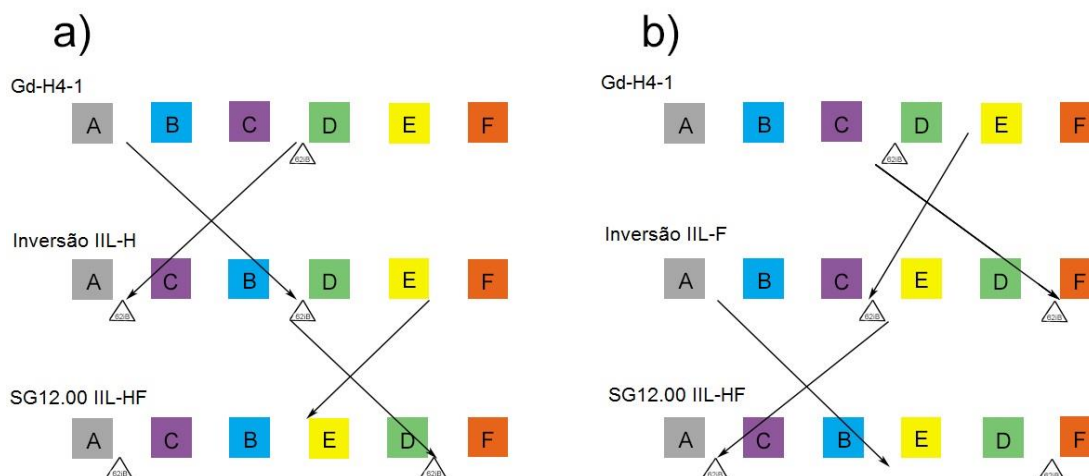


Figura III.4.13: Possíveis cenários para a origem das inversões IIL-H e IIL-F na linhagem SG12.00, de acordo com os resultados obtidos nesse estudo. **a)** Primeiro cenário onde a inversão IIL-H ocorreu anteriormente duplicando a sequência da sonda 62iB. Posteriormente, a inversão IIL-F reutilizaria esta sequência levando-a para seu ponto de quebra proximal na linhagem SG12.00. **b)** Segundo cenário onde a inversão IIL-F teria ocorrido anteriormente, duplicando a sequência da sonda 62iB, posteriormente esta região seria reutilizada para a quebra pelo ponto de quebra proximal da inversão IIL-H, e desta vez levando consigo a região 62iB. Triângulos representam a sequência da sonda 62iB.

Os resultados apresentados na presente análise convergem para o mais provável reuso, pelas inversões IIL-H e IIL-F, da sequência compreendida pela sonda 62iB na linhagem SG12.00. Situações semelhantes também foram encontradas por Calvete *et al.* (2012) na análise dos pontos de quebra das inversões $2m$ e $2n$ em *D. buzzatii*, e por Puerma *et al.* (2014) na análise dos pontos de quebra das inversões E_1 e E_2 em *D. subobscura*.

As causas do reuso dos pontos de quebras é muito pouco esclarecida. Uma de suas principais características é estar associado à sequências que sejam propensas a participarem de eventos que origem rearranjos

cromossômicos. Dentre estas sequências se destacam as duplicações segmentares (Ruiz-Herrera et al., 2006) e sequências ricas em AT, as quais se denominam sítios frágeis (Zhang & Freudenreich, 2007). As análises *in silico* de evolução cromossômica e reuso baseadas nos genomas das espécies de *Drosophila* disponíveis, feita por von Grotthuss et al. (2010), apontaram que os sítios frágeis são as sequências predominantemente envolvidas nos pontos de quebras de inversões, e assim, provavelmente os maiores responsáveis pela reorganização gênica no gênero.

A sonda 62iB delimitou o ponto de quebra proximal da inversão IIL-H em aproximadamente 2.816 pb no genoma de *D. willistoni*. As buscas no banco de dados *Repbase* (Jurka et al., 1996) para a ocorrência de *TEs* nessa região, mostrou a presença de cinco similaridades: 76,60% de similaridade com o transposon *hAT* de *Crassostrea gigas* (ostra do Pacífico, 130 pb) (Bao & Jurka, 2013); 85,11% de similaridade com o *TE Helitron-N4B_ALy* (*Arabidopsis lyrata*, 45 pb) (Kojima & Jurka, 2013); 76,19% de similaridade com o retrotransposon *Copia-52_GM-I* (*Glycine max*, 81 pb) (Schmutz et al., 2010); 82,81% de similaridade com o transposon *Kolobok-N2_HM* (*Hydra magnipapillata*, 72 pb) (Bao, 2015) e 77,36% de similaridade com o retrotransposon *Tx1-12_DR* (*Danio rerio*, 52 pb) (Bao & Jurka, 2014), este último situado na sonda 62iB. Essa busca também mostrou a ocorrência de duas repetições simples (TTGCA)_n e (TTGTTC)_n com 39 e 34 pb respectivamente. De maneira geral, as regiões de ambos os pontos de quebra da inversão IIL-H se apresentam, de acordo com os dados disponíveis, com poucas sequências repetitivas no genoma de *D. willistoni*.

A análise da sintonia da região do ponto de quebra proximal da inversão IIL-H no genoma da linhagem Gd-H4-1, com relação aos respectivos genes ortólogos em *D. melanogaster*, mostrou que assim como no ponto de quebra distal, a manutenção da paridade desses genes nesta espécie. Uma inspeção mais detalhada desta relação de sintonia feita manualmente, expandiu esta comparação para 200 kb nas regiões dos pontos de quebras distal (AB) e proximal (CD) nos genomas de *D. willistoni* e *D. melanogaster*. Esta análise apontou o seguimento da alta conservação destas duas regiões entre as duas espécies, com a presença de prováveis micro-inversões em alguns casos.

Entretanto, a região ortóloga do ponto de quebra distal da inversão IIL-H se encontra em uma região mais proximal no cromossomo 2L dessa espécie. Já a região proximal do ponto de quebra da inversão IIL-H, em *D. melanogaster* se encontra em uma região distal do seu cromossomo 2L.

As espécies do gênero *Drosophila* são reconhecidas por terem elevado grau de reordenação do seu conteúdo gênico nos diferentes Elementos de Muller. Sendo assim, a quebra da sintenia entre espécies distintas é devido às diversas inversões paracêntricas ocorridas no decorrer da história evolutiva destas (Bhutkar *et al.*, 2008). A manutenção da sintenia nas regiões dos pontos de quebra entre *D. willistoni* e *D. melanogaster*, se somam aos possíveis cenários obtidos para a origem das inversões IIL-H e IIL-F nessa análise (**Figura III.4.13**), e implicam em uma mais provável condição basal do arranjo cromossômico da linhagem Gd-H4-1 com relação ao arranjo da linhagem SG12.00.

A delimitação dos pontos de quebras das inversões $2m$ e $2n$ em *D. buzzatti* por Calvete *et al.* (2012), as quais compartilhavam um dos pontos de quebra citologicamente, foi primeiramente baseada no mapa físico de BACs dessa espécie (González *et al.* 2005); e no genoma da espécie relacionada *D. mojavensis* (*Drosophila* 12 genomes Consortium, 2007), a qual não apresentava as inversões. Por meio de caminhada cromossômica por hibridação *in situ*, utilizando como sondas os BACs que delimitavam o ponto de quebra dessas inversões, foi possível estabelecer quais desses clones continham as regiões dos três pontos de quebra na linhagem invertida (AC, BE, DF). Estes BACs positivos tiveram suas porções terminais sequenciadas e estas sequências serviram de base para delimitar os pontos de quebra (AB, CD, EF) no genoma de *D. mojavensis*. Subsequentemente, sondas baseadas no genoma de *D. mojavensis* foram mapeadas fisicamente nos cromossomos politênicos de *D. buzzatii*, permitindo assim a delimitação gênica dos três pontos de quebra das inversões $2m$ e $2n$.

Corbett-Detig *et al.* (2012) estabeleceram um método *in silico* que permite a delimitação e a análise dos pontos de quebra das inversões segregantes de *D. melanogaster*. Esse método se baseia nos genomas

disponíveis pelo *Drosophila Population Genomic Project* (<http://www.dpgp.org>). Conforme explicado resumidamente pelos autores, leituras de 76 pb terminais, separadas por 300 pb de diferentes isolinhagens oriundas, nessa análise, de populações da África e de uma população da França, são alinhadas com o genoma referência de *D. melanogaster* (Adams, 2000). Esses alinhamentos seguem parâmetros de eliminação para leituras de sequências com as mesmas coordenadas e as que apresentem excesso, ou pouca indentidade entre si, dentre outros. Esse estudo tinha a intenção de comprovar que o método é capaz de estabelecer os pontos de quebra das inversões em espécies de *Drosophila*. Sendo assim, ele foi capaz de delimitar, para ambos os pontos de quebra, cinco inversões na espécie *D. melanogaster*, os quais são corroborados por estudos anteriores (Andolfatto *et al.* 1999; Matzkin *et al.* 2005; Ranz *et al.*, 2007). Entretanto, os autores ressaltam que este método tem limitações com relação às regiões com a presença de sequências repetitivas, ou *TEs*; bem como o método requer um genoma de referência altamente caracterizado por mapeamento físico.

As inversões E_1 e E_2 de *D. subobscura* também compartilham citologicamente um dos seus pontos de quebra. Puerma *et al.* (2014) estabeleceram a delimitação gênica destes pontos de quebra em uma linhagem padrão (sem as inversões), e em uma linhagem E_{1+2} , por caminhada cromossômica, com base em quatro fagos recombinantes que já haviam sido mapeados fisicamente próximos à estas regiões (Papaceit *et al.* 2006). Destes fagos foram feitas subclonagens. Estas subclonagens tiveram suas porções terminais sequenciadas e alinhadas com o genoma da espécie relacionada *D. pseudoobscura* (Richards *et al.*, 2005). Esta estratégia serviu como base para a delimitação de outras sondas que foram hibridadas nos cromossomos politêncos de *D. subobscura*; estabeleceu os pontos de quebra dessas inversões na espécie *D. pseudoobscura*, bem como a delimitação gênica dos três pontos de quebras envolvendo as inversões E_1 e E_2 .

A delimitação dos pontos de quebra da inversão IIL-H de *D. willistoni*, no presente estudo, se baseou no estabelecimento de marcadores físicos no cromossomo II dessa espécie (Garcia *et al.*, 2015; **Tópico III.3** e **Capítulo V** desta Tese), e no estabelecimento de sondas com base no genoma da

linhagem Gd-H4-1 (*Drosophila* 12 Genomes Consortium, 2007) para caminhada cromossômica nos cromossomos politênicos da linhagem SG12.00. Como não se dispunha de um mapa físico bem delimitado, ou mesmo o genoma montado de outra linhagem, esta análise não pode se espelhar totalmente nos mais recentes estudos feitos por Calvete *et al.* (2012), Corbett-Detig *et al.* (2012) e Puerma *et al.* (2014).

Um fato também constatado na caminhada cromossômica, foi de que as sondas 50b e 71b apresentaram sinais de hibridação *in situ* no cromossomo III, implicando possíveis quimerismo na montagem da região do ponto de quebra proximal da inversão IIL-H.

Finalizando, os resultados encontrados na presente Tese reforçam a premissa do poder de resolução da hibridação *in situ* não fluorescente nos cromossomos politênicos das espécies de *Drosophila*, uma vez que essa técnica fornece bases para a melhor montagem de genomas, para delimitar pontos de quebras de inversões e demonstra a ocorrência de reuso de sequências nos pontos de quebras de inversões (*Drosophila* 12 Genomes Consortium, 2007; Prada *et al.*, 2001).

A inversão IIL-H no genoma de *D. willistoni* consiste em aproximadamente em 5,51 Mb. Devido ao uso da linhagem SG12.00 na caminhada cromossômica para a delimitação dos pontos de quebra dessa inversão, a presente análise também resultou em um mapeamento físico de alta resolução dessa região na referida linhagem.

Estando aqui os pontos de quebra da inversão IIL-H delimitados no genoma da linhagem Gd-H4-1, e também citologicamente na linhagem SG12.00, as análises futuras se dedicarão ao rastreamento na biblioteca genômica de fagos λ contruída para a linhagem SG12.00, com as sondas 74 e 75 (**Tópico II.7, Capítulo II**). A construção da biblioteca genômica já foi realizada pela empresa *Rx Biosciences* (EUA). As sondas foram estabelecidas na presente Tese e serão enviadas para a mesma empresa, que também fará o rastreamento.

Este rastreamento será para isolar os fagos positivos que contenham as regiões AC e BE dos pontos de quebra da inversão IIL-H na linhagem

SG12.00. Esta abordagem permitirá o sequenciamento destes fagos positivos, para a análise comparativa da região desses pontos de quebra entre as duas linhagens analisadas nesta Tese.

Nosso objetivo final com isto é estabelecer o mecanismo envolvido na origem da inversão ILL-H em *Drosophila willistoni*.

CAPÍTULO IV

Conclusões e Perspectivas

Diante dos resultados do presente trabalho, conclui-se que:

- A linhagem sequenciada de *Drosophila willistoni*, a qual é oriunda da Ilha Guadalupe (Caribe), apresenta distinto arranjo cromossômico com relação ao braço IIL do fotomapa aprimorado da espécie.
- Estas distinções são devidas à presença das inversões IIL-A e IIL-F, ocorrentes na linhagem Gd-H4-1 em homozigose. Como a análise dos pontos de quebra de inversões requer um alinhamento entre o padrão citológico e o padrão genômico das linhagens utilizadas, este estudo passou a considerar a linhagem sequenciada Gd-H4-1 como a padrão para o desenvolvimento deste tipo de análise na espécie.
- Uma vez que o presente estudo visa delimitar os pontos de quebra da inversão IIL-H de *Drosophila willistoni*, esta análise requer outra linhagem dessa espécie, a qual tenha a inversão IIL-H fixada com relação à linhagem Gd-H4-1. Sendo assim, escolheu-se a linhagem SG12.00 do Uruguai, que apresentava seu arranjo cromossômico determinado com relação ao fotomapa da espécie.
- As diferenças de arranjo cromossômico no braço IIL entre a linhagem sequenciada e a linhagem SG12.00 foram estabelecidas por meio de cruzamentos recíprocos, e mostrou que estas diferem pela ocorrência das inversões IIL-A, IIL-F e IIL-H fixadas em SG12.00.
- Dada a escassez de marcadores genéticos disponíveis para *Drosophila willistoni*, sugerimos mais uma vez que a montagem do seu genoma deve ser ainda vista com cautela. Em vista disto, foi feita uma revisão da ordem e da

orientação dos *scaffolds* do cromossomo II de *Drosophila willistoni*, via o estabelecimento de 18 novos marcadores físicos para a linhagem sequenciada Gd-H4-1. Esta análise mostrou que em realidade os braços cromossômicos IIL e IIR de *Drosophila willistoni* correspondem aos elementos de Muller B e C respectivamente, ao contrário do que se vinha tradicionalmente inferindo.

- O estabelecimento do ponto de quebra distal da inversão IIL-H deu-se por meio de caminhada cromossômica por hibridação *in situ* não fluorescente, nos cromossomos politênicos da linhagem SG12.00, e pela técnica da *PCR*, delimitando esse ponto de quebra entre os genes *Dwil\GK21048* e *Dwil\GK21115*.

- Já o estabelecimento do ponto de quebra proximal deu-se por meio de caminhada cromossômica por hibridação *in situ* não fluorescente, nos cromossomos politênicos da linhagem SG12.00. Esta análise mostrou que uma sequência de aproximadamente 1.212 pb está duplicada no ponto de quebra proximal da inversão IIL-F, e no ponto de quebra distal da inversão IIL-H. Este resultado implica em um possível reuso desta região por ambas as inversões.

- Tanto o ponto de quebra distal da inversão IIL-H, quanto o seu ponto de quebra proximal, apresentaram no genoma baixa ocorrência de sequências similares a Elementos de Transposição, ou quaisquer outras sequências repetitivas.

- A análise da sintenia das regiões dos pontos de quebra com relação aos genes ortólogos na espécie *Drosophila melanogaster*, mostrou alta conservação entre esta e a linhagem Gd-H4-1; uma situação que implica a possível condição basal do arranjo cromossômico da linhagem sequenciada, com relação ao arranjo da linhagem SG12.00.

Perspectivas:

- Analisar as regiões dos pontos de quebra da inversão IIL-H na linhagem SG12.00. Para isto, se mandou construir uma biblioteca genômica de fagos λ na empresa Rx Biosciences (EUA), e se estabeleceram sondas para o isolamento dos fagos positivos que contenham regiões A e B do ponto de quebra distal. A análise comparativa dessas regiões em ambas as linhagens será determinante para a inferência do provável mecanismo envolvido na origem desta inversão. Esta abordagem será concluída imediatamente após a defesa desta Tese.
- Estabelecer os genes que delimitam os pontos de quebra da inversão IIL-F na linhagem sequenciada Gd-H4-1 e na linhagem SG12.00 de *Drosophila willistoni*, para futuramente tentar inferir o provável mecanismo de origem desta inversão.
- Analisar o efeito de ambas as inversões (IIL-H e IIL-F) na expressão gênica e recombinação genética das sequências envolvidas.

CAPÍTULO V

Reassignment of *Drosophila willistoni* genome scaffolds to chromosome II arms

Carolina Garcia^{*§}, Alejandra Delprat[§], Alfredo Ruiz[§], Vera L S Valente^{*}

* Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Brazil 15053

§ Departament de Genètica i de Microbiologia, Facultat de Biociències, Universitat Autònoma de Barcelona, Barcelona, Spain 08193

Running title: Chromosome II scaffolds in *Drosophila willistoni*

Key words: *Drosophila willistoni*, polytene chromosome II, physical markers, *in situ* hybridization, Muller elements

Corresponding autor: Vera L S Valente, Departamento de Genética, Av. Bento Gonçalves, 9500, 43323M/210, Postal Code: 91501-970 - Porto Alegre, RS, Brasil, Phone (55-51) 3308-6713. E-mail: 00004887@ufrgs.br

Artigo publicado na revista G3: Genes, Genomes, Genetics

G3 g3.115.021311; Early Online October 4, 2015, doi:10.1534/g3.115.021311

Reassignment of *Drosophila willistoni* Genome Scaffolds to Chromosome II Arms

Carolina Garcia,^{*,†} Alejandra Delprat,[†] Alfredo Ruiz,[†] and Vera L. S. Valente^{*,†}

^{*}Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Brazil 15053, and

[†]Departament de Genètica i de Microbiologia, Facultat de Biociències, Universitat Autònoma de Barcelona, Barcelona, Spain 08193

ABSTRACT *Drosophila willistoni* is a geographically widespread Neotropical species. The genome of strain Gd-H4-1 from Guadeloupe Island (Caribbean) was sequenced in 2007 as part of the 12 *Drosophila* Genomes Project. The assembled scaffolds were joined based on conserved linkage and assigned to polytene chromosomes based on a handful of genetic and physical markers. This paucity of markers was particularly striking in the metacentric chromosome II, comprised two similarly sized arms, IIL and IIR, traditionally considered homologous to Muller elements C and B, respectively. In this paper we present the cytological mapping of 22 new gene markers to increase the number of markers mapped by *in situ* hybridization and to test the assignment of scaffolds to the polytene chromosome II arms. For this purpose, we generated, by polymerase chain reaction amplification, one or two gene probes from each scaffold assigned to the chromosome II arms and mapped these probes to the Gd-H4-1 strain's polytene chromosomes by non-fluorescent *in situ* hybridization. Our findings show that chromosome arms IIL and IIR correspond to Muller elements B and C, respectively, directly contrasting the current homology assignments in *D. willistoni* and constituting a major reassignment of the scaffolds to chromosome II arms.

KEYWORDS

Drosophila willistoni
polytene
chromosome II
physical markers
in situ
hybridization
Muller elements

Drosophila willistoni belongs to the *willistoni* species group of the subgenus *Sophophora*. With a widespread geographic distribution in the Neotropical region (Val *et al.* 1981; Bächli 2015), it is one of the most polymorphic species for chromosomal inversions in the entire genus *Drosophila* (Sperlich and Pfriem 1986; Rohde and Valente 2012) and is a classic model organism for research in population and evolutionary genetics, speciation, ecological genetics, and molecular evolution (Ehrman and Powell 1982; Krimbas and Powell 1992; Markow and O'Grady 2007). The karyotype of this species consists of two metacentric chromosomes (X and II) and a rod chromosome (III) (Dobzhansky and Pavan 1943).

Muller (1940) proposed that, in *Drosophila*, the genic content of chromosome arms has been conserved despite numerous rearrangements that have occurred during the evolution of the genus, a notion that has been supported widely in subsequent studies (Bhutkar *et al.* 2008). The current chromosomal arm homologies for *D. willistoni* are as follows (Schaeffer *et al.* 2008): XL and XR are homologous to *D. melanogaster* X and 3L (Muller elements A and D), IIL and IIR are homologous to *D. melanogaster* 2R and 2L (Muller elements C and B), whereas chromosome III is homologous to *D. melanogaster* 3R and 4 (Muller elements E and F).

The genome of *D. willistoni* (strain Gd-H4-1, Guadeloupe Island, Caribbean) was completely sequenced and assembled in 2007 by the *Drosophila* 12 Genomes Consortium. The scaffolds were joined based on conserved linkage in the orthologous genes located at the scaffold ends (Schaeffer *et al.* 2008). Subsequently, scaffolds were anchored to the polytene chromosomes based on a handful of genetic markers, mainly enzyme loci and other genes taken from classical studies (Ayala *et al.* 1972; Lakovaara and Saura 1972; Dobzhansky and Powell 1975), and less than 10 gene markers mapped by *in situ* hybridization (ISH). There is scarcity of gene markers occurring in the metacentric chromosome II, which shows a high level of segregating inversions, mainly in the IIL arm (Rohde and Valente 2012). This chromosome has two similarly sized arms named IIL and IIR (Dobzhansky 1950), which have been traditionally considered homologous to Muller elements C and B, respectively (Lakovaara and Saura 1972; review in Sorsa 1988). Eight scaffolds

Copyright © 2015 Garcia *et al.*

doi: 10.1534/g3.115.021311

Manuscript received August 13, 2015; accepted for publication September 17, 2015; published Early Online October 4, 2015.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supporting information is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.115.021311/-/DC1

[†]Corresponding author: Departamento de Genética, Av. Bento Gonçalves, 9500, 43323M/210, Postal Code: 91501-970 - Porto Alegre, RS, Brazil. E-mail: 00004887@ufrgs.br

■ **Table 1** Gene markers used for scaffold reassignment to *Drosophila willistoni* chromosome II listed from centromere to telomere

<i>D. willistoni</i> Gene	<i>D. melanogaster</i> Ortholog Gene	Scaffold Number	Scaffold Orientation	Scaffold Position of Gene	Cytological Position	Primers F and R (5'-3')
III arm (Muller B element)						
<i>Dwil</i> GK18743	<i>Dmel</i> Lamp1	4884	+	878,652.0.883,516	37B	GTCAAGCAGTAGCAGCACCA GCCACGCGAAGTTGATCGAC
<i>Dwil</i> GK24519	<i>Dmel</i> CG9515	4585	+	282,915.0.284,232	38A	ATTCATTACAGCACAAACC GGACTCAATCGGAACTATG
<i>Dwil</i> GK18260	<i>Dmel</i> Yuri	4577	-	1,906,481.0.1,918,823	45A	GTGAAGAGCCTACACACAGC CTTCTGAGATGATCCACGAC
<i>Dwil</i> Adh	<i>Dmel</i> Adh	4577	-	2,980,068.0.2,982,042	44A	CATGGAACGTGTTAAGTGCC AGTTCACAGCAATGGTACGC
<i>Dwil</i> GK23840	<i>Dmel</i> CG17549	4516	+	60,623.0.62,773	46A	GGAGGATATGCTGGTGGTTA GTGCTGACTTGCTCCAACCTG
<i>Dwil</i> GK15054	<i>Dmel</i> CG7371	4521	+	716,961.0.719,512	47A	ATCCTGAGCCTGAGTTCAC CGCCAAGAGAATTGTATCCG
<i>Dwil</i> GK12721	<i>Dmel</i> CG13127	4752	+	52,463.0.54,862	55A	GCAGCTCGATGAACTCTATG TCTCCGAAGACTGTGTACTC
<i>Dwil</i> GK18432	<i>Dmel</i> fusl	4945	+	417,864.0.419,442	55A	GCATCAGCCTCATATCCATC GTCAACACTCTCGGCTCCAG
<i>Dwil</i> GK21099	<i>Dmel</i> Rab3-GAP	4851	-	226,612.0.230,981	55C	CTGGAGCAGTCAAGCGGAGA ATCCAAGCATCTAAGCGTG
IIIR arm (Muller C element)						
<i>Dwil</i> GK23049	<i>Dmel</i> bw	4954	+	2,790,903.0.2,793,354	56C	CAGTAGTAACCACTCCGATG GCGGACACATTGTCTACCAG
<i>Dwil</i> GK22144	-	4558	+	73,772.0.80,277	58C	CATTGACGATCTCAGCAAC TCACTTCGGACTACTCCAGC
<i>Dwil</i> GK22138	<i>Dmel</i> px	4558	+	211,240.0.284,571	59B	GCTGCATTAGATCCTCATAG GGCAGCCAACAGTCCATACA
<i>Dwil</i> vlc	<i>Dmel</i> vlc	4382	-	1,240,131.0.1,244,696	62A	CAACGCCACTATCTGTGAAG TCTCATTGCACTCACCTACA
<i>Dwil</i> GK17912	<i>Dmel</i> FLASH	4822	+	150,635.0.154,598	64B	ACGATGATCTGGATGAGTTC TACAACATACCTAGTTCCGC
<i>Dwil</i> GK20645	<i>Dmel</i> CG2269	4510	-	3,566,941.0.3,573,564	68C	CTGATGGACACCACAGAGTG ACTCGTACAACATGGCGGAC
<i>Dwil</i> GK15808	<i>Dmel</i> Ir60a	4514	-	205,020.0.207,611	71A	AACGAGGCGAGTCACCGATAC CATATCGGACGCTCTTGAAC
<i>Dwil</i> Adam	<i>Dmel</i> Adam	4513	+	22,848.0.24,439	72A	GGTGAGGATGACGATGAGGA TCCGAATGTAAGAGTCCAC
<i>Dwil</i> GK19495	<i>Dmel</i> Hsf	4512	+	1,798,077.0.1,801,918	77A	GGCTACCGTCAATAGATCAG AGAACATACGTGGACGTCAG

The scaffold number corresponds to the last four numbers of the scaffolds, which all start with scf2_110000000. The scaffold number and scaffold position of genes correspond to the material available in the FlyBase database (St. Pierre *et al.* 2014). The scaffold orientation is according to Schaeffer *et al.* (2008)

were assigned to each of the two arms of this chromosome in the Gd-H4-1 strain, based on linkage information for five genetic markers and the physical mapping of another two markers. Thus, the current assignment of scaffolds to *D. willistoni* chromosome II arms should be viewed with caution and as provisional, as stated by Schaeffer *et al.* (2008).

To increase the number of markers mapped by ISH and to test the assignment of scaffolds to chromosome II in strain Gd-H4-1, we mapped 18 new markers for chromosome II by using nonfluorescent ISH. Additionally, three gene markers were physically mapped to chromosome X and one gene marker mapped to chromosome III, constituting a total of 22 new gene markers.

MATERIALS AND METHODS

Drosophila stock and cytological preparations

The strain of *D. willistoni* used in this study was Gd-H4-1. This strain comes from Guadeloupe Island in the Caribbean, and it was used for sequencing the complete genome of this species (*Drosophila* 12 Genomes Consortium

2007). For that purpose, the high degree of polymorphism for inversions naturally present in this species (Rohde and Valente 2012) was virtually solved by successive inbreeding (*Drosophila* 12 Genomes Consortium 2007). The stock was purchased from the UC San Diego Stock Center and maintained at 21° in standard cornmeal–agar–yeast culture medium.

The cytological preparations of polytene chromosomes for the nonfluorescent ISH were made with salivary glands from larvae in the third stage of development fed with extra fresh yeast. Salivary glands were dissected in saline solution and immediately transferred to a coverslip with a drop of 45% acetic acid for 3 min for fixation of the glands. Then, the excess of acetic acid was removed, and a drop of acetic acid:water:lactic acid mixture in 3:2:1 (v/v) ratio was added and rested for 4 min. The material was then placed on a slide, spread, and squashed. All the analyses of chromosomal preparations were done on a phase contrast NIKON Optiphot-2- microscope.

Gene probes

Release 1.04 of genome of *D. willistoni* available in FlyBase database (St. Pierre *et al.* 2014), was used to generate 22 gene probes. Two probes

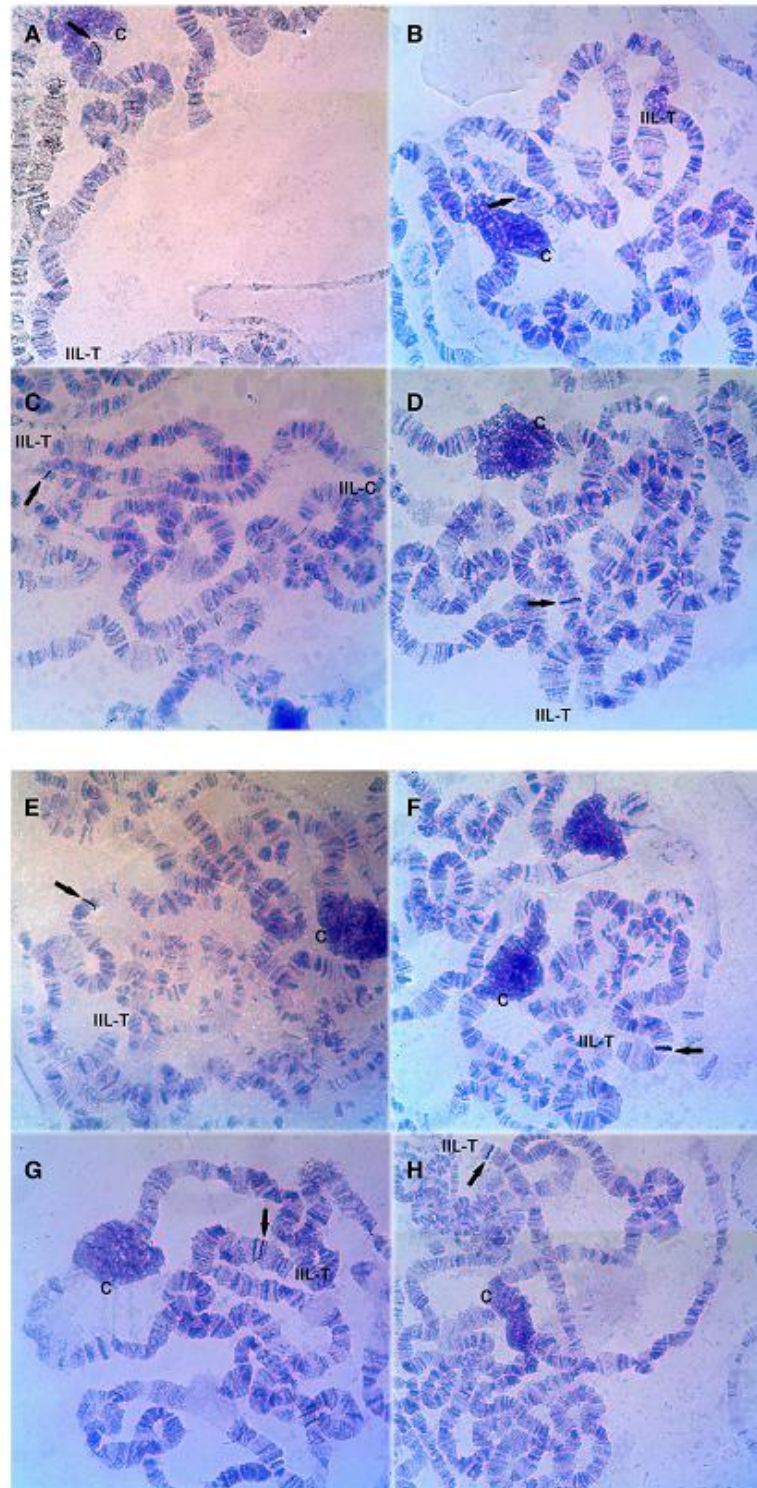


Figure 1 Gene probes hybridized to the *D. willistoni* chromosome III arm ordered from centromere to telomere (A) *Dwi*\GK18743 (scaffold 4884); (B) *Dwi*\GK24519 (scaffold 4585); (C) *Dwi*\GK18260 (scaffold 4577); (D) *Dwi*\GK23840 (scaffold 4516); (E) *Dwi*\GK15054 (scaffold 4521); (F) *Dwi*\GK12721 (scaffold 4752); (G) *Dwi*\GK18432 (scaffold 4945); and (H) *Dwi*\GK21099 (scaffold 4851). The black arrows indicate hybridization signals. ILL-T, III arm telomere; ILL-C, III arm centromere; C, chromocenter.

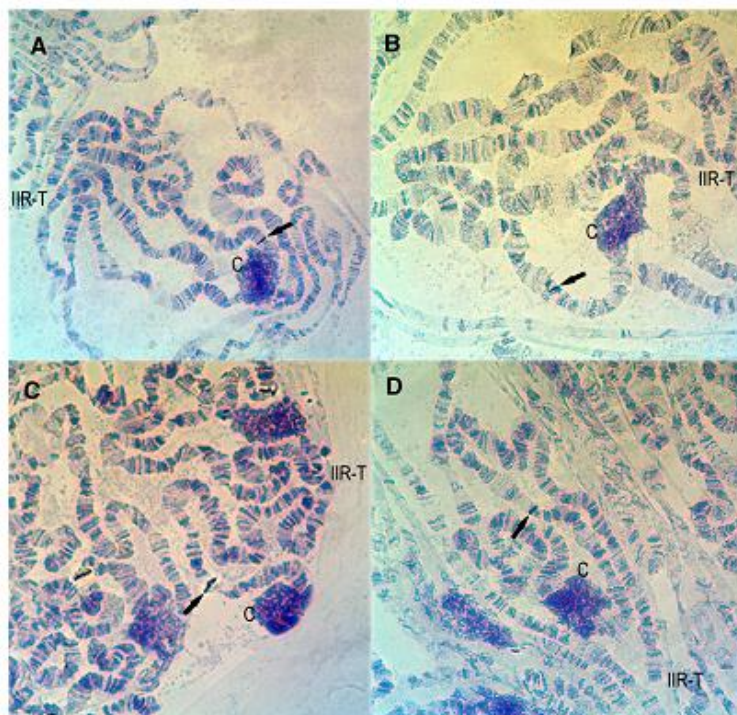


Figure 2 Gene probes hybridized to *D. willistoni* IIR arm ordered from centromere to telomere (A) *Dwi*\GK23049 (scaffold 4954); (B) *Dwi*\GK22144 (scaffold 4558); (C) *Dwi*\GK22138 (scaffold 4558); (D) *Dwi**vlc* (scaffold 4382); (E) *Dwi*\GK17912 (scaffold 4822); (F) *Dwi*\GK20645 (scaffold 4510); (G) *Dwi*\GK15808 (scaffold 4514); (H) *Dwi**Adam* (scaffold 4513); and (I) *Dwi*\GK19495 (scaffold 4512). The black arrows indicate hybridization signals. IIR-T, IIR arm telomere; C, chromocenter.

were designed on scaffolds 4577 (III) and 4558 (IIR); for the rest of the chromosome II scaffolds, only one gene was chosen (Table 1). Additionally, three probes were designed for chromosome X and one for chromosome III (supporting information, Table S1). All the polymerase chain reaction (PCR) primers were designed with Primer Designer v.1.01 (Scientific and Educational Software).

PCR amplifications were performed with Taq DNA Polymerase, recombinant (Invitrogen) following the manufacturer's instructions. PCR products were purified with Exonuclease I (USB) and Shrimp Alkaline Phosphatase (USB) and sequenced by Macrogen INC (Seoul, Korea). All sequences confirmed by BLASTN in the FlyBase database (St. Pierre *et al.* 2014).

Probes were labeled with a Biotin PCR Labeling Core Kit (Jena Bioscience) per the manufacturer instructions, with minor adjustments. The labeled products were purified with a NucleoSpin Gel and PCR Clean-up Kit (MACHEREY-NAGEL).

ISH and mapping

Nonfluorescent ISHs were carried out at 37° using 100–150 ng of biotin-labeled probe on each chromosomal slide, following the protocol of Biémont *et al.* (2004) with minor modifications: 1) The slides were washed in 70% ethanol before the later washing with 95% ethanol in the pre-hybridization step; 2) two washes with 2× saline sodium citrate at 37° for 10 min and one wash for a few seconds in 2× saline sodium citrate at room temperature after the hybridization step; 3) one wash with 1× phosphate-buffered saline (PBS) just after the washes with 0.1% Triton in 1× PBS after the detection step; 4) one wash with 0.1% Triton in 1× PBS before the wash with 1× PBS after the revelation step.

Hybridization signals were detected and revealed using a Vectastain ABC KIT (Vector Laboratories) and DAB Substrate (Roche), respectively. Then, the slides were stained with Giemsa solution (5%) for 5 min and made permanent with EUKITT (Panreac). Finally, the positives hybridization signals were mapped in the polytene chromosome of the sequenced strain Gd-H4-1 and recorded with a Moticam Package (Moticam3 with 3.0 MP and Motic Images Plus 2.0). Their physical locations were determined according to current photomap of the species (Schaeffer *et al.* 2008) and the chromosomal analysis of this strain by Rohde and Valente (2012).

Data availability

All cytological materials are available for viewing at *Drosophila* Laboratory of Universidade Federal do Rio Grande do Sul, Brazil. More detailed data are available upon request.

RESULTS AND DISCUSSION

With a stringency of approximately 77% (Schwarzacher and Heslop-Harrison 2000) and considering the use of homologous probes, each gene probe mapped in this study showed a single hybridization signal in the polytene chromosomes of *D. willistoni*. Remarkably, all probes designed for scaffolds previously assigned to the III chromosome arm hybridized to the IIR chromosome arm. In turn, all probes designed for scaffolds previously assigned to the IIR chromosome arm hybridized to the III chromosome arm. In both arms, the scaffolds are currently ordered from centromere (scaffold 1) to telomere (scaffold 8) (Schaeffer *et al.* 2008). In the III arm, our results indicate that the same order of scaffolds from

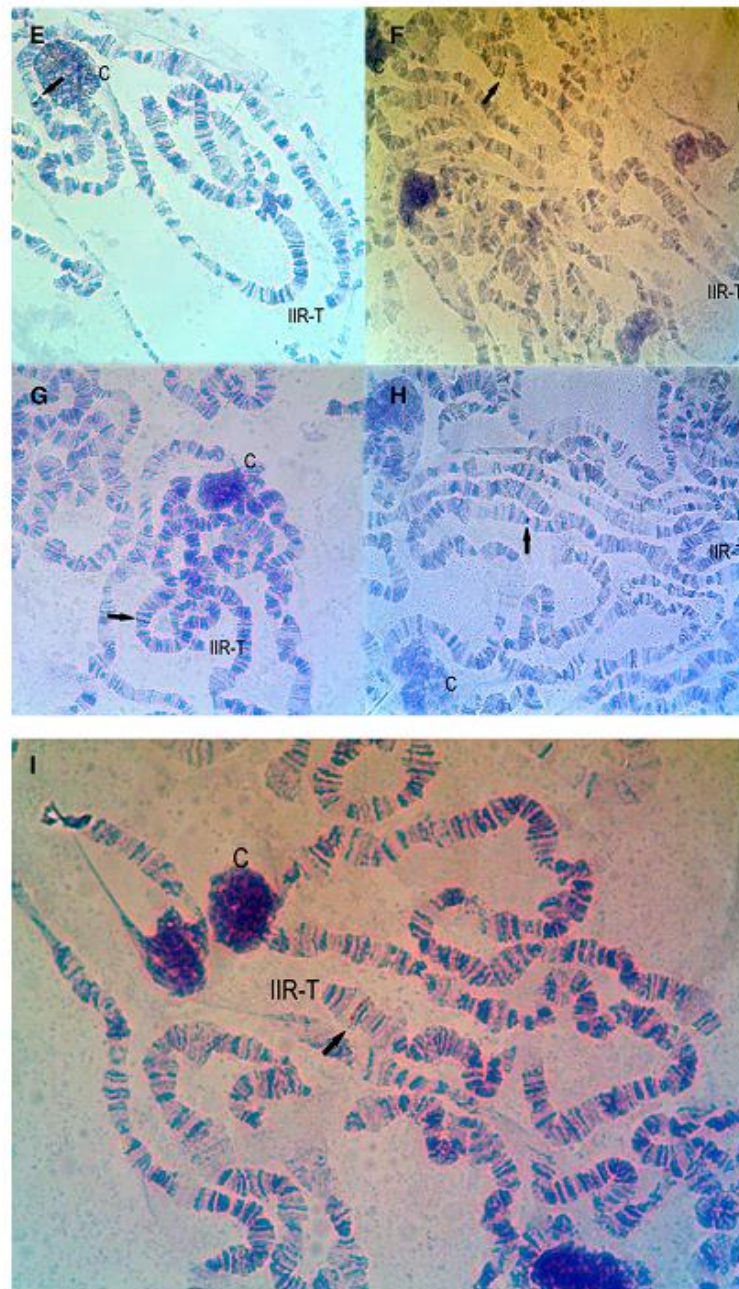


Figure 2 Continued.

centromere to telomere holds (Table 1 and Figure 1). In contrast, in the IIR arm, our results show that the order is reversed, with scaffold 8 located near the centromere and scaffold 1 located near the telomere (Table 1 and Figure 2). Our analysis corroborated these orientations in the two cases where two gene probes were mapped for the same scaffold (scaffold 4577 in the III arm and scaffold 4558 in the IIR arm; see Table 1).

In the case of the IIR arm, Schaeffer *et al.* (2008) inferred that scaffold 4558 has a negative orientation. According to their study, the

cytological position of the *GK22138* gene would appear before that of the other hybridized gene, *GK22144*. As seen in Figure 2, B and C and in Table 1, our hybridizations resulted in the *GK22144* gene being positioned before the *GK22138* gene. A comparison of our results with the previous observations used for anchoring scaffolds (Schaeffer *et al.* 2008) seems appropriate. One caveat should be pointed out from the beginning: the variability of gene arrangements within *D. willistoni* makes linkage mapping an arduous task. Mutants have been often

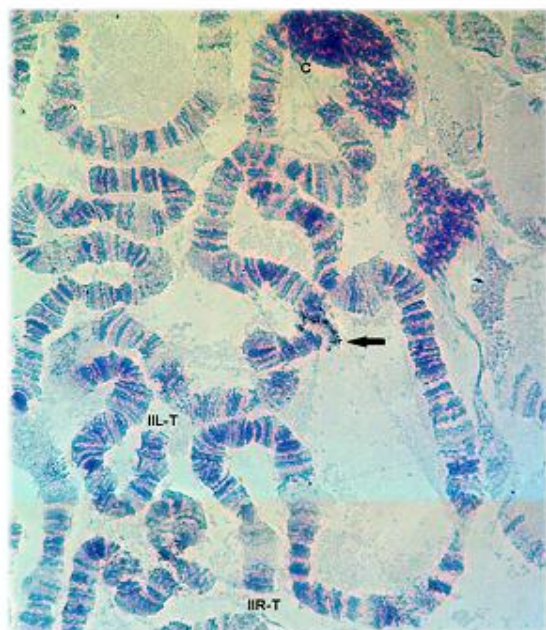


Figure 3 *In situ* hybridization of the gene probe of *Adh* (scaffold 4577) in *Drosophila willistoni*. In the present study, this gene hybridized to section 44A in chromosome arm IIL, in contrast to what was reported in previous studies (Lakovaara and Saura 1972; Rohde et al. 1995). The image shows chromosome II with the two arms IIL and IIR intact for analysis. The black arrow indicates the site of the single hybridization signal. IIL-T, IIL arm telomere; IIR-T, IIR arm telomere; C, chromocenter.

isolated or generated from different strains and crossings to map those mutants may involve inversions hindering or reducing recombination. Additionally, chromosomal differences between populations may lead to different results when using different

strains. Spassky and Dozhansky (1950) mapped six mutants to chromosome II of *D. willistoni* (*bw*, *abb*, *px*, *hk*, *pw*, *Sv*, and *S*). Most of these mutants were obtained from a strain from Belem (Pará, Brazil), but some (e.g., *bw* and *hk*) were generated in X-rayed cultures from strains of a different origin. The eye-color mutant *brown* (*bw*) was located at one end of the chromosome and was considered the starting point of the genetic map. However, Spassky and Dozhansky (1950) gave no indication as to the correspondence between chromosome II arms and Muller elements. An addition to this map was made later by Lakovaara and Saura (1972), who determined the linkage map location of four chromosome II enzyme loci (*Est-5*, *Gpdh*, *Mdh-2*, and *Adh*). Most of the *D. willistoni* stocks used by them originated in the Leeward Islands (West Indies). They were the first to report that *Est-5*, *Gpdh*, *Mdh-2*, and *Adh* are located in the right arm of *D. willistoni* chromosome II and that this arm corresponds to *D. melanogaster* arm 2L (Muller element B).

To anchor the scaffolds to *D. willistoni* chromosome II, Schaeffer et al. (2008) used linkage information from five genetic markers (*bw*, *px*, *hk*, *Gpdh*, and *Adh*) and ISH information from two other markers (*Adh* and *Cf*). The five genetic markers were chosen to maximize reliability but were taken from two different studies. The locations of the five markers in the combined linkage map are as follows: *bw* (0 cM) – *px* (28 cM) – *hk* (31 cM) – *Gpdh* (59 cM) – *Adh* (66 cM). The first two markers were assigned to chromosome arm IIL, whereas the latter three were assigned to chromosome arm IIR. However, in addition to the fact that the strains used by Lakovaara and Saura (1972) were different from those used by Spassky and Dobzhansky (1950), they were different from the *D. willistoni* sequenced strain from Guadeloupe Island.

We have hybridized PCR probes for the three genes considered by Schaeffer et al. (2008): *brown* (*bw*), *plexus* (*px*), and *Alcohol dehydrogenase* (*Adh*). In *D. willistoni*, the orthologous genes for *D. melanogaster* *bw* and *px* are named *Dwi*\GK23049 and *Dwi*\GK22138, respectively, in the FlyBase database (St. Pierre et al. 2014). Our ISH analysis mapped *Dwi*\GK23049 and *Dwi*\GK22138 to sections 56C

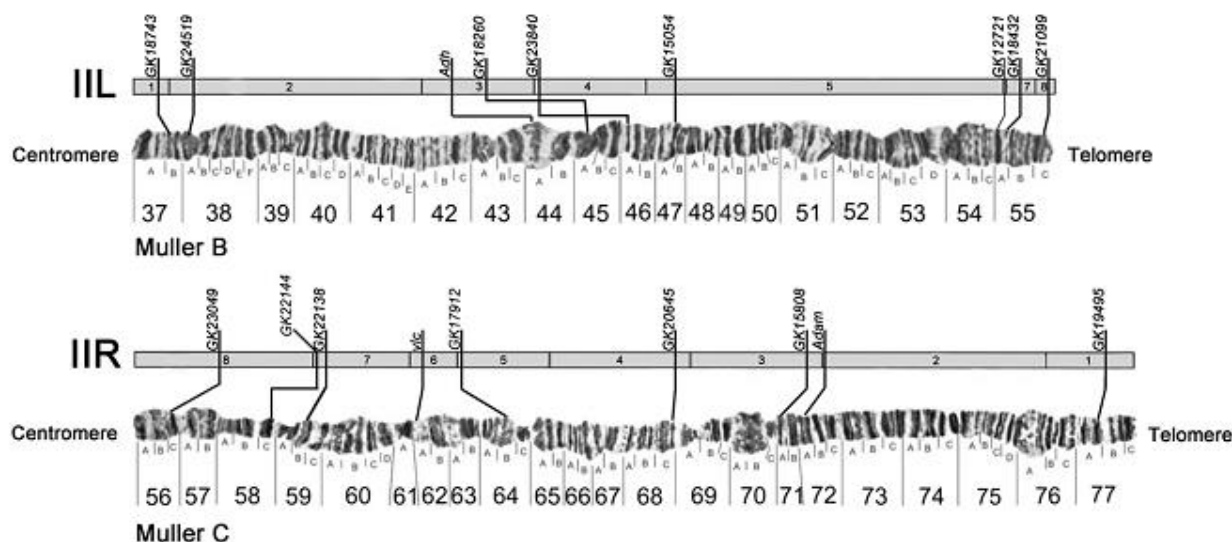


Figure 4 Reassignments of the sequence scaffolds and gene markers in chromosome II of *Drosophila willistoni* (Schaeffer et al. 2008) according to this study (Table 1). Numbers 1–8 in chromosome arm IIL correspond to scaffolds 4884, 4585, 4577, 4516, 4521, 4752, 4945, and 4851 in Table 1. Numbers 1–8 in chromosome arm IIR correspond to scaffolds 4512, 4513, 4514, 4510, 4822, 4382, 4558, and 4954 in Table 1. The position of the genes on the scaffolds matches closest position possible in the scale of genome.

and 59B of the *D. willistoni* chromosome IIR arm (Table 1, Figure 2, A and C). Thus, *bw* and *px* are located in the sequenced *D. willistoni* strain in arm IIR near the centromere and not in the tip of arm III as assumed by Schaeffer *et al.* (2008). Because these two genes also served as a basis for the orientations of the scaffolds, these findings reinforce the idea that scaffolds of IIR arm are in the opposite orientation to the one first proposed by Schaeffer *et al.* (2008).

Furthermore, we mapped the *Adh* gene unequivocally to section 44A in arm III (Figure 3), far away from the positions in arm IIR suggested by both the linkage map and the physical localization. Because this was one of the few genes physically localized before *D. willistoni* genome sequencing, it played a key role in the anchoring of scaffolds to *D. willistoni* chromosomes and thus deserves further comment. As stated previously, *Adh* was first assigned to IIR arm by Lakovaara and Saura (1972) by linkage mapping. Rohde *et al.* (1995) in turn mapped the *Adh* gene using ISH to the IIR arm in seven species of the *willistoni* species group, including *D. willistoni*. This analysis was done with a cloned segment of DNA from *D. melanogaster*, probe SAC-PAT (Moses 1986). Because this probe presented multiple hybridization signals on the chromosomes of *D. willistoni*, the inference of the main physical location was made by quantification. Hence, of 28 chromosome IIR arms analyzed in that study for this species, ~18% indicated hybridization of the *Adh* gene probe consistently in section 67 of this chromosomal arm. Given the previous observation by Lakovaara and Saura (1972), these findings were interpreted as further evidence that the III and IIR arms of *D. willistoni* would be equivalent to the 2R and 2L arms of *D. melanogaster*, respectively (V. L. S. Valente, personal communication). Given the importance of this gene for scaffold assignment in *D. willistoni*, we must emphasize that our technique with a specific gene probe, high stringency, and the presence of a single positive hybridization signal in the chromosomes resulted in a reliable determination of its physical location (Figure 3).

Our results change the assignment of *D. willistoni* chromosome II arms to Muller elements, such that the III chromosomal arm is homologous to *D. melanogaster* 2L (Muller element B) and the IIR chromosomal arm is homologous to *D. melanogaster* 2R (Muller element C) (Figure 4). This result is contrary to what is traditionally accepted. The paucity of genetic markers and the ambiguous interpretation from the classic studies with crossings of those highly polymorphic species have certainly been the determinants for maintaining this mistake for so long.

In addition to aforementioned, the gene markers of *D. willistoni* were established in different populations of its wide overall geographical distribution. Because of the high degree of paracentric inversions present in this species, the order of the genetic and physical markers already established in many cases differs from that occurring in the Gd-H4-1 lineage (Schaeffer *et al.* 2008), which emphasizes the importance of establishing physical markers in the sequenced lineage and correctly aligning their genomic and cytological maps.

In view of this, to briefly examine the assignment of scaffolds to the other chromosomes, we performed additional ISHs of three genes of the X chromosome (one probe in the XL arm and two in the XR arm) and one gene on chromosome III of *D. willistoni*.

In chromosomal arm XL, a probe was designed for the gene *Dwil* \GK16707. This gene is orthologous to *Dmelunc* in *D. melanogaster* and was chosen because it is located in the most telomeric scaffold (4963) of this arm in *D. willistoni*. In this study, its physical position was in section 1C (Figure S1), the most centromeric section in the XL arm. This indicates a disagreement with the position of this gene established in the initial assembly of the scaffold. Note that the *Dmelunc* gene in *D. melanogaster* maps in the region 19F1 of the X chromosome, also next to the centromere.

As reported by Schaeffer *et al.* (2008), scaffold 4822 of the IIR arm (Table 1 and Table S1) had an assembly error, and a small part of it exhibited homology with Muller element D and thus belonging to chromosomal arm XR. We designed a probe for gene *Dwil* \GK17758 that is located in this small part of scaffold 4822; that probe hybridized in section 27C of arm XR as expected (Table S1 and Figure S2). The other gene physically mapped in this arm was *Dwil* \GK16749. This gene is located in the most telomeric scaffold of this arm (4511) and was mapped to section 34B (Figure S3). Thus, our results for the chromosome XR arm are fully congruent with the scaffold assignments that Schaeffer *et al.* (2008) established on the basis of linkage information for six genetic markers and ISH information for another five genes (*E74*, *E75*, *Sod*, *Hsp83*, and *Hsp27*), a relatively high number of markers in comparison to chromosome arm XL and chromosome II.

A dot chromosome is absent in *D. willistoni*, and chromosome III was hypothesized by Sturtevant and Novitski (1941) to have originated from the fusion of Muller elements E and F. Years later, Papacit and Juan (1998) mapped three genes in *D. willistoni* by ISH, *cubitus interruptus* (*ci*), *ankyrin* (*Ank*), and *eyeless* (*ey*), which are located in the *D. melanogaster* dot chromosome (Muller element F). They located the three genes in the centromeric region of *D. willistoni* chromosome III, corroborating with molecular techniques, the fusion of Muller elements E and F. This event has already been confirmed for all *willistoni* subgroup species (Powell *et al.* 2011; Pita *et al.* 2014) with a timing of occurrence estimated at 15 Mya. In view of the good determination of the physical markers for the base of this chromosome we carried out an ISH with the gene *Dwil* \GK22422 located in scaffold 4921 (Table S1), the most telomeric region. This gene was mapped to section 99D (Figure S4) confirming the localization of this gene in this scaffold.

In general, the physical markers used in this study for chromosomes X and III of *D. willistoni* suggest an assignment of scaffolds in these chromosomes without major relocations of genomic content compared with those observed in the case of the Muller elements B and C. However, the sequenced Gd-H4-1 strain still needs better resolution for these chromosomes by establishing physical markers from ISH as presented here for chromosome II.

This study contributes to a more accurate characterization of the genome of *D. willistoni*, providing a better basis for further genetic studies associated with it. As *D. willistoni* is an intriguing model for chromosomal and evolutionary studies, in particular, the results described here will help in analyses of interspecific chromosome evolution and genome evolution (as in Ruiz *et al.* 1997; Ranz *et al.* 2001; Papacit *et al.* 2006 with other species) and also in the characterization of chromosomal inversion breakpoints (as in Delprat *et al.* 2009).

ACKNOWLEDGMENTS

We thank Dr. Victor Hugo Valiati (Universidade do Vale do Rio dos Sinos) for all the help with the images in Brazil. We also thank MSc Nuria Rius and anonymous reviewers for useful comments on a previous version of the manuscript. This work was supported by a grant (BFU2011-30476) to A.R. from the Ministerio de Ciencia e Innovación, Spain. We gratefully acknowledge grant and fellowships fundings from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES process PDSE 18467-12-1 to C.G.), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq to C.G. and V.L.S.V.) and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS process 10/0028-7 to V.L.S.V.) from Brazil.

LITERATURE CITED

- Ayala, F. J., J. R. Powell, M. L. Tracey, C. A. Mourão, and S. Pérez-Salas, 1972 Enzyme variability in the *Drosophila willistoni* group. IV. Genic variation in natural populations of *Drosophila willistoni*. *Genetics* 70: 113–139.
- Bächli, G., 2015 Taxodros, the database on taxonomy of *Drosophilidae*. Available at: http://www.taxodros.uzh.ch/search/prt_rawfile.php?prt=SPECIES-LIST_GR_SR. Accessed October 13, 2015.
- Bhutkar, A., S. W. Schaeffer, S. M. Russo, M. Xu, T. F. Smith *et al.*, 2008 Chromosomal rearrangement inferred from comparisons of 12 *Drosophila* genomes. *Genetics* 179: 1657–1680.
- Biémont, C., L. Monti-Dedieu, and F. Lemeunier, 2004 Detection of transposable elements in *Drosophila* salivary gland polytene chromosomes by *in situ* hybridization. *Methods Mol. Biol.* 260: 21–28.
- Delprat, A., B. Negre, M. Puig, and A. Ruiz, 2009 The transposon *Galileo* generates natural chromosomal inversions in *Drosophila* by ectopic recombination. *PLoS One* 4: e7883.
- Dobzhansky, T., 1950 The chromosomes of *Drosophila willistoni*. *J. Hered.* 41: 156–158.
- Dobzhansky, T., and C. Pavan, 1943 Studies on Brazilian species of *Drosophila*. *Bol. Facul. Fil. Cien e Letr. Univ. S. Paulo* 36: 7–72.
- Dobzhansky, T., and J. R. Powell, 1975 The *willistoni* group of sibling species of *Drosophila*, pp. 589–622 in *Handbook of Genetics*, Vol. 3, edited by King, R. C. Plenum Press, New York, NY.
- Drosophila* 12 Genomes Consortium, 2007 Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450: 203–218.
- Ehrman, L., and J. R. Powell, 1982 The *Drosophila willistoni* species group, pp. 193–225 in *The Genetics and Biology of Drosophila*, Vol. 3b, edited by Ashburner, M., H. L. Carson, and J. N. Thompson, Jr. Academic Press, New York, NY.
- Krimbas, C. B., and J. R. Powell, 1992 Chromosomal polymorphism in natural and experimental populations, pp. 2–52 in *Drosophila Inversion Polymorphism*, edited by Krimbas, C. B., and J. R. Powell. CRC Press, Boca Raton, FL.
- Lakovaara, S., and A. Saura, 1972 Location of enzyme loci in chromosomes of *Drosophila willistoni*. *Experientia* 28: 355–356.
- Markow, T. A., and P. M. O'Grady, 2007 *Drosophila* biology in the genomic age. *Genetics* 177: 1269–1276.
- Moses, K., 1986 Interspecific expression of the alcohol dehydrogenase gene in *Drosophila*. Ph.D. Thesis, University of Cambridge.
- Muller, H. J., 1940 Bearings of the *Drosophila* work on systematics, pp. 185–268 in *The New Systematics*, edited by Huxley, J., Clarendon Press, Oxford.
- Papaceit, M., and E. Juan, 1998 Fate of dot chromosome genes in *Drosophila willistoni* and *Scaptodrosophila lebanonensis* determined by *in situ* hybridization. *Chromosome Res.* 6: 49–54.
- Papaceit, M., M. Aguadé, and C. Segarra, 2006 Chromosomal evolution of elements B and C in the *Sophophora* subgenus of *Drosophila*: evolutionary rate and polymorphism. *Evolution* 60: 768–781.
- Pita, S., Y. Panzera, V. L. S. Valente, Z. G. S. de Melo, C. Garcia *et al.*, 2014 Cytogenetic mapping of the Muller F element genes in *Drosophila willistoni* group. *Genetica* 142: 397–403.
- Powell, J. R., K. Dion, M. Papaceit, M. Aguadé, S. Vicario *et al.*, 2011 Nonrecombining genes in a recombination environment: the *Drosophila* “dot” chromosome. *Mol. Biol. Evol.* 28: 825–833.
- Ranz, J. M., F. Casals, and A. Ruiz, 2001 How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res.* 11: 230–239.
- Ruiz, A., J. M. Ranz, M. Caceres, and C. Segarra, 1997 Chromosomal evolution and comparative gene mapping in the *Drosophila repleta* species group. *Braz. J. Genet.* 20: 553–565.
- Rohde, C., and V. L. S. Valente, 2012 Three decades of studies on chromosomal polymorphism of *Drosophila willistoni* and description of fifty different rearrangements. *Genet. Mol. Biol.* 35: 966–979.
- Rohde, C., E. Abdelhay, H. Pinto Júnior, A. Schrank, and V. L. S. Valente, 1995 Analysis and *in situ* mapping of the *Adh* locus in species of the *willistoni* group of *Drosophila*. *Cytobios* 81: 37–47.
- Schaeffer, S. W., A. Bhutkar, B. F. McAllister, M. Matsuda, L. M. Matzkin *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.
- Schwarzacher, T., and P. Heslop-Harrison, 2000 *Practical In situ Hybridization*. BIOS Scientific Publishers, Oxford.
- Sorsa, V., 1988 Polytene chromosome and Evolution of *Drosophila*, pp. 139–153 in *Chromosome Maps of Drosophila*, Vol. 1, edited by Sorsa, V. CRC Press, Boca Raton, FL.
- Spassky, B., and T. Dobzhansky., 1950 Comparative genetics of *Drosophila willistoni*. *Heredity* 4: 201–215.
- Sperlich, D., and P. Pfriem, 1986 Chromosomal polymorphism in natural and experimental populations, pp. 257–309 in *The Genetics and Biology of Drosophila*, Vol. 3e, edited by Ashburner, M., H. L. Carson, and J. N. Thomson, Jr. Academic Press, New York, NY.
- St. Pierre, S. E., L. Ponting, R. Stefancsik, P. McQuilton, and the FlyBase Consortium, 2014 FlyBase 102—advanced approaches to interrogating FlyBase. *Nucl. Acids Res.* 42(Database issue): D780–788.
- Sturtevant, A. H., and E. Novitski, 1941 The homologies of the chromosome elements in the genus *Drosophila*. *Genetics* 26: 517–541.
- Val, F. C., C. R. Vilela, and M. D. Marques, 1981 *Drosophilidae* of the Neotropical Region, pp. 123–168 in *The Genetics and Biology of Drosophila*, Vol. 3a, edited by Ashburner, M., H. L. Carson, and J. N. Thompson, Jr. Academic Press, New York, NY.

Communicating editor: S. Celniker

REASSIGNMENT OF *DROSOPHILA WILLISTONI* GENOME SCAFFOLDS TO CHROMOSOME II ARMS

Carolina Garcia^{*§}, Alejandra Delprat[§], Alfredo Ruiz[§], Vera L S Valente^{*}

^{*} Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Brazil 15053

[§] Departament de Genètica i de Microbiologia, Facultat de Biociències, Universitat Autònoma de Barcelona, Barcelona, Spain 08193

Corresponding autor: Vera L S Valente, Departamento de Genética, Av. Bento Gonçalves, 9500, 43323M/210, Postal

Code: 91501-970 - Porto Alegre, RS, Brasil, Phone (55-51) 3308-6713. E-mail: 00004887@ufrgs.br

Table S1 Gene markers used for chromosomes X and III of *Drosophila willistoni*. The scaffold number corresponds to the last four numbers of the scaffolds, which all start with scf2_110000000. The scaffold number and scaffold position of genes corresponds to the material available in the FlyBase database (St. Pierre *et al.* 2014).

<i>D. willistoni</i> Gene	<i>D. melanogaster</i> Ortholog Gene	Scaffold number	Scaffold position of gene	Cytological position/chromosome	Primers F and R (5'-3')
<i>Dwil</i> \GK16707	<i>Dmel</i> \unc	4963	432,088..435,746	1C/XL arm	ACTCAGTCTTCGACGGAAGC AGTTGTATCGGATTCTACCA
<i>Dwil</i> \GK17758	<i>Dmel</i> \ida	4822	3,033,141..3,041,719	27C/XR arm	GCTGCATTAGATCCTCATAG GGCAGCCAACAGTCCATACA
<i>Dwil</i> \GK16749	<i>Dmel</i> \CG13313	4511	7,841,949..7,843,999	34B/XR arm	GCTATCAGTCACCGTGTAGA GGCAGTTGCTCCACCATCAC
<i>Dwil</i> \GK22422	<i>Dmel</i> \CG31204	4921	3,260,674..3,262,239	99D/chromosome III	GAGTCAATGCGTCCATACCA GGATAATCCTCACGAGACTG

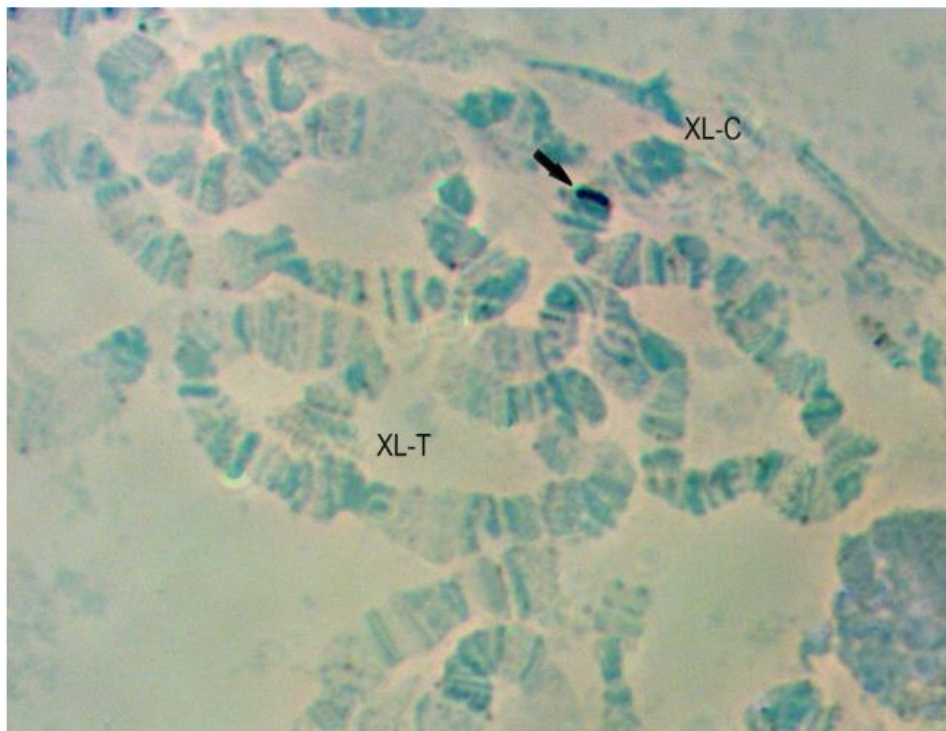


FIGURE S1 *In situ* hybridization of the *Dwi\GK16707* gene (scaffold 4963) to the *D. willistoni* chromosome XL arm. The black arrow indicates the hybridization signal site in section 1C. **XL-T**: XL arm telomere. **XL-C**: XL arm centromere.

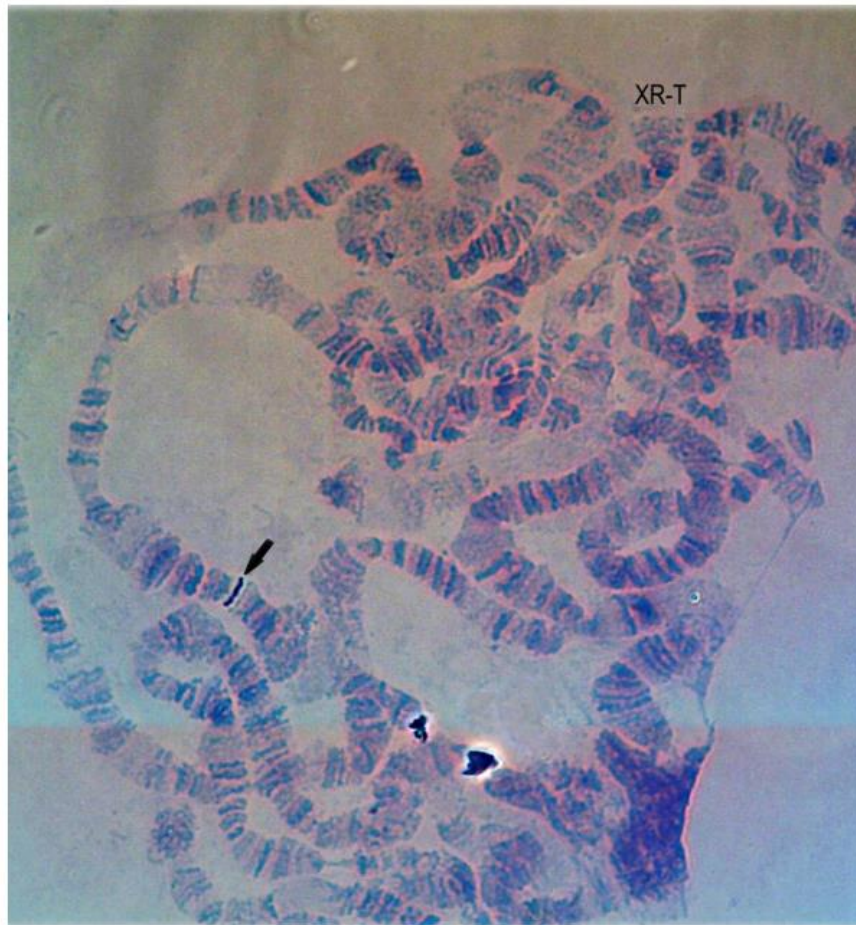


FIGURE S2 *In situ* hybridization of the *Dwil\GK17758* gene to the *D. willistoni* chromosome XR arm. This gene is located in the chimeric scaffold 4822, which was split into two Muller elements: a large portion in the IIR arm (see Table 1) and a smaller portion, containing the *Dwil\GK17758* gene, in the XR arm. The black arrow indicate hybridization signal site in section 27C. XR-T: XR arm telomere.

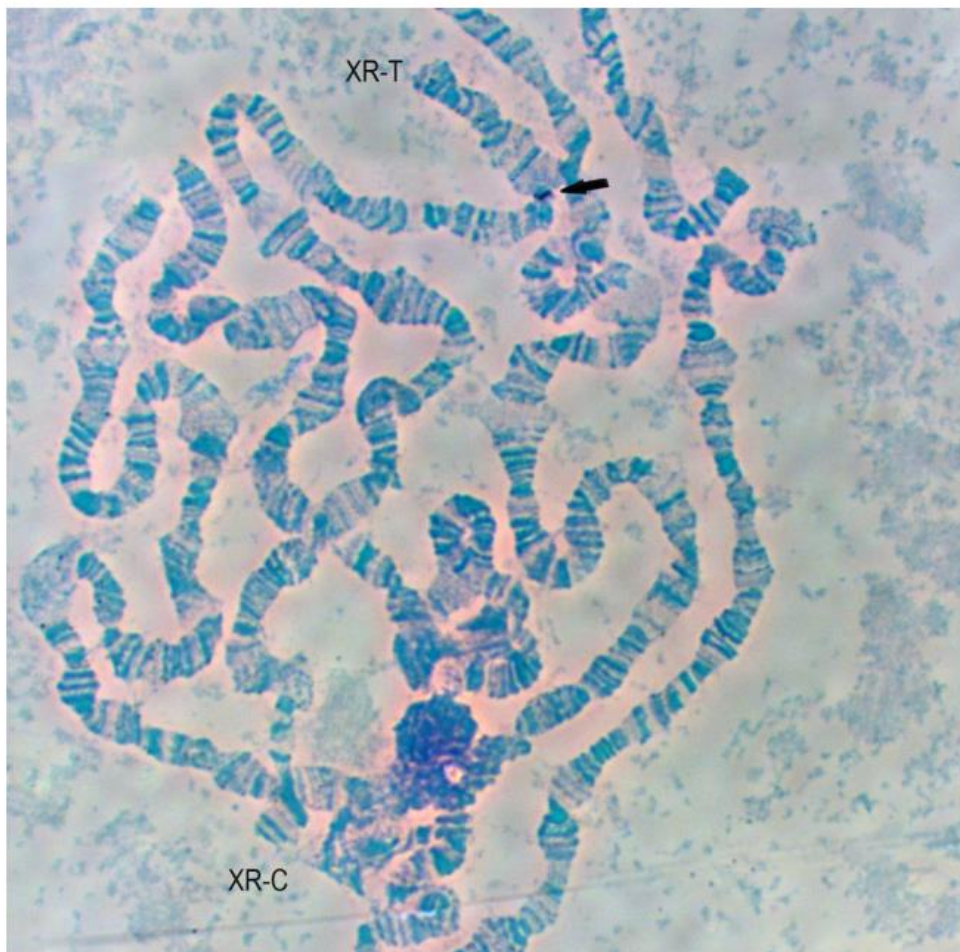


FIGURE S3 *In situ* hybridization of the *Dwil\GK16749* gene (scaffold 4511). This scaffold is the most telomeric in this chromosome arm. The black arrow indicates the hybridization signal in section 34B. **XR-T**: XR arm telomere. **XR-C**: XR arm centromere.

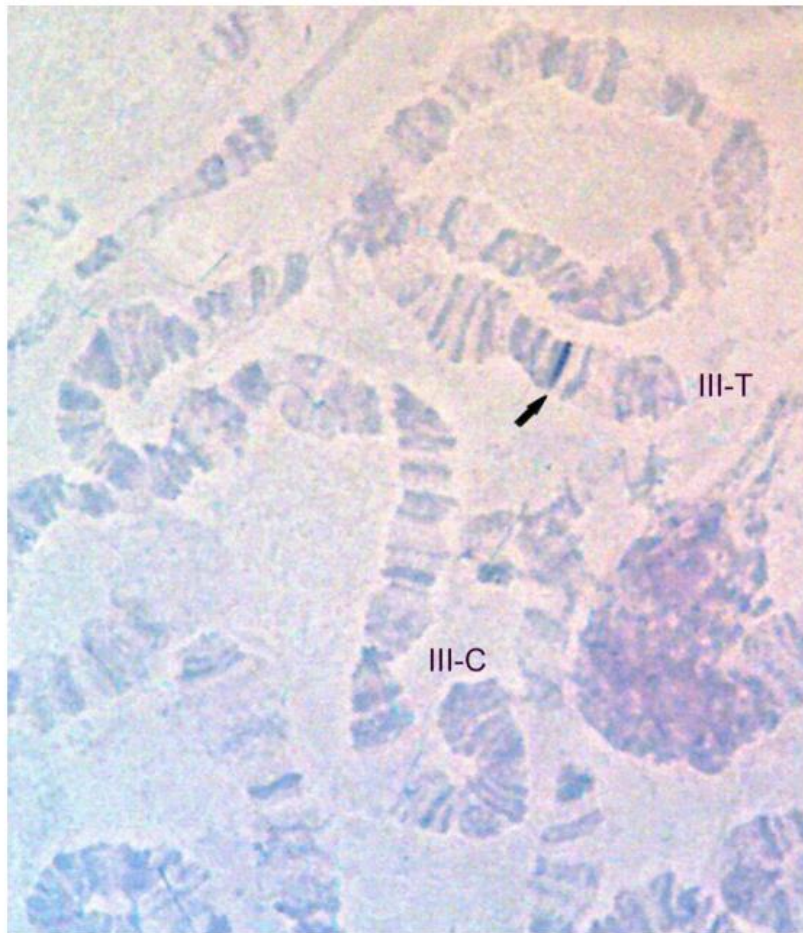


FIGURE S4 *In situ* hybridization of the *Dwil\GK22422* gene (scaffold 4921) to chromosome III. The black arrow indicates the hybridization signal in section 94D of this chromosome. This gene is located in the most telomeric scaffold (4921) and its cytological localization confirms its position in the scaffold. **III-T**: chromosome III telomere. **III-C**: chromosome III centromere.

REFERÊNCIAS BIBLIOGRÁFICAS DA TESE

Adam MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, *et al.* (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287:2185-2195.

Anderson CL, Carew EA, Powell JR (1993) Evolution of the Adh locus in the *Drosophila willistoni* group: the loss of an intron, and shift in codon usage. *Mol Biol Evol.* 10(3):605-618.

Anderson WW, Arnold J, Baldwin DG, Beckenbach AT, Brown CJ, Bryant SH, Coyne JA, Harshman LG, Heed WB and Jeffery DE (1991) Four decades of inversion polymorphism in *Drosophila pseudoobscura*. *PNAS* 88 (22): 10367-10371.

Andolfatto P, Wall JD, Kreitman M (1999) Unusual haplotype structure at the proximal breakpoint of *In(2L)t* in a natural population of *Drosophila melanogaster*. *Genetics* 153: 1297–1311.

Andolfatto P, Depaulis F, Navarro A (2001) Inversion polymorphisms and nucleotide variability in *Drosophila*. *Genet Res* 77(1):1-8.

Aulard S, David JR and Lemeunier F (2002) Chromosomal inversion polymorphism in Afrotropical populations of *Drosophila melanogaster*. *Genet Res* 79(1):49-63.

Ayala FJ, Powell JR, Dobzhansky T (1971) Polymorphism in continental and island populations of *Drosophila willistoni*. *PNAS USA* 68: 2480-2483.

Ayala D, Guerrero RF, Kirkpatrick M (2012) Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito. *Evolution* 67(4): 946-958.

Bächli G (2015) Taxodros, the database on taxonomy of Drosophilidae. Disponível em: <http://www.taxodros.unizh.ch/>. Acessado em agosto de 2015.

Bao W, Jurka, J (2013) DNA transposons from the Pacific oyster genome. *Rebase Reports* 13(1): 660-660.

Bao W, Jurka, J (2014) Non-LTR retrotransposons from zebrafish. *Rebase Reports* 14(6):1829-1829.

Bao W (2015) DNA transposons from *Hydra magnipapillata*. *Rebase Reports* 15(12): 3629-3629.

Battaglia, B, Birch LC (1956) Crossing-over in *Drosophila willistoni*. *Nature* 178: 1005.

Betrán E, Quezada-Díaz JE, Ruiz A, Santos M, Fontdevila A (1995) The evolutionary history of *Drosophila buzzatii*. XXXII. Linkage disequilibrium between allozymes and chromosome inversions in two colonizing populations. *Heredity* 74 (Pt 2): 188-199.

Besansky NJ, Krzywinski J, Lehmann T, Simard F, Kern M, *et al.* (2003) Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: evidence from multilocus DNA sequence variation. *Proc Natl Acad Sci U S A.* 16; 100(19):10818-10823.

Biémont C, Monti-Dedieu L, Lemeunier F (2004) Detection of transposable elements in *Drosophila* salivary gland polytene chromosomes by *in situ* hybridization. *Methods Mol. Biol.* 260: 21-28.

Bonorino CB, Couto e Silva T, Abdelhay E, Valente VL (1993) Heat shock genes in the *willistoni* group of *Drosophila*: induced puffs and proteins. *Cytobios.* 73(292): 49-64.

Bridges CB (1935) Salivary chromosome maps with a key to the banding of the chromosomes of *Drosophila melanogaster*. *J Hered* 26 (2): 60-64.

Brcic D, Prevosti A, Budnik M, Monclús M, Ocaña J (1981) Colonization of *Drosophila subobscura* in Chile I. First population and cytogenetic studies. *Genetica*, 56(1): 3-9.

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

Bhutkar A, Schaeffer SW, Russo SM, *et al.* (2008) Chromosomal rearrangement inferred from comparisons of 12 *Drosophila* genomes. *Genetics* 179(3):1657- 1680.

Cáceres M, Barbadilla A, Ruiz A (1997) Inversion Length and Breakpoint Distribution in the *Drosophila buzzatii* Species Complex: Is Inversion Length a Selected Trait? *Evolution* 51(4): 1149-1155.

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.

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

Cáceres M; National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program, Sullivan RT, Thomas JW (2007) A recurrent inversion on the eutherian X chromosome. *Proc Natl Acad Sci U S A.* 20;104(47):18571-18576.

Casals F, Cáceres M, Ruiz A (2003) The foldback-like transposon *Galileo* is involved in the generation of two different natural chromosomal inversions of *Drosophila buzzatii*. *Mol Biol Evol* 20: 674–685.

Casals F, Navarro A (2007) Inversions: The chicken or the egg. *Heredity* 99: 479–480.

Calvete O, González J, Betrán E, Ruiz A (2012) Segmental duplication, microinversion, and gene loss associated with a complex inversion breakpoint region in *Drosophila*. *Mol Biol Evol.* 29(7):1875-89.

Carson HL, Kaneshiro KY (1976) *Drosophila* of Hawaii: Systematics and Ecological Genetics. Annual Review of Ecology and Systematics 7: 311-345.

Cavasini R, Carvalho KA, Klaczko LB (2010). Absence of recombination in males of *Drosophila mediopunctata*. Dros Info Service 93: 122-124.

Chapple CE, Guigó R (2008) Relaxation of Selective Constraints Causes Independent Selenoprotein Extinction in Insect Genomes. Plos ONE 3 (8): e2968.
Chovnick A (1973) Gene conversion and transfer of genetic information within the inverted region of inversion heterozygotes. Genetics 75(1): 123-131.

Cirera S, Martin-Campos JM, Segarra C, Aguade M (1995) Molecular characterization of the breakpoints of an inversion fixed between *Drosophila melanogaster* and *D. subobscura*. Genetics 139: 321–326.

Cirulli ET, Noor MA (2007) Localization and characterization of X chromosome inversion breakpoints separating *Drosophila mojavensis* and *Drosophila arizonae*. J Hered. doi:10.1093/jhered/esl065

Corbett-Detig RB, Cardeno C, Langley CH (2012) Sequence-based detection and breakpoint assembly of polymorphic inversions. Genetics 192:131–137.

Corbett-Detig RB, Hartl DL (2012) Population genomics of inversion polymorphisms in *Drosophila melanogaster*. PLoS Genet 8(12): e1003056.

Da Cunha AB, Burla H and Dobzhansky T (1950) Adaptive Chromosomal Polymorphism in *Drosophila willistoni*. Evolution 4 (3): 212-235.

Da Cunha AB and Dobzhansky T (1954) A further study of chromosomal polymorphism in *Drosophila willistoni* in its relation to environment. Evolution, 8: 119-134.

Darwin C (1859) On the Origin of Species by means of natural selection, or the preservation of favoured races in the struggle for life. Jonh Murray, Albermale Street, London, 502p.

De Souza HML, Da Cunha AB, Dos Santos EP (1970) Adaptive Polymorphism of Behavior Evolved in Laboratory Populations of *Drosophila willistoni*. The American Naturalist 104: 175-189.

Delprat A, Negre B, Puig M, Ruiz A (2009) The transposon Galileo generates natural chromosomal inversions in *Drosophila* by ectopic recombination. PLoS One 4:e7883.

Deonier RC, Tavaré S, Waterman M. (2005). Computational genome analysis: an introduction. Springer-Verlag Science & Business Media. ISBN 0-387-98785-1.

Deprá M, Ludwig A, Valente VLS, Loreto ELS (2012) Mar, a MITE family of hAT transposons in *Drosophila*. Mob DNA 3: 13. doi: 10.1186/1759-8753-3-13.

Diniz NM, Sene FM (2004) Chromosomal phylogeny of the *Drosophila fasciola* species subgroup revisited (Diptera, Drosophilidae). Genetics and Molecular Biology 27 (4): 561-556.

Dobzhansky T (1937, 1941, 1951) *Genetics and the Origin of Species*. Ed. 1, 2, 3. Columbia University Press, New York.

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

Dobzhansky T, Epling C (1944) *Contributions to the Genetics, Taxonomy, and Ecology of Drosophila pseudoobscura and Its Relatives*. Carnegie Institution of Washington, Washington, 554 pp.

Dobzhansky T (1949) On some of the problems of population genetics and evolution. *Symposium sui Fattori Ecologici e Genetici della Speciazione Degli Animali*. *La Ricerca Scientifica*, 19(Suppl.), 11.

Dobzhansky T (1950) The chromosomes of *Drosophila willistoni*. *J Heredity* 41: 156-158.

Dobzhansky T (1957) Genetics of natural populations. XXVI. Chromosomal variability in island and continental populations of *Drosophila willistoni* from Central America and the West Indies. *Evolution* 11: 280-293.

Dobzhansky T (1970) *Genetics of the Evolutionary Process*, Columbia University Press, New York.

Dobzhansky T, Pavlovsky O (1955) An extreme case of heterosis in a central american population of *Drosophila tropicalis*. *Proc Natl Acad Sci* 15;41(5): 289-895.

Dobzhansky T, Pavlovsky O (1972) A comparative study of the chromosomes in the incipient species of the *Drosophila paulistorum* complex. *Chromosoma* 13: 196-218.

Drosophila 12 Genomes Consortium (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 480: 203-218.

Eickbush TH (2002) R2 and related site-specific non-long terminal repeat retrotransposons. *In Mobile DNA II*, N. L. Craig, R. Craigie, M. Gellert and A. M. Lambowitz (eds), American Society of Microbiology Press, Washington pp. 813–835.

Engels WR, Preston CR (1984) Formation of chromosome rearrangements by *P* factors in *Drosophila*. *Genetics* 107: 657-678.

Ehrman L, Powell JR (1982) The *Drosophila willistoni* species group. *In: The Genetics and Biology of Drosophila*. Ashburner M, Carson HL and Thompson Jr. JN (eds), Academic Press, New York, vol 3b: 193-225.

Evans AL, Mena PA, McAllister BF (2007) Positive selection near an inversion breakpoint on the neo-X chromosome of *Drosophila americana*. *Genetics* 177(3): 1303-1319.

Evgen'ev MB, Zelentsova H, Shostak N, Kozitsina M, Barskyi V, *et al.* (1997) *Penelope*, a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*. *Proc Natl Acad Sci USA* 94(1):196-201.

Evgen'ev M, Zelentsova H, Mnjoian L, Poluectova H, Kidwell MG (2000) Invasion of *Drosophila virilis* by the Penelope transposable element. *Chromosoma* 109 (5):350-357.

Feder JL, Roethele FB, Filchak K, Niedbalski J, Romero-Severson J (2003) Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. *Genetics* 163: 939–953.

Ferry RM, Lancefield RC, Metz CW (1923) Additional Mutant Characters in *Drosophila willistoni*. *J Hered* 14 (8): 373-384.

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

Fonseca NA, Vieira CP, Schlötterer C, Vieira J (2012) The *DAIBAM MITE* element is involved in the origin of one fixed and two polymorphic *Drosophila virilis* phylad inversions. *Fly (Austin)* 6(2): 71-74.

Fontdevila A, Ruiz A, Ocaña J, Alonso G (1982) Evolutionary History of *Drosophila buzzatii*. II. How Much Has Chromosomal Polymorphism Changed in Colonization? *Evolution* Vol. 36, N° 4, pp. 843-851.

Fragata I, Lopes-Cunha M, Bárbaro M, Kellen B, Lima M, Santos MA, *et al.* (2014) How much can history constrain adaptive evolution? A real-time evolutionary approach of inversion polymorphisms in *Drosophila subobscura*. *J Evol Biol.* 27(12): 2727-2738.

França ZM, Da Cunha AB (1968) Crossing-over between heterozygous inversions and its relation with polymorphism in *Drosophila willistoni*. *Revista Brasileira de Biologia* 28: 495–497.

França ZM, Da Cunha AB, Garrido MC (1968) Recombination in *Drosophila willistoni*. *Heredity* 23: 199–204.

Fyrberg EA, Kindle KL, Davidson N, Sodja A (1980) The actin genes of *Drosophila*: a dispersed multigenefamily. *Cell* 19: 365-378.

Fuller ZL, Haynes GD, Zhu D, Batterton M, Chao H, *et al.* (2014) Evidence for stabilizing selection on codon usage in chromosomal rearrangements of *Drosophila pseudoobscura*. *G3 (Bethesda)* 17; 4(12): 2433-2449.

Garcia CF, Delprat A, Ruiz A, Valente VLS (2015) Reassignment of *Drosophila willistoni* genome scaffolds to chromosome II arms. *G3 (Bethesda)* pii: 4;5(12): 2559-2566. doi: 10.1534/g3.115.021311

Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH., *et al.* (2010) The NCBI BioSystems database. *Nucleic Acids Res.* 38 (Database issue): D492-6, doi: 10.1093/nar/gkp858.

Gershenson SM (1930) Phenomenon of reinversion in the sex-chromosome of *Drosophila melanogaster*. *Reports IV Congress Zool., Kiev (Russian)* p. 7.

Goñi B, Matsuda M, Yamamoto MT, Vilela CR, Tobar YN (2012) Crossing over does occur in males of *Drosophila ananassae* from natural populations. *Genome* 55: 505–511.

Goñi B, Valente VLS (2012) Polytene chromosome analysis in color eye mutants of *Drosophila willistoni* mutants and their hybrids. The H inversion. *Drosophila Information Service*, v. 95, pp. 81-85.

Goldberg DA (1980) Isolation and partial characterization of the *Drosophila* alcohol dehydrogenase gene. *Proc Natl Acad Sci U S A* 77(10): 5794-5798.

González J, Casals F, Ruiz A (2007) Testing chromosomal phylogenies and inversion breakpoint reuse in *Drosophila*. *Genetics* 175(1): 167-77.

Gray YH, Tanaka MM, Sved JA (1996) *P*-element-induced recombination in *Drosophila melanogaster*: hybrid element insertion. *Genetics* 144(4): 1601-10.

Gray YH (2000) It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements. *Trends Genet* 16: 461–468.

Griffiths AJF, Wessler SR, Lewontin RC, Carroll SB (2008) Mutaç o cromossomica I: Mudanç a na Estrutura dos cromossomos. *In*: Griffiths AJF (ed) *Introduç o   Gen tica* 9th edition. Guanabara Koogan, Rio de Janeiro, pp 490-503.

Gr neberg H (1936) A case of complete reversion of a chromosomal rearrangement in *Drosophila melanogaster*. *Nature* 138:508.

Guerra M (1988) *Introduç o   Citogen tica Geral*. Guanabara Koogan, 142p.

Guerra M (2004) *FISH: Conceitos e aplicaç es*. Sociedade Brasileira de Gen tica, Ribeir o Preto.

Guill n Y, Ruiz A (2012) Gene alterations at *Drosophila* inversion breakpoints provide prima facie evidence for natural selection as an explanation for rapid chromosomal evolution. *BMC Genomics* 13:53. doi: 10.1186/1471-2164-13-53.

Hartl DL, Jones EW (2009) *Genetics of Chromosomal Inversions*. *In*: *Genetics: analysis of genes and genomes*. 7th edition. Jones and Bartlett, London, pp 275-277.

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

Henderson SA, Woodruff RC, Thompson JN (1978) Spontaneous Chromosome Breakage at Male Meiosis Associated with Male Recombination in *Drosophila Melanogaster*. *Genetics* 88(1): 93-107.

Hinton CW, Lucchesi JC (1960) A cytogenetic study of crossing over in inversion heterozygotes of *D. melanogaster*. *Genetics* 45: 87-94.

Hoffmann AA, Sgr  CM, Weeks AR (2004) Chromosomal inversion polymorphisms and adaptation. *Trends Ecol Evol.* 19(9):482-8.

Hoffmann AA, Rieseberg LH (2008) Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? *Annu Rev Ecol Evol Syst* 1; 39: 21–42.

Holmgren R, Livak K, Morimoto R, Freund R, Meselson K (1979) Studies of cloned sequences from four *Drosophila* heat shock loci. *Cell* 18: 1359-1370.

Horton IH (1939) A Comparison of the Salivary Gland Chromosomes of *Drosophila melanogaster* and *D. simulans*. *Genetics* 24(2): 234-243.

Jurka J, Klonowski P, Dagman V, Pelton P (1996) CENSOR - a program for identification and elimination of repetitive elements from DNA sequences. *Computers and Chemistry* Vol. 20 (No. 1): 119-122.

Kapitonov VV, Jurka J (2007) "Chapaev - a novel superfamily of DNA transposons." *Rebase Reports* 7(9): 779-779.

Kastritsis CD (1966) A comparative chromosome study in the incipient species of the *Drosophila paulistorum* complex. *Chromosoma* 19(2): 208-22.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, *et al.* (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12): 1647-1649.

Kenig B, Kurbalija Novičić Z, Patenković A, Stamenković-Radak M, Anđelković M (2015) Adaptive Role of Inversion Polymorphism of *Drosophila subobscura* in Lead Stressed Environment. *PLoS One* 23; 10(6): e0131270.

Kidwell MG, Kidwell JF, Sved JA (1977) Hybrid Dysgenesis in *Drosophila melanogaster*: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination. *Genetics* 86(4): 813-833.

Kidwell MG (2005) Transposable elements and the Evolution of Dipteran Genomes. *In: The evolutionary biology of flies*. Yeates DK and Wiegmann BM (eds), Columbia University, New York, pp 145-173.

Kikkawa H (1937) Spontaneous crossing-over in the male of *Drosophila ananassae*. *Zool. Mag.* 49: 159–160.

Kojima KK, Jurka J (2013) Helitrons from *Arabidopsis lyrata*. *Rebase Reports* 13(2): 1033-1033.

Kohany O, Gentles AJ, Hankus L, Jurka J (2006) Annotation, submission and screening of repetitive elements in Rebase: Rebase Submitter and Censor. *BMC Bioinformatics* 25; 7:474.

Krimbas CB, Loukas M (1980) The inversion polymorphism of *Drosophila subobscura*. *In Hecht MK, Steere WC, Wallace B (eds). Evolutionary biology*. Vol.12. Plenum Press, New York, pp. 163-234.

Krimbas CB, Powell JR (1992). Chromosomal Polymorphism in Natural and Experimental Populations *In: Drosophila Inversion Polymorphism*. Krimbas CB and Powell JR (eds), CRC, Florida, USA, pp 2-52.

Lake CM, Hawley SR (2012) The Molecular Control of Meiotic Chromosomal Behavior: Events in Early Meiotic Prophase in *Drosophila* Oocytes. *Annu. Rev. Physiol.* 74: 425–51.

Lakovaara S, Saura A (1972) Location of enzyme loci in chromosomes of *Drosophila willistoni*. *Experientia* 28(3): 355-356.

Lancefield RC, Metz CW (1922) The sex-linked group of mutant characters in *Drosophila willistoni*. *American Naturalist* 56 (644): 211-241.

Lemeunier F, Ashburner M (1984) Relationships within the melanogaster species subgroup of the genus *Drosophila* (Sophophora) IV. The chromosomes of two new species. *Chromosoma* 89 (5): 343-351.

Lemeunier F, Aulard S (1992) Inversion polymorphism in *Drosophila melanogaster*. In Krimbas CB & Powell JR (eds) *Drosophila inversion polymorphism*. CRC Press, Boca Raton, FL, pp. 339-405.

Lim JK (1988) Intrachromosomal rearrangements mediated by *hobo* transposons in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 85(23):9153-9157.

Lim JK, Simmons MJ (1994) Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster*. *Bioessays* 16(4): 269-275.

Loukas M, Kafatos FC (1986) The actin loci in the genus *Drosophila*: establishment of chromosomal homologies among distantly related species by *in situ* hybridization. *Chromosoma* 94: 297-308.

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

Lyon MF (2003) Transmission ratio distortion in mice. *Annu. Rev. Genet* 37: 393–408.

Machado CA, Haselkorn TS, Noor MA (2007) Evaluation of the genomic extent of effects of fixed inversion differences on intraspecific variation and interspecific gene flow in *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 175(3): 1289-306.

Marques EK, Napp M, Winge H, Cordeiro AR (1966) A corn meal, soybean flour, wheat germ medium for *Drosophila*. *Dros Inf Serv* 41:1 87.

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. *PNAS* 105 (8): 2957-2962.

Matzkin LM, Merritt TJ, Zhu CT, Eanes WF (2005) The structure and population genetics of the breakpoints associated with the cosmopolitan chromosomal inversion *In(3R)Payne* in *Drosophila melanogaster*. *Genetics* 170: 1143–1152.

McAllister BF (2002) Chromosomal and allelic variation in *Drosophila americana*: selective maintenance of a chromosomal cline. *Genome*. 45(1):13-21.

McClintock B (1941) The Stability of Broken Ends of Chromosomes in *Zea mays*. *Genetics* 26(2): 234-282.

McClintock B (1946) Maize genetics. *Carnegie Inst of Wash Year Book* 45,176-186.

McClintock B (1947) Cytogenetic studies of maize and *Neurospora*. *Carnegie Inst of Wash Year Book* 46, 146-152.

McGaugh SE, Noor MA (2012) Genomic impacts of chromosomal inversions in parapatric *Drosophila* species. *Philos Trans R Soc Lond B Biol Sci* 5; 367(1587): 422-429.

Morgan TH (1912) Complete Linkage in the Second Chromosome of the Male of *Drosophila*. *New Series, Vol. 36, N^o. 934*: 719-720.

Moses K (1986) Interspecific expression of the alcohol dehydrogenase gene in *Drosophila*. Ph.D. Thesis, University of Cambridge.

Muller HJ (1940) Bearings of the *Drosophila* work on systematics. *In The New Systematics*, edited by Huxley J, Clarendon Press, Oxford, pp.185–268.

Murphy WJ, Larkin DM, Everts-van der Wind A, Bourque G, Tesler G *et al.* (2005) Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. *Science* 309: 613–617.

Nadeau JH, Taylor BA (1984) Lengths of chromosomal segments conserved since divergence of man and mouse. *Proc Natl Acad Sci U S A.* 81(3): 814-818.

Navarro A, Betran E, Barbadilla A, Ruiz A (1997) Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* 146: 695–709.

Navarro A, Ruiz A (1997) On the fertility effects of pericentric inversions. *Genetics* 147(2): 931-933.

Olvera O, Powell JR, de la Rosa ME, Salceda VM, *et al.* (1979) Populations genetics of Mexican *Drosophila*. VI. Cytogenetics aspects of the inversion polymorphism in *Drosophila pseudoobscura*. *Evolution* 33: 381-395.

Ohno S (1973) Ancient linkage groups and frozen accidents. *Nature* 3;244(5414): 259-262.

Orr-Weaver TL (1995) Meiosis in *Drosophila*: Seeing is believing. *Proc. Natl. Acad. Sci. USA* 92: 10443-10449,

Painter TS (1933) A new method for the study of chromosome aberrations and the plotting of chromosome maps. *Science* 78: 585–586.

Papacit M, Juan E (1998) Fate of dot chromosome genes in *Drosophila willistoni* and *Scaptodrosophila lebanonensis* determined by *in situ* hybridization. *Chromosome Res* 6(1): 49-54.

Papaceit M, Aguadé M, Segarra C (2006) Chromosomal evolution of elements B and C in the *Sophophora* subgenus of *Drosophila*: evolutionary rate and polymorphism. *Evolution* 60: 768–781.

Papaceit M, Segarra C, Aguadé M (2012) Structure and population genetics of the breakpoints of a polymorphic inversion in *Drosophila subobscura*. *Evolution* 67: 66–79.

Pavan C (1946) Chromosomal Variation in *Drosophila nebulosa*. *Genetics* 31(6): 546-557.

Pegueroles C, Ordóñez V, Mestres F, Pascual M (2010) Recombination and selection in the maintenance of the adaptive value of inversions. *J Evol Biol* 23(12): 2709-2717.

Pevzner P, Tesler G (2003) Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. *Proc Natl Acad Sci U S A*. 24; 100(13): 7672-7677.

Pita S, Panzera Y, Valente VLS, de Melo GCS, C. Garcia et al. (2014) Cytogenetic mapping of the Muller F element genes in *Drosophila willistoni* group. *Genetica* 142: 397–403.

Prada CF, Delprat A, Ruiz A (2011) Testing chromosomal phylogenies and inversion breakpoint reuse in *Drosophila*. The *martensis* cluster revisited. *Chromosome Res.* 19(2): 251-265.

Prazeres da Costa O, González J, Ruiz A (2009) Cloning and sequencing of the breakpoint regions of inversion 5g fixed in *Drosophila buzzatii*. *Chromosoma*. 118(3): 349-360.

Powell JR (1997) Population Genetics: Inversions. *In: Progress and prospects in evolutionary biology: the Drosophila model*. Oxford University, New York, pp 48-80.

Powell JR, Dion K, Papaceit M, Aguadé M, Vicario S, et al. (2011) Nonrecombining genes in a recombination environment: the *Drosophila* “dot” chromosome. *Mol. Biol. Evol.* 28: 825–833.

Puerma E, Orengo DJ, Salguero D, Papaceit M, Segarra C, Aguadé M (2014) Characterization of the breakpoints of a polymorphic inversion complex detects strict and broad breakpoint reuse at the molecular level. *Mol Biol Evol* 31(9): 2331-2341.

Puig M, Caceres M, Ruiz A (2004) Silencing of a gene adjacent to the breakpoint of a widespread *Drosophila* inversion by a transposon-induced antisense RNA. *PNAS* 101 (24): 9013–9018.

Rane RV, Rako L, Kapun M, Lee SF, Hoffmann AA (2015) Genomic evidence for role of inversion 3RP of *Drosophila melanogaster* in facilitating climate change adaptation. *Mol Ecol* 24(10): 2423-2432.

Ranz JM, Segarra C, Ruiz A (1997) Chromosomal homology and molecular organization of Muller's elements D and E in the *Drosophila repleta* species group. *Genetics* 145(2): 281-295.

Ranz JM, Casals F, Ruiz A (2001) How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res.* 11: 230–239.

Ranz JM, González J, Casals F, Ruiz A (2003) Low occurrence of gene transposition events during the evolution of the genus *Drosophila*. *Evolution* 57: 1325–1335.

Ranz JM, Maurin D, Chan YS, von Grotthuss M, Hillier LW, *et al.* (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. *Journal of Heredity* 87: 191-198.

Reinhardt JA, Kolaczowski B, Jones CD, Begun DJ, Kern AD (2014) Parallel geographic variation in *Drosophila melanogaster*. *Genetics* 197(1): 361-73. doi: 10.1534/genetics.114.161463.

Richards S, Liu Y, Bettencourt BR, Hradecky P, Letovsky PS, *et al.* (2005) Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution. *Genome Res.* 15: 1–18.

Rieger TT (1999) Mapeamento por hibridização *in situ* e expressão no desenvolvimento de genes de resposta a estresses e ao hormônio ecdisona em espécies do grupo *willistoni* de *Drosophila*. Tese de Doutorado, Universidade Federal do Rio Grande do Sul. Porto Alegre/RS, 138pp.

Rohde C, Abdelhay E, Pinto Júnior H, Schrank A, Valente VL (1995) Analysis and *in situ* mapping of the *Adh* locus in species of the *willistoni* group of *Drosophila*. *Cytobios* 81(324): 37-47.

Rohde C, De Toni DC, Valente VLS (1998) Um caso raro de inversão pericêntrica em populações naturais de *Drosophila willistoni*. *Genetics and Molecular Biology* 21 (Suppl.): 54.

Rohde C, Degrandi TH, De Toni DC, Valente VLS (2005) *Drosophila willistoni* polytene chromosomes. I. Pericentric inversion on X chromosome. *Caryologia* 58(3): 249-254.

Rohde C, Garcia AC, Valiati VH, Valente VL (2006) Chromosomal evolution of sibling species of the *Drosophila willistoni* group. I. Chromosomal arm IIR (Muller's element B). *Genetica* 126(1-2): 77-88.

Rohde C, Valente VLS (2012) Three decades of studies on chromosomal polymorphism of *Drosophila willistoni* and description of fifty different rearrangements. *Genet Mol Biol* 35(4):966-79

- Runcie DE, Noor MA (2009) Sequence signatures of a recent chromosomal rearrangement in *Drosophila mojavensis*. *Genetica* 136(1): 5-11.
- Ruiz A, Fontdevilla A Wasserman M (1982) The Evolutionary History of *Drosophila buzzatii*. III. Cytogenetic Relationships between Two Sibling Species of the *buzzatii* Cluster. *Genetics* 101(3-4): 503-518.
- Ruiz A, Wasserman M (1993) Evolutionary cytogenetics of the *Drosophila buzzatii* species complex. *Heredity* 70: 582–596.
- Ruiz A, Ranz JM, Caceres M, Segarra C (1997) Chromosomal evolution and comparative gene mapping in the *Drosophila repleta* species group. *Braz. J. Genet.* 20: 553–565.
- Ruiz-Herrera A, Castresana J, Robinson TJ (2006) Is mammalian chromosomal evolution driven by regions of genome fragility? *Genome Biol* 7: R115. doi: 10.1186/gb-2006-7-12-r115.
- Sachs RK, Levy D, Chen AM, Simpson PJ, Cornforth MN, et al. (2000) Random breakage and reunion chromosome aberration formation model; an interaction-distance version based on chromatin geometry. *Int J Radiat Biol.* 76(12): 1579-1588.
- Sambrook J, Russel DW (2004) *Molecular Cloning – a laboratory manual*. 3th edition. Vol. 1. Cold Spring Harbor Laboratory.
- Sandmeyer SB, Aye M, Menees T (2002). Ty3, a position-specific gypsy-like element in *Saccharomyces cerevisiae*. In *Mobile DNA II*, NL Craig, R Craigie, M Gellert and AM Lambowitz (eds), American Society of Microbiology Press, Washington pp. 663–683.
- Santos RA, Valente VLS (1990) On the occurrence of *Drosophila paulistorum* Dobzhansky & Pavan (Diptera, Drosophilidae) in urban environment: ecological and cytological observations. *Evol Biol* 4: 253-268.
- Santos-Colares MC, Degrand TH, Valente, VLS (2004) Cytological detection of male recombination in *Drosophila willistoni*. *Cytologia (Tokyo)*, 69(4): 359–365. doi:10.1508/cytologia.69.359.
- Schaeffer SW, Goetting-Minesky MP, Kovacevic M, Peoples JR, Graybill JL, et al. (2003) Evolutionary genomics of inversions in *Drosophila pseudoobscura*: evidence for epistasis. *Proc Natl Acad Sci U S A* 8;100(14): 8319-8324.
- Schaeffer SW, Bhutkar A, McAllister BF, 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.
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, et al. (2010) "Genome sequence of the palaeopolyploid soybean." *Nature* 463(7278), 178-183.
- Schoen DJ (2000) Comparative genomics, marker density and statistical analysis of chromosome rearrangements. *Genetics* 154(2): 943-952.

Schultz J (1936) Variegation in *Drosophila* and the Inert Chromosome Regions. Proc Natl Acad Sci U S A. 22(1): 27–33.

Singh BN, Das A (1992) Further evidence for latitudinal inversion clines in natural populations of *Drosophila melanogaster* from India. J Hered 83(3): 227-230.

Singh BN (2000) *Drosophila ananassae*: A species characterized by several unusual genetic features. Current Science, 78 (4): 391-398.

Soler AM, B Goñi 2012. Linkage relationships of spontaneous mutations in *Drosophila willistoni*. (Mutation Note). Dros. Inf. Serv. 95: 129-139.

Sonoda E, Hochegger H, Saberi A, Taniguchi Y, Takeda S (2007) Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. DNA Repair 5: 1021-1029.

Souvorov A, Kapustin Y, Kiryutin B, Chetvernin V, Tatusova T, Lipman D (2010) Gnomon - NCBI Eukaryotic gene prediction tool. Bethesda MD: National Center for Biotechnology Information.

Spassky B, Dobzhansky T (1950) Comparative genetics of *Drosophila willistoni*. Heredity 4(2): 201-15.

Spassky B, Richmond RC, Perez-Salas S, Pavlovsky O, Mourão CA, *et al.* (1971) Geography of the sibling species related to *Drosophila willistoni*, and of the semispecies of the *Drosophila paulistorum* complex. Evolution 25: 129-143.

Sperlich D, Pfriem P (1986) Chromosomal polymorphism in natural and experimental population. *In: The Genetics and Biology of Drosophila*. Ashburner M, Carson HL and Thompson Jr. JN (eds.). Vol. 3c. Academic Inc., London, pp. 257-309.

St. Pierre SE, Ponting L, Stefanicsik R, McQuilton P, the FlyBase Consortium (2014) FlyBase 102—advanced approaches to interrogating FlyBase. Nucl. Acids Res. 42(Database issue): D780–788.

Stefansson H, Helgason A, Thorleifsson G, Steinthorsdottir V, Masson G, *et al.* (2005) A common inversion under selection in Europeans. Nat. Genet 37:129–37.

Sturtevant AH (1931) Known and probable inverted sections of the autosomes of *Drosophila melanogaster*. Pub. Carnegie Instn. 421: 21-27.

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 Sci U S A. 22(7): 448-450.

Sturtevant AH, Nivitski E (1941) The homologies of the chromosome elements in the genus *Drosophila*. Genetics 26: 517-541.

Svoboda YH, Robson MK, Sved JA (1995) *P*-element-induced male recombination can be produced in *Drosophila melanogaster* by combining end-deficient elements in *trans*. Genetics 139(4):1601-1610.

Swanson CP, Merz T, Young WJ (1967) Citogenética, São Paulo, Ed. Polígono, 244p.

Swift H (1962) Nucleic acids and cell morphology in dipteran salivary glands. pp. 73-125. The Molecular Control of Cellular Activity. Edited by J. M. Allen. McGraw Hill, New York.

Talbert P B, Henikoff S (2000). A Reexamination of Spreading of Position-Effect Variegation in the *white-roughest* Region of *Drosophila melanogaster*. Genetics 154: 259–272.

Thibaud-Nissen F, Souvorov A, Murphy T, *et al.* (2013) Eukaryotic Genome Annotation Pipeline. *In: The NCBI Handbook* [Internet]. 2nd edition. Bethesda (MD): National Center for Biotechnology Information (US); 2013-. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK169439/>

Throckmorton LH (1982) The *virilis* species group. *In: The genetics and biology of Drosophila*. Vol. 3b. Ashburner M, Carson HL, Thompson JN (eds). London: Academic Press; pp: 227–296.

Tonzetich J, Lyttle TW, Carson HL (1988) Induced and natural break sites in the chromosomes of Hawaiian *Drosophila*. Proc Natl Acad Sci U S A. 85(5): 1717-1721.

Townsend JI (1952) Genetics of marginal populations of *Drosophila willistoni*. Evolution 6: 428-442.

Townsend JI (1958) Chromosomal polymorphism in Caribbean Island populations of *Drosophila willistoni*. PNAS USA 44: 38-42.

Tubio MJ, Costas CJ, Naveira FH (2003) "GYPSY14_AG, a member of the mdg1 lineage of the Ty3/gypsy group of LTR retrotransposons." Rebase Reports 3(9): 170-170.

Valente VLS, Morales NB (1985) New inversions and qualitative description of inversion heterozygotes in natural populations of *Drosophila willistoni* inhabiting two different regions in the State of Rio Grande do Sul, Brazil. Braz J Genet 8: 167-173.

Valente VLS, Araújo AM (1985) Observations on the chromosomal polymorphism of natural populations of *Drosophila willistoni* and its association with the choice of feeding and breeding sites. Braz J Genet 8: 271-284.

Valente VLS, Araújo AM (1986) Chromosomal polymorphism, climatic factors, and variation in population size of *Drosophila willistoni* in southern Brazil. Heredity 57: 149-159.

Valiati VH, Valente VLS (1997) Chromosomal polymorphism in urban populations of *Drosophila paulistorum*. Braz J Genet 20 (4): 567-581.

Valente VLS, Rohde C, Valiati V, Morales NB, Goñi B (2001) Chromosome inversions occurring in Uruguayan populations of *Drosophila willistoni*. DIS 84: 55-59.

Valente VLS, Goñi B, Valiati VH, Rohde C, Morales NB (2003) Chromosomal polymorphism in *Drosophila willistoni* populations from Uruguay. *Genetics and Molecular Biology* 26(2): 163-173.

Vavilov NI (1926) *Studies on the origin of cultivated plants*. Leningrad: Institut Botanique Appliqué et d'Amélioration des Plantes.

von Grotthuss M, Ashburner M, Ranz JM (2010) Fragile regions and not functional constraints predominate in shaping gene organization in the genus *Drosophila*. *Genome Res.* 20(8): 1084-1096.

Wallace AG, Detweiler D, Schaeffer SW (2011) Evolutionary history of the third chromosome gene arrangements of *Drosophila pseudoobscura* inferred from inversion breakpoints. *Mol Biol Evol.* 28(8): 2219-2229.

Wasserman M (1960) Wasserman, M. (1960). Cytological and phylogenetic relationships in the *repleta* group of the genus *Drosophila*. *Proc Natl Acad Sci USA*, 46(6): 842-859.

Wesley CS, Eanes WF (1994) Isolation and analysis of the breakpoint sequences of chromosome inversion In(3L)Payne in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 12; 91(8): 3132–3136.

Zhang Y, Sturgill D, Parisi M, Kumar S, Oliver B (2007) Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. *Nature* 450: 233–237.

Zhang J, Yu C, Pulletikurti V, Lamb J, Danilova T, Weber DF, Birchler J, Peterson T (2009) Alternative *Ac/Ds* transposition induces major chromosomal rearrangements in maize. *Genes Dev* 23: 755–765.

Zelentsova H, Poluectova H, Mnjoian L, Lyozin G, Veleikodvorskaja V, *et al.* (1999) Distribution and evolution of mobile elements in the *virilis* species group of *Drosophila*. *Chromosoma* 108(7): 443-56.

Zhimulev, Igor F, Koryakov DE (2009) *Polytene Chromosomes*. In: eLS. John Wiley & Sons Ltd, Chichester. <http://www.els.net> [doi: 10.1002/9780470015902.a0001183.pub2].

Zivanovic G, Mestres F (2012) Changes in chromosomal polymorphism and global warming: The case of *Drosophila subobscura* from Apatin (Serbia). *Genetics and Molecular Biology*, 34 (3): 489-495.

Anexo I

Rearranjos cromossômicos estruturais em organismos diploides: Consequências meióticas e mecanismos de origem.

Os cromossomos eucarióticos são estruturas que se encontram no núcleo das células em uma máxima taxa de condensação, compostos por ácido desoxirribonucleico e proteínas associadas. Esta organização abriga o código da vida: os genes.

Os rearranjos cromossômicos estruturais (também denominados Alterações/Aberrações cromossômicas estruturais) compreendem as modificações em larga escala da composição estrutural cromossômica, visíveis microscopicamente e oriundas de quebras seguidas da reconstituição do cromossomo em uma composição anormal. Essas alterações são estáveis quando são capazes de serem transmitidas, de maneira inalterada, por meio de várias divisões celulares.

Os diferentes rearranjos estruturais classificam-se em rearranjos intracromossômicos, nos quais as alterações envolvem somente um cromossomo; ou rearranjos intercromossômicos, envolvendo dois ou mais cromossomos simultaneamente. Compreendem os tipos (Revisões em Macgregor, 1993; Gupta, 2007; Griffiths *et al.* 2008; Auger & Sheridan, 2011) (Figura 1):

- **Deleções:** Perda de um segmento do cromossomo podendo esta ser terminal, resultante de uma quebra e acarretando a perda da porção terminal de um cromossomo; ou intersticial resultante de duas quebras. Quando da ocorrência de deleções terminais em ambas as extremidades teloméricas de um cromossomo, estas extremidades rompidas podem se fundir dando origem a uma estrutura circular denominada cromossomo em anel. Por sua vez, as deleções que afetam o centrômero são perdidas, uma vez que sem o centrômero o cromossomo não é hábil de se mover durante o ciclo celular. O efeito da deleção é diretamente relacionado ao tamanho do fragmento perdido, sendo que algumas deleções são viáveis pela combinação com seu respectivo alelo normal; bem como deleções de

alelos selvagens acabam por dar oportunidade de expressão fenotípica para alelos recessivos deletérios, o que se denomina Pseudodominância, fenômeno comprovado principalmente pelos estudos de mutações induzidas (Chikushi, *et al.* 1971). Este rearranjo pode ocorrer no estado homocigoto ou heterocigoto, neste último caso resultando na formação de uma alça para o pareamento dos cromossomos homólogos na meiose de uma deleção intersticial. Estas regiões apresentam taxas bastante diminuídas de recombinação na meiose, uma vez que os gametas portando grandes deleções são letais, assim como as deleções em estado homocigoto na prole.

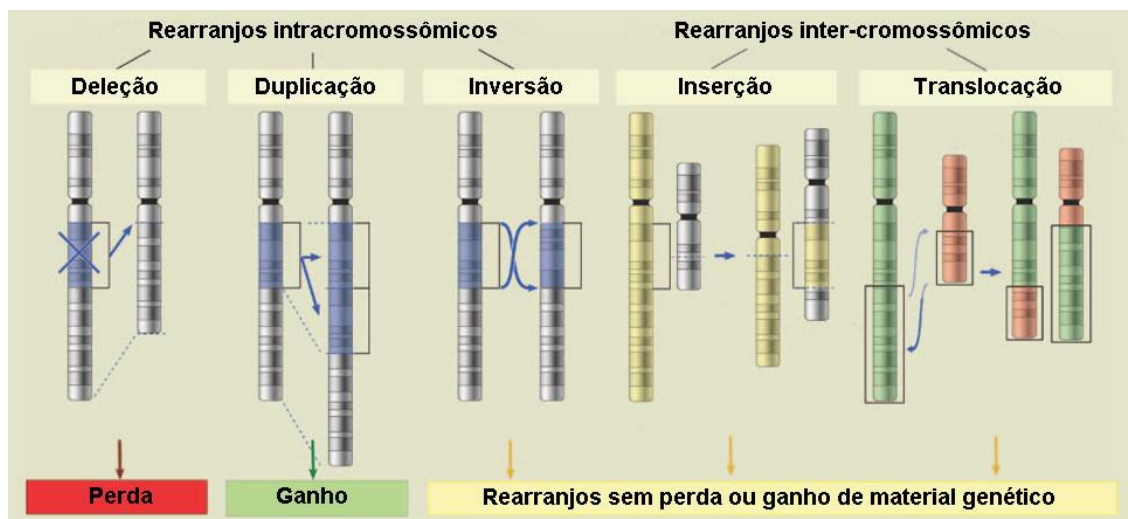


Figura 1: Representação esquemática dos principais tipos de rearranjos cromossômicos estruturais. Retirado e modificado de Alqallaf *et al.* (2013).

- **Inserções:** Envolve a transposição de um segmento de um dado cromossomo para outra posição do mesmo cromossomo, ou de um cromossomo distinto (inserção não homóloga). Se a inserção não acarretar em ganho ou perda de material genético, será estável. Entretanto, a inserção de um segmento pode não resultar na deleção do fragmento em seu sítio original, resultando na duplicação deste, estando assim fortemente relacionada aos eventos de duplicação. Como as deleções, em estado heterocigoto, grandes inserções formam uma alça para o pareamento na meiose.

- Duplicações:** Repetição de um segmento cromossômico que pode estar ou no mesmo cromossomo da sequência original, ou inserida em outro cromossomo, geralmente sendo mais bem aceitas nos diferentes organismos do que as deleções. As duplicações podem ser inseridas: (i) de maneira contígua na mesma ordem e no mesmo braço cromossômico (duplicações *in tandem*), tendo como exemplo clássico a mutação dominante *Bar* no cromossomo X de *Drosophila melanogaster* (Sturtevant, 1925); (ii) de maneira contígua no mesmo braço cromossômico, mas com a ordem invertida (duplicações reversas, as quais envolvem uma inversão); (iii) no mesmo braço cromossômico, porém em sítios afastados; (iv) em diferentes braços cromossômicos, do mesmo cromossomo; (v) em diferentes cromossomos. Na meiose, as duplicações podem assumir diferentes configurações, em divisões distintas (Figura 2).

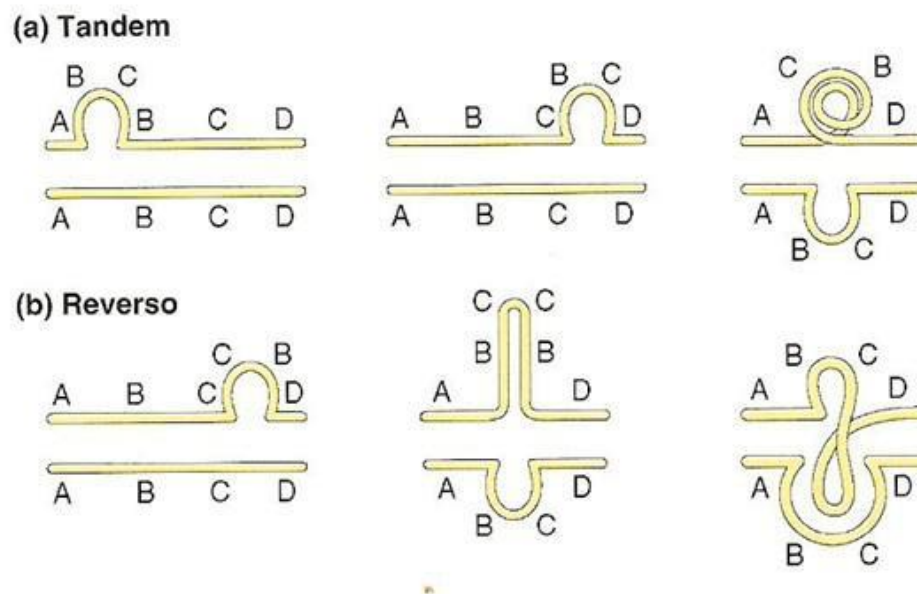


Figura 2: Representação esquemática das diferentes configurações meióticas de duplicações cromossômicas heterozigotas. a) duplicações *in tandem* com mesma ordem. b) Duplicações *in tandem* com ordem reversa. Retirado de Griffiths *et al.* (2008).

Apesar dos efeitos deletérios que as duplicações podem acarretar, este rearranjo tem sido considerado um escape da poderosa pressão da seleção natural à manutenção dos genes com funções vitais. Isto porque as cópias oriundas de

duplicações não sofrem tamanha pressão, podendo com isto se diferenciar a tal ponto de originar novos genes com novas funções, ou funções mais especializadas; e assim, desempenhando um importante papel na evolução dos seres vivos (Ohno, 1970), sendo esta característica melhor visualizada em categorias taxonômicas superiores, tais como filo e ordem.

- **Inversões:** é o rearranjo cromossômico resultante de duas quebras e o posterior giro de 180° da sequência envolvida, sendo assim reconstituídas com a ordem invertida. Podem ser de dois tipos: paracêntricas (as quais não envolvem o centrômero na formação da inversão, ocorrendo no mesmo braço cromossômico), e pericêntricas (as quais envolvem o centrômero e mais de um braço cromossômico). Na meiose podem ser visualizadas no estado homizigoto ou no estado heterozigoto. Neste último estado ocorre a formação de uma alça de inversão para o pareamento dos cromossomos homólogos. A ocorrência de recombinação meiótica, dentro da alça de uma inversão heterozigota paracêntrica, leva a formação na meiose I de um cromossomo dicêntrico (presença de dois centrômeros) e um fragmento acêntrico (ausência do centrômero), os quais resultam em gametas com deleções e duplicações. Já a recombinação envolvendo inversões pericêntricas dá origem a gametas portando deleções e duplicações (**Figura I.2, Capítulo I**). A proporção de produção de gametas inviáveis vai depender da taxa de recombinação meiótica ocorrente. Este rearranjo em si não implica em ganho ou perda de material genético, sendo assim mais viáveis que as demais alterações. Entretanto, o caso de inversões ocorrentes em genes com funções essenciais acarretará em mutações gênicas letais ou em efeito de posição.

- **Isocromossomos:** Formam-se, em geral, quando a divisão dos centrômeros dá-se transversalmente, quando deveria ocorrer longitudinalmente. Forma-se então uma cromátide somente com braços curtos, e outra somente com braços longos. O cromossomo que perdeu um de seus braços o substitui com uma cópia exata do seu outro braço. O portador terá monossomia de um dos braços e trissomia parcial do outro braço cromossômico.

- **Translocações:** De dois tipos principais: (i) Translocação recíproca, a mais comum, onde ocorre a troca mútua de segmentos cromossômicos entre cromossomos não homólogos. Podem alterar de maneira drástica a morfologia dos

cromossomos; (ii) Translocação robertsoniana, um tipo especial de translocação, ocorre somente nos cromossomos acrocêntricos. As quebras que ocorrem próximas aos centrômeros destes cromossomos são sucedidas pela fusão dos braços longos originando um cromossomo metacêntrico ou submetacêntrico. Na meiose, os cromossomos portadores de translocações recíprocas costumam parrear-se em configuração quadrivalente (em cruz) (Figura 3). Desta maneira, três padrões de segregação cromossômica, sem ocorrência de recombinação meiótica, podem ocorrer: segregação alternada ($AB+DC$ ou $AD+BC$, Figura 3C), segregação adjacente I ($AB+BC$ ou $AD+DC$) e segregação adjacente II ($BC+CD$ ou $BA+AD$). Os gametas resultantes das segregações adjacente-1 e adjacente-2 não são viáveis devido às deleções e duplicações presentes nestes, e que acarretam em um desequilíbrio do conteúdo gênico.

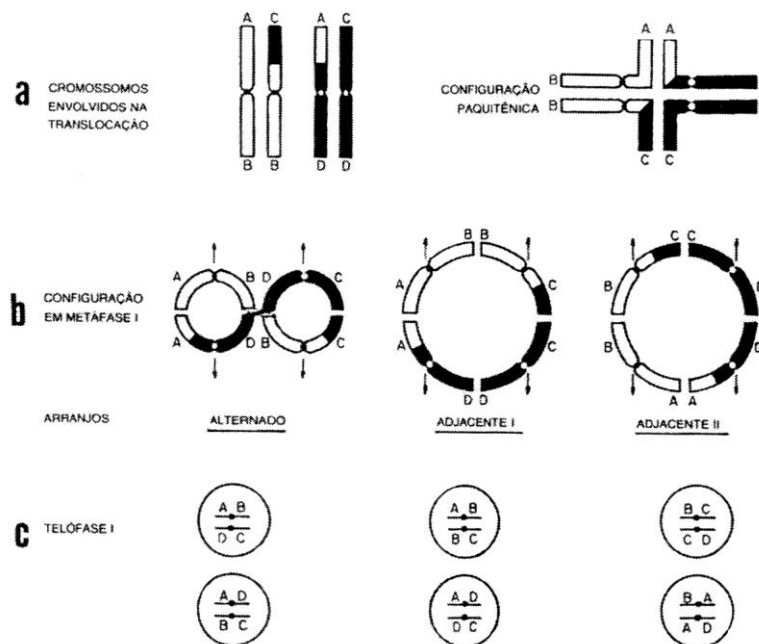


Figura 3: Representação esquemática do pareamento meiótico de cromossomos não homólogos com translocação recíproca (prófase I) (a), possíveis orientações dos centrômeros (setas) obtidas na metáfase I (b) e gametas formados ao final da meiose II a partir das diferentes segregações cromossômicas (Modificado de Guerra, 1988).

A compreensão da origem dos diferentes tipos de rearranjos cromossômicos estruturais necessita levar em consideração alguns pontos-chaves (Griffiths *et al.*, 2008):

- 1) Cada cromossomo é uma única molécula de DNA de fita dupla.
- 2) O primeiro evento que vai culminar na origem de um rearranjo cromossômico estrutural é a geração de duas ou mais quebras nos cromossomos de uma célula. Estas quebras podem ser simples, ocorrendo em uma das fitas do DNA (em inglês: *Single Strand Break, SSB*); ou pode ser uma quebra dupla, ocorrendo em ambas as fitas do DNA (em inglês: *Double Strand Break, DSB*).
- 3) Estas quebras são potencialmente letais, a menos que sejam reparadas.
- 4) As células possuem sistemas de reparo que corrigem estas quebras, reunindo novamente as pontas quebradas.
- 5) Se o mecanismo de reparo reúne as duas pontas da mesma quebra, a ordem original do DNA é restaurada. Entretanto, se o reparo reúne as pontas de duas quebras distintas, isto acarreta na origem de um rearranjo cromossômico estrutural.
- 6) Os diferentes rearranjos cromossômicos estruturais para sobreviverem à meiose devem ter um único centrômero e dois telômeros. Caso estas características não ocorram, o rearranjo estrutural vai se perder no ciclo celular. Cabe ressaltar que a ocorrência de cromossomos com centrômeros difusos, denominado cromossomo holocêntrico, é relatada para muitas espécies (Mandrioli & Manicardi, 2012).
- 7) Caso estas alterações ocasionem a duplicação ou a deleção de algum segmento do cromossomo, acarretará no desequilíbrio do balanço gênico, sendo este desequilíbrio diretamente relacionado ao tamanho da alteração estrutural.

De acordo com a premissa de que os rearranjos cromossômicos estruturais se originam da quebra cromossômica, e reorganização deste de uma maneira diferente da inicial, surge a intrigante questão: quais os mecanismos que originam estes rearranjos? No que concerne à origem das quebras, tradicionalmente vem

sendo classificada entre induzida ou espontânea. As alterações com quebras cromossômicas induzidas seriam decorrentes de algum fator externo que causasse dano ao material genético dos diferentes organismos, tais como substâncias químicas e radiações ionizantes. Por sua vez as alterações cromossômicas espontâneas se originam devido às quebras ocorrentes no material genético nos diferentes processos biológicos que este passa no interior das células (Darlington & Koller, 1947).

Na categoria das alterações induzidas, os diferentes tipos de rearranjos cromossômicos estruturais acima citados podem ser originados via a exposição às radiações ionizantes raios- γ e raios-X. Estas radiações ionizantes, devido à emissão de seus elétrons, alcançam energia suficiente para ionizar, ou seja, arrancar elétrons; de diferentes átomos, dentre estes os átomos da molécula de DNA nos cromossomos. Isto resulta na formação do par íon negativo (elétron arrancado) e íon positivo (átomo que perdeu um elétron). Este elétron arrancado pela ionização origina uma desestruturação da molécula afetada, rompendo suas ligações químicas, causando assim, no caso da molécula de DNA, mutações genéticas, morte celular, efeitos bioquímicos e fisiológicos (alterações morfológicas e funcionais dos órgãos) (Okuno, 2013). Muito do conhecimento inicial acerca dos tipos e efeitos dos rearranjos cromossômicos estruturais são devidos aos estudos que utilizaram este recurso. Os estudos que envolviam a análise citogenética das diferentes alterações estruturais foram muito explorados em diferentes organismos modelos da citogenética, destacando-se entre estes a *Drosophila melanogaster*.

Com a descrição das alterações cromossômicas estruturais induzidas nos cromossomos politênicos de *Drosophila melanogaster* por Bridges (1923), uma vez que os rearranjos ocorrentes de forma espontâneas eram muito raros, corroboraram as descobertas pioneiras de Muller (1927), de que os Raios-X acarretavam em alterações cromossômicas estruturais, as quais eram visíveis nas descendências de machos da espécie *D. melanogaster* que foram submetidos a diferentes intensidades de radiação. Herman Joseph Muller, que desenvolvia seus estudos acerca da recombinação meiótica (*crossing over*), mostrou que com esta técnica, em poucas semanas, era possível obter uma gama de mutantes em nível cromossômico e fenotípico, resultando na principal forma de análise até o advento das técnicas moleculares entre 1960/1970. Com isto deu-se a experimentação em

larga escala de muitos problemas genéticos, e os estudos dos rearranjos cromossômicos estruturais originados por agentes químicos e físicos proporcionaram à ciência, um largo alcance de conhecimento acerca do comportamento meiótico e mitótico destes rearranjos ocorrentes nos mais diversos seres vivos na natureza (Schulz-Schaeffer, 1980). Atualmente, sabe-se que os mecanismos subsequentes às quebras cromossômicas induzidas são os mesmos que dão origem aos rearranjos espontâneos. O avanço no conhecimento acerca da biologia celular e molecular veio elucidar a origem destas quebras cromossômicas espontâneas, e os mecanismos que originam as alterações (Haber, 2000).

As quebras de fita dupla é uma característica integrante de vários processos que ocorrem no âmbito biológico da célula entre eles: quebras causadas por subprodutos metabólicos da respiração celular, Recombinação V(D)J, estresse físico durante o fuso meiótico ou mitótico, e a quebras durante o ciclo meiótico que objetivam eventos de recombinação entre os cromossomos homólogos. Estas quebras podem resultar na morte celular, ou na origem de várias alterações genéticas (Shrivastav *et al.*, 2008).

As quebras de fita dupla durante a meiose ocorrem na Prófase I, por meio da clivagem inicial da enzima conservada *tipo-topoisomerase II Spo11*, uma para cada fita. Após os cortes e a dissociação de *Spo11*, as pontas 5' são processadas e sofrem uma diminuição nucleotídica (perda de uma pequena porção terminal), expondo a ponta 3' de cada fita de DNA, um processo que envolve pelo menos uma dezena de proteínas (Lam & Keeney, 2014). Com isto, ocorre uma disputa entre os principais mecanismos de reparo de quebras de fita dupla: recombinação homóloga (do inglês *Homologous recombination, HR*), Recombinação homóloga não alélica (do inglês "*Nonallelic homologous recombination*" NAHR) e União das extremidades não homólogas (do inglês, *non-homologous end-joining, NHEJ*), com o uso de sofisticadas maquinarias proteicas para cada via de reparo (Shrivastav *et al.*, 2008).

Recombinação homóloga: é considerada uma via de reparo livre de erro. Ocorre preferencialmente durante a fase S e início da fase G2, após a replicação do material genético. Em *Saccharomyces cerevisiae* (e extremamente conservado entre os demais eucariotos), subsequente à criação das quebras de fita dupla, uma

das fitas simples com extremidade 3' livre e diminuída pode invadir a cromátide irmã, sendo que na meiose a preferência é a invasão do cromossomo homólogo. Esta invasão requer a participação da proteína *Rad51p*, a qual atua na permuta entre as fitas de DNA, junto a outras proteínas associadas, tais como *Rad52p*, *Rad54p*, *Tid1p* e *Dmc1p*. Com as sequências complementares alinhadas, ocorre a síntese de DNA a partir da extremidade 3' livre, utilizando o DNA da cromátide irmã, ou cromossomo homólogo não danificado como molde. O reparo por recombinação homóloga pode levar ao evento de recombinação meiótica (*crossing over*) entre os cromossomos homólogos. A enzima *Spo11* desencadeia múltiplas quebras de fita dupla ao longo de todos os cromossomos, entretanto, poucos reparos por recombinação homóloga vão culminar em recombinação meiótica. Somente aqueles casos onde a fita simples invasora encontra em processo de síntese a subsequente extremidade 3' livre, dando origem às Junções de Holliday (Figura 4).

A recombinação meiótica tem sido implicada com a ausência da proteína *Rad51p*, bem como com a existência de sítios preferenciais para a sua ocorrência (pontos quentes de recombinação) (Haber 2000; Neale & Keeney, 2006; Wolf-Dietrich *et al.*, 2010). No fungo *S. cerevisiae* 10 pontos quentes de recombinação já foram caracterizados em alta resolução molecular, e associações significativas foram estabelecidas com regiões intergênicas com hipersensibilidade às ações das nucleases, e regiões com alto conteúdo G + C somadas à modificação da cromatina (Petes, 2001).

Recombinação homóloga não alélica (do inglês “*nonallelic homologous recombination*” NAHR): também denominada Recombinação ectópica, é um mecanismo de reparo que utiliza a maquinaria enzimática da recombinação homóloga (*HR*) entre cromossomos homólogos e entre cromátides irmãs. Para a ocorrência de *NAHR* deve haver segmentos compartilhando alta similaridade, tais como os segmentos de DNA em baixo número de cópias (como as duplicações segmentares em humanos), ou regiões altamente repetitivas (tais como os elementos de transposição e DNAs satélites). Estes segmentos repetitivos encontram-se ou *in tandem*, ou dispersos pelo genoma e a recombinação ocorre entre cópias não alélicas, podendo estas estar localizadas em regiões intercromossômicas, intracromossômicas e intracromátides.

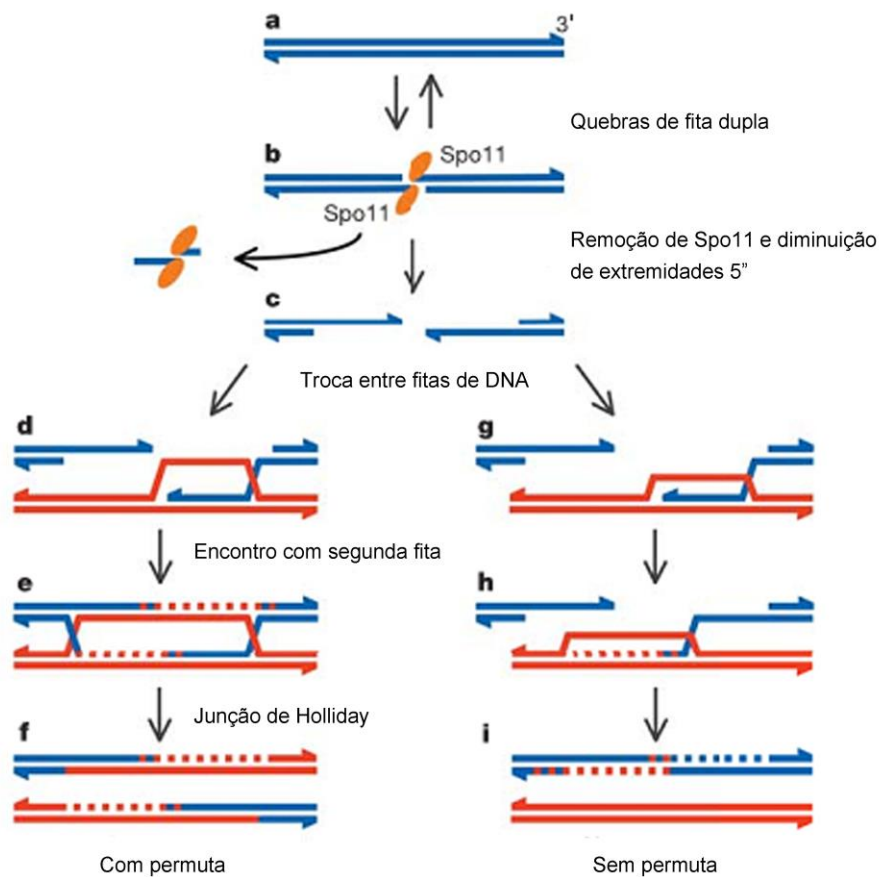


Figura 4: Representação esquemática dos eventos da recombinação homóloga. **a-c:** formação das quebras de dupla fita e diminuição de extremidades 5' (pontas em seta). **d-f:** recombinação meiótica com reparo e permuta. **h-i:** recombinação meiótica com reparo e sem permuta. Modificado de Lam & Keeney (2015).

Estas recombinações ectópicas potencialmente implicam na origem de diferentes rearranjos cromossômicos. A identidade mínima entre duas sequências necessárias para uma recombinação ectópica é chamada de *MEPS* (do inglês *Minimal effective processing segment*). Modelagem *in silico* infere que uma *MEPS* eficiente para *NAHR* é de aproximadamente 250 pb. No entanto, *in vivo* este parâmetro ainda não é efetivamente elucidado. Os diferentes rearranjos cromossômicos estruturais são dependentes da posição genômica das cópias, bem como da orientação destas (Figura 5) (Stankiewicz & Lupski, 2002; Hoang *et al.*, 2010).

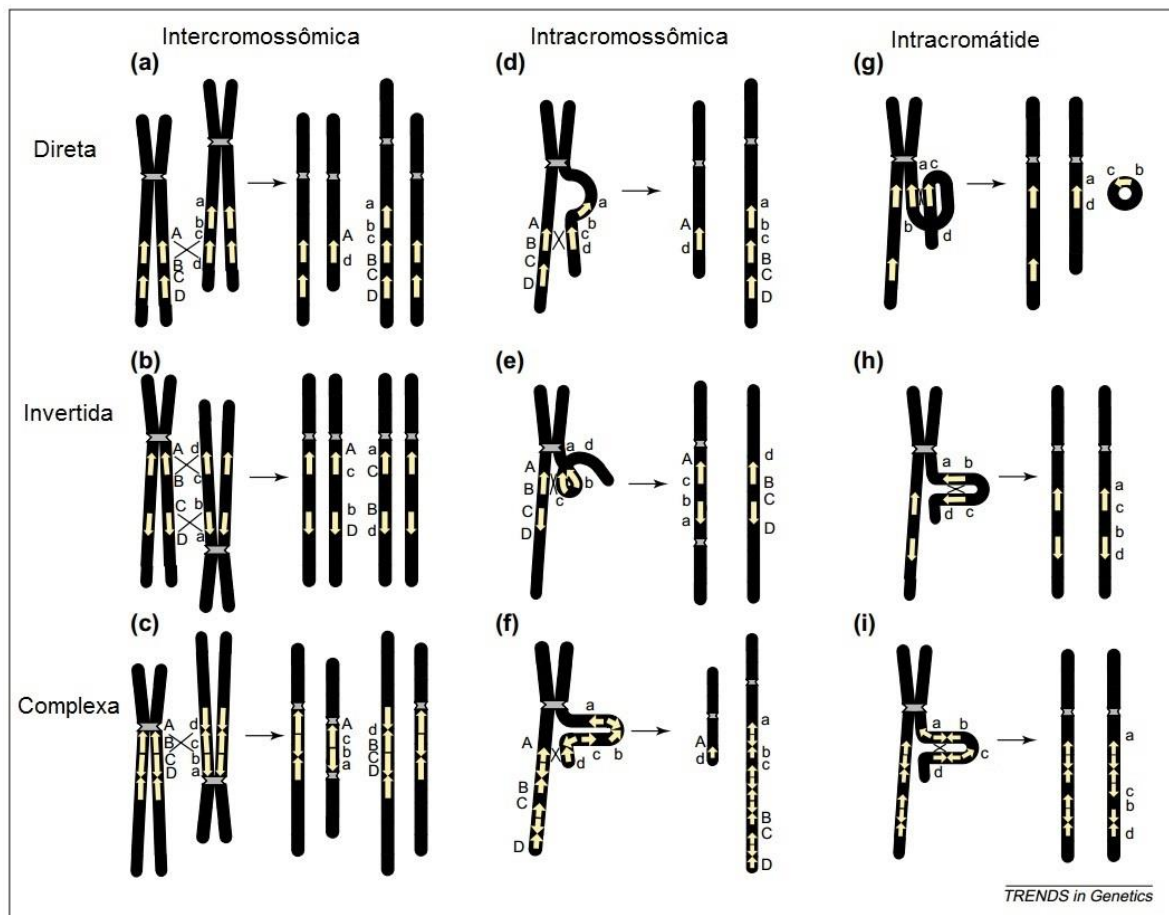


Figura 5: Representação esquemática dos rearranjos cromossômicos estruturais originados por eventos de recombinação homóloga não alélica. Repetições são representadas por setas dentro dos cromossomos. a, d, f) deleções e duplicações; b, h, i) inversão; c, e) Inversão e duplicação; g) Deleção e fragmento acêntrico. Retirado de Stankiewicz & Lupski (2002).

O mecanismo de recombinação ectópica tem sido implicado na origem de rearranjos cromossômicos estruturais entre indivíduos não relacionados. Inclusive, podendo não ser dependente de homologia para a recombinação. A recorrente translocação recíproca constitucional $t(11; 22)$ em seres humanos tem sua origem via recombinação não homóloga utilizando sequências palindrômicas ricas em AT nos cromossomos 11q e 22q, formando uma estrutura secundária em forma de grampo (*hairpin*) (Edelmann *et al.*, 2001).

União das extremidades não homólogas (do inglês, *non-homologous end-joining*, *NHEJ*): o mecanismo de reparo pela união de extremidades não

homólogas atua em todos os tipos de células, desde as bactérias até o homem, e está envolvido em muitos processos diferentes, tais como a reparação do DNA, a manutenção dos telômeros, e o reparo das quebras do sistema de recombinação V(D)J. Este mecanismo é primariamente acionado durante a fase G1 do ciclo celular, sendo que após esta fase, *HR* preferencialmente atua no reparo das quebras de fita dupla (Pastwa & Błasiak, 2003).

A via clássica (*cNHEJ*), de leveduras a mamíferos, requer o mesmo conjunto básico de proteínas: as proteínas de ligação às porções terminais de DNA *Ku70/80*, bem como a *DNA-ligase IV* e a sua proteína associada *XRCC4*. Células de vertebrados também requerem *DNA-PKcs*, as quais não têm homólogos em levedura. Em contrapartida, em levedura existe também uma exigência para *Rad50*, *Mre11* e *XRS2*, três proteínas com atividades de endo e de exonuclease, e a proteína *Artemis*. A ligação precoce das proteínas *Ku70/80* impede que as extremidades 5' e 3' sofram ação de exonucleases, sendo um ponto decisivo na escolha de *cNHEJ* para o reparo pela maquinaria celular. Não necessita de homologias para fazer a junção entre as pontas das fitas duplas de DNA (ligação de terminais cegos), característica garantida pela *XRCC4:ligase IV*; porém, geralmente deleções e inserções de nucleotídeos são observadas nas junções reparadas. O mecanismo de reparo *cNHEJ* evoluiu para reparar *DSBs* que necessitam serem unidas diretamente. Em organismos haploides em fase estacionária, não existe qualquer doador de homologia, e *RH* não é uma opção. Evolutivamente, assumindo que muitos organismos foram e são haploides, *cNHEJ* provavelmente representa umas das primeiras estratégias evolutivas para o reparo do DNA (Figura 6).

Esta via de reparo está fortemente relacionada ao ciclo de quebra, fusão e ponte (do inglês *breakage-fusion-bridge cycle* - *BFB*) (McClintock, 1941). Neste ciclo, quebras de fita dupla originam a quebra da porção terminal de um cromossomo (telômero), deixando este com sua porção terminal desprotegida. Após a duplicação na meiose I, via o mecanismo de reparo *cNHEJ*, estas pontas terminais quebradas das cromátides irmãs são fusionadas, dando origem a um cromossomo dicêntrico formando uma ponte, na qual cada centrômero é puxado em direção aos pólos distintos de uma só vez durante a anáfase II. Isto resulta na quebra deste cromossomo dicêntrico, e as extremidades quebradas permanecem

adesivas, fusionando-se subsequentemente à duplicação na prófase mitótica das células filhas.

O resultado é o surgimento recorrentes cromossomos dicêntricos, dando continuidade ao ciclo, com os cromossomos quebrando novamente a cada mitose. Uma vez que a cada ciclo, a quebra provavelmente ocorre em um local diferente do que nos anteriores, acarretará em um reagrupamento dos locos genéticos com deleções e duplicações gênicas. Os rearranjos cromossomicos estruturais resultantes deste ciclo têm sido encontrados em uma gama de tumores malignos analisados em humanos (Gisselsson *et al.*, 2000; Bunting & Nussenzweig, 2013).

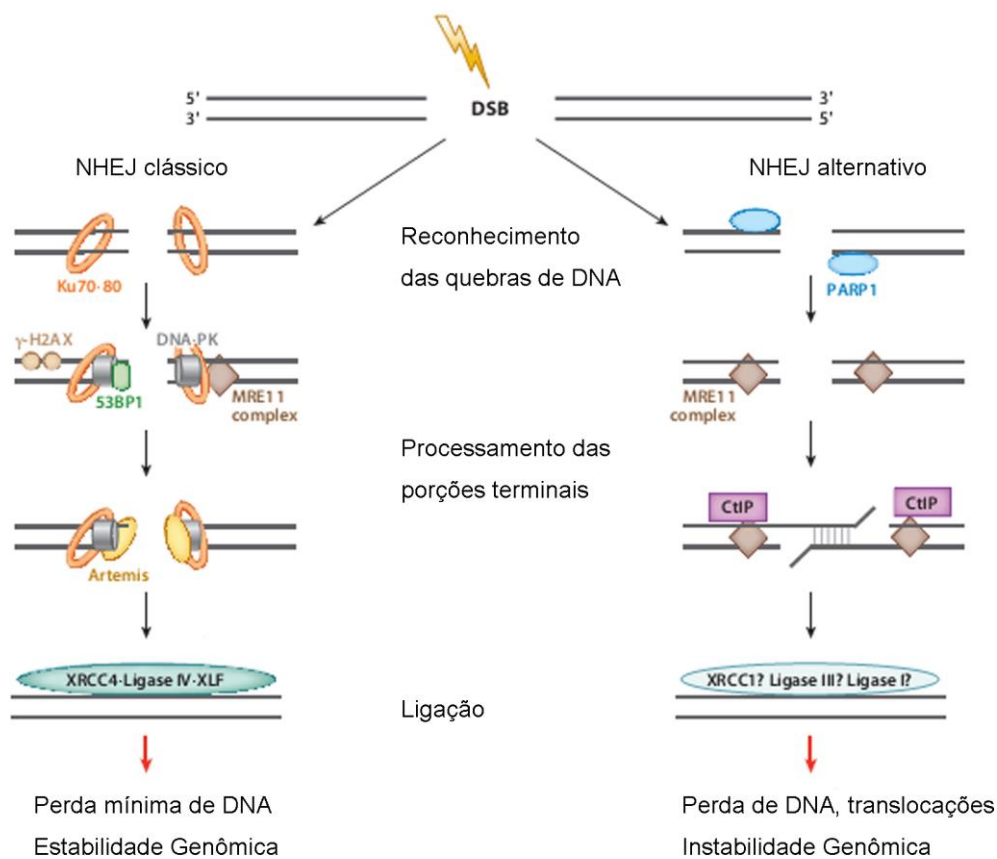


Figura 6: Representação esquemática das vias clássica e alternativa do mecanismo de reparo de quebras de fita dupla (DSB) União de extremidades não homólogas. Modificada de Deriano & Roth, 2013.

A via alternativa (*aNHEJ*) foi descoberta em leveduras e mamíferos aproximadamente ao mesmo tempo, como uma forma de sistema capaz de reparar quebras de fita dupla em células com deficiências para um, ou vários fatores críticos para *cNHEJ*. Entretanto a descoberta deste mecanismo em células eficientes também para *cNHEJ* indicou que este mecanismo não se trata somente de uma via alternativa utilizadas por células que tenham deficiência em *cNHEJ*, mas sim, que provavelmente se trata de um mecanismo sofisticado e com suas características inerentes (Figura 6).

De acordo com Decottignies (2013) a via alternativa do mecanismo *NHEJ* divide-se em duas categorias: *MMEJ* (do inglês *microhomology-mediated end-joining*). Esta via repara junções que exibam tanto pequenas deleções e micro homologias de 3 a 16 pb que se sobreponham (Figura 6). E o tipo *Alt-NHEJ*, mecanismo que não requer homologias pré-existentes. Dada à habilidade de ligar de maneira inapropriada sequências não contíguas, a origem de rearranjos cromossômicos tais como deleções, duplicações e translocações é fortemente relacionada aos mecanismos de *aNHEJ*, principalmente em células com a via *cNHEJ* defectiva. (Kasperek & Humphrey, 2011; Bunting & Nussenzweig, 2013).

Transposição aberrante: Elementos de Transposição (do inglês *Transposable elements – TEs*) são sequências constituintes da fração repetitiva do genoma, em praticamente todos os organismos eucariotos até então estudados. Intrinsecamente, eles possuem maquinaria enzimática para se mobilizarem de um sítio a outro dentro do genoma hospedeiro, de acordo com características inerentes (Feschotte & Pritham, 2007; Wicker *et al.*, 2007). As excisões, inserções ou duplicações dos *TES* originam quebras das fitas de DNA, que de acordo com o já exposto acima, podem desencadear a origem de diferentes rearranjos cromossômicos estruturais por meio de mecanismos de reparo. No entanto, o mecanismo de transposição diretamente responsável pelas quebras e junção do fragmento transposto surge como uma alternativa para mediar rearranjos cromossômicos. Sendo assim, de alguma forma, os *TEs* reorganizariam os genomas por si mesmos, como um subproduto da sua função normal (Weil, 2009). Este mecanismo tem sido denominado Transposição Alternativa/Transposição Aberrante, e tem sido descrito para os *TEs* de classe II (transposons), os quais utilizam um intermediário de DNA para se transpor (Gray, 2000).

Transposição alternativa envolve as porções terminais 5' de um *TE* e 3' de outro *TE*, formando um elemento híbrido. Estes podem estar na mesma cromátide, ou na cromátide irmã, ou no cromossomo homólogo parcialmente ou totalmente replicados. No entanto, em tais casos, o que é "excisado" e "reinserido" não é um típico transposon apenas, mas sim, uma região de DNA cromossômica. Dependendo do que está sendo inserido, sua orientação e local no genoma, rearranjos tais como duplicações, deleções, inversões e translocações podem ser originados (Figura 7) (Zhang & Peterson, 2004; Weil, 2009).

O mecanismo de transposição alternativa tem sido inferido particularmente por meio da associação dos elementos *Activator (Ac)/Dissociation (Ds)*, e sua consequente origem de rearranjos cromossômicos estruturais já foram descritos em milho (Zhang & Peterson, 2004; Zhang *et al.*, 2009;), *Arabidopsis* (Krishnaswamy *et al.*, 2008) e arroz (Yu *et al.*, 2012; Xuan *et al.*, 2015).

Mecanismos na replicação: Quebras de fita dupla também podem ser criadas por diversos problemas ocorrentes durante o mecanismo de replicação do DNA, podendo vir a desencadear a utilização de sistemas de reparo por recombinação. Entretanto, devido à presença de vários pontos de quebras envolvidos, e frequentemente a grande distância entre eles, a origem dos diferentes rearranjos cromossômicos observados não podem simplesmente serem devidas a eventos de recombinação, tais como *NAHR* ou *NHEJ*. Atualmente, processos alternativos baseados nos erros da replicação de DNA, têm sido propostos (Liu *et al.*, 2012). Dentre estes, se destaca o chamado "*Fork Stalling and Template Switching*" (*FoSTeS*) (Lee *et al.*, 2007).

No modelo *FoSTeS*, uma das fitas da forquilha de replicação de DNA desengata a partir do modelo original, em princípio devido ao encontro de um dano físico, estrutura secundária, ou DNA polimerase defectiva. Posteriormente, emparelha com outra forquilha de replicação em proximidade física, mas podendo efetivamente ser distante ou em outro cromossomo, em virtude de micro-homologia na porção terminal 3' (a qual pode ser fornecida pelas sequências repetitivas de DNA); reiniciando, ou começando uma nova síntese de DNA (Stankiewicz & Lupski, 2010).

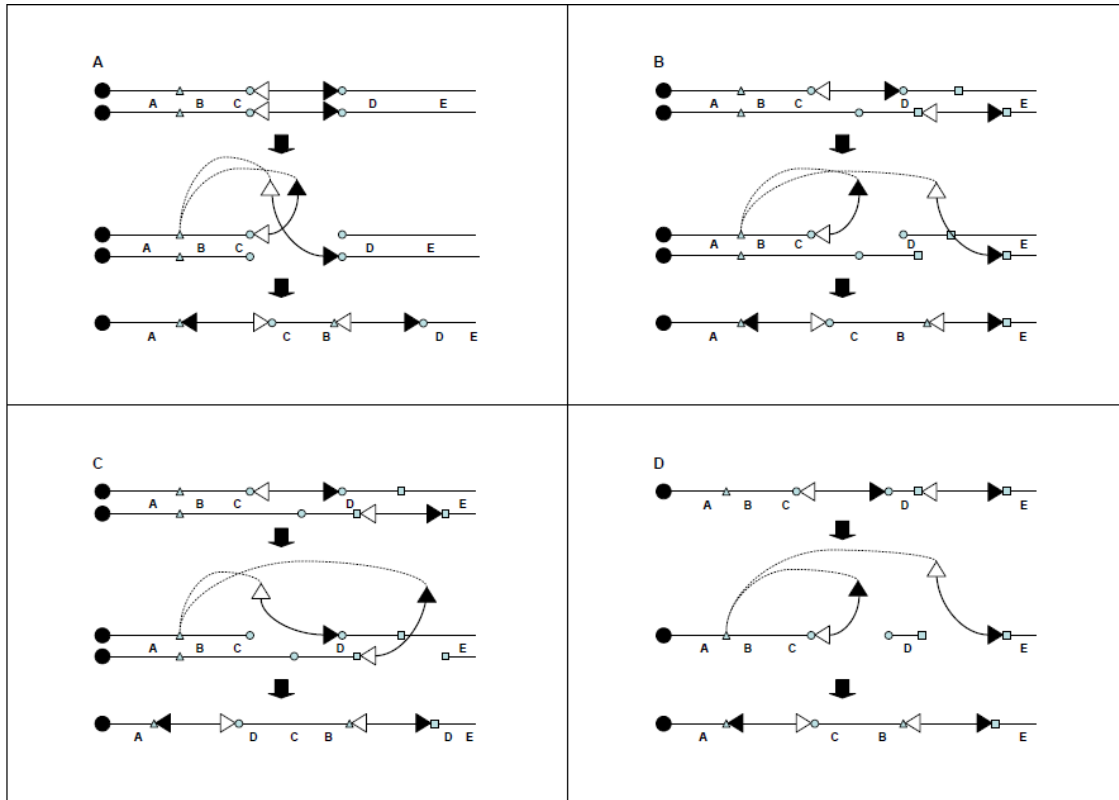


Figura 7: Representação esquemática da origem de rearranjos cromossômicos estruturais via o mecanismo de transposição alternativa (não representada todas as possibilidades). A) *TEs* localizados na mesma cromátide irmã/ou cromossomo homólogo, resultando em uma inversão (CB) flanqueada pelos dois *TEs*. B) *TEs* em sítios distintos nos cromossomos homólogos resultando em uma inversão (CB) e uma deleção (D). C) Mesma conformação que em B, entretanto com o envolvimento das outras regiões terminais dos distintos *TEs*, resultando em uma inversão (CB) e uma duplicação (D). D) *TEs* na mesma cromátide resultando em uma inversão (CB) e uma deleção (D). Duplicação do sítio alvo representada por círculos e quadrados. Sítio de inserção do elemento representada por triângulo. Retirado de Delprat *et al.* (2009).

Esta fita posteriormente realinha com sua fita modelo original, resultando em uma alça de replicação com toda sequência não homóloga obtida por meio das invasões. Esta alça pode escapar dos mecanismos de reparo, e em um ciclo de replicação posterior ter seu conteúdo duplicado. Esse mecanismo pode dar origem a duplicações, ou até mesmo triplicações de genes individuais, com deleções e

inversões associadas (Zhang *et al.*, 2009). Esse processo está implicado com a origem de rearranjos cromossômicos não recorrentes em uma população, e sua descrição foi inferida para a duplicação do gene da proteína proteolípida *PLP1* no cromossomo X em humanos, o qual causa a doença Pelizaeus-Merzbacher caracterizada pela hipomielinização do sistema nervoso central. (Lee *et al.*, 2007).

Rearranjos cromossômicos estruturais, enquanto anormalidades cromossômicas adquiridas referem-se àqueles que surgem durante o desenvolvimento, ou durante a vida do organismo. Suas implicações se restringem somente ao indivíduo afetado. Por sua vez, os rearranjos cromossômicos estruturais constitucionais referem-se àqueles que são herdados de um dos parentais portador, ou ocorrerem via mecanismo “de novo” nos gametas que resultam no zigoto, sendo assim passados para as futuras gerações.

Implicações evolutivas dos Rearranjos cromossômicos estruturais.

Os Rearranjos cromossômicos desempenham um importante papel na evolução de todos os seres vivos, fornecendo variabilidade para a ação da seleção natural. Os rearranjos podem alterar a posição dos genes com relação a seus respectivos elementos reguladores, excluir vários genes ou parte de um único gene (Eichler & Sankoff, 2003).

As taxas de ocorrência são variadas entre as diferentes espécies, e entre os cromossomos sexuais e autossômicos. Para vertebrados estima-se taxas entre 0,2 até 1-2 rearranjos por milhões de anos. Já para invertebrados essa taxa é estimada entre 7 a 50 rearranjos por milhões de anos, uma discrepância altamente influenciada pelo tempo de geração, estratégias de reprodução e história evolutiva das espécies (Eichler & Sankoff, 2003; Coghlan *et al.*, 2005).

A contribuição dos rearranjos cromossômicos estruturais para a especiação dos seres vivos tem sido alvo de debate há mais de cinco décadas. Há uma gama de modelos propondo que rearranjos cromossômicos aceleram a diferenciação genética entre populações e, deste modo, facilitam a especiação (Ayala & Coluzzi,

2005). Dentre estes, dois modelos de especiação cromossômica têm se destacado: o modelo clássico dos híbridos estéreis, e o modelo recombinação-supressão.

No modelo de esterilidade do híbrido, indivíduos heterozigotos para um rearranjo cromossômico são parcialmente ou totalmente inférteis devido a problemas na segregação dos seus gametas, ou à produção de gametas desbalanceados oriundos de eventos de recombinação envolvendo as regiões cromossômicas rearranjadas. Sendo assim, por diminuírem o valor adaptativo dos seus portadores, os rearranjos cromossômicos seriam desvantajosos e a fixação destes em uma população estaria submetida a efeito fundador ou deriva genética em uma população pequena e isolada. Caso ocorra a fixação do rearranjo, a seleção natural agiria favorecendo mutações que reduzam a probabilidade de fluxo gênico entre populações portando diferentes rearranjos cromossômicos, podendo acarretar em isolamento reprodutivo completo.

Já o modelo da recombinação-supressão, de maneira geral, propõe que os rearranjos cromossômicos reduzem o valor adaptativo dos seus portadores devido à supressão da recombinação nas regiões cromossômicas reorganizadas. Neste caso, os gametas recombinantes, por serem portadores de rearranjos cromossômicos deletérios, são eliminados e não passados adiante na prole. Com isto, somente os gametas parentais que não sofreram recombinação seriam fecundados. Assim, as diferentes combinações alélicas envolvidas no rearranjo cromossômico são mantidas, resultando em diferenças genéticas entre as populações, desfavorecendo com o tempo o fluxo gênico e acarretando em isolamento reprodutivo entre estas (Faria & Navarro, 2010; Brown & O'Neill, 2010, Feder *et al.*, 2011).

A divergência genética associada às diferenças cariotípicas foram analisadas em 41 pares de espécies irmãs (que possuem um ancestral comum muito recente em termos evolutivos) das famílias Cricetidae e Muridae de roedores (Castiglia, 2014). Este estudo mostrou uma prevalência de maiores diferenças cariotípicas entre as espécies irmãs vivendo em simpatria, do que as espécies vivendo em alopatria; entretanto sem maior correlação estatística com a divergência genética. O autor conclui que, na ocorrência de um contato secundário,

espécies que divergem cariotipicamente têm uma maior chance de permanecerem distintas do que as espécies que divergem somente geneticamente.

A redução da recombinação associada às inversões foi inferida por Farré *et al.* (2013). Neste estudo os autores analisaram as taxas de recombinação nos cromossomos lineares e rearranjados entre humanos e chimpanzés. As taxas de recombinação eram estatisticamente reduzidas nos cromossomos rearranjados, ao quais eram fixados nas linhagens de humanos e fixados nas linhagens de chimpanzés, garantindo diferenças inerentes de cada espécie.

Entretanto, estabelecer um modelo fixo da atuação dos diferentes rearranjos cromossômicos na especiação das diferentes espécies, parece uma tarefa extremamente rígida frente à grande diversidade em todos os fatores que a história da vida oferece. Jackson (2011) ressalta que as análises futuras devem tomar em conta a falta de exclusividade de um modelo, mas sim a interação dos diferentes mecanismos destes. Outra questão é agregar os fatores históricos biogeográficos em que se passam a origem dos rearranjos cromossômicos e a especiação. Dentre estes fatores, destacam-se as zonas de hibridização, e sua influência no fluxo gênico e diferenciação das espécies (Barton, 2013). A disponibilidade de genomas completos de diferentes organismos traz aos dias atuais, uma mais ampla possibilidade de inferências acerca dos rearranjos cromossômicos estruturais relacionados à especiação.

Referências Bibliográficas Anexo I

Ayala FJ, Coluzzi M (2005) Chromosome speciation: humans, *Drosophila* and mosquitoes. *Proc Natl Acad Sci USA* 102: 6535–6542

Alqallaf AK, Fuad M, Alkoot and Mash'el S. Aldabbous (2013). Discovering the Genetics of Autism, Recent Advances in Autism Spectrum Disorders - Volume I, Prof. Michael Fitzgerald (Ed.), ISBN: 978-953-51-1021-7, InTech, DOI: 10.5772/53797.

Auger DL, Sheridan WF (2011). Plant Chromosomal Deletions, Insertions and Rearrangements. *In: Plant Cytogenetics: Genome Structure and Chromosome Function*. Hass HW and Birchler JA (eds), Springer, New York, 355pp.

Barton NH (20013) Does hybridization influence speciation? *J. Evol. Biol.* 26: 267–269.

Bridges CB (1923) Aberrations in chromosomal materials. *Sc.Pap. Second Intern. Congre. Eugenetic*, 1: 76-81.

Brown JD, O'Neill RJ (2010) Chromosomes, conflict, and epigenetics: chromosomal speciation revisited. *Annu. Rev. Genomics Hum. Genet* 11: 291-316.

Bunting SF, Nussenzweig A (2013) End-joining, translocations and cancer. *Nature Reviews Cancer* 13: 443–454.

Castiglia R (2014) Sympatric sister species in rodents are more chromosomally differentiated than allopatric ones: implications for the role of chromosomal rearrangements in speciation. *Mammal Review* 44: 1–4.

Chikushi H, Doira H, Sakaguchi B (1971). Genetical studies of the deletion, “retarded growth”, in the third chromosome of *Bombyx mori*. *Japan. J. Genetics* 46(5): 301-307.

Coghlan A, Eichler EE, Oliver SG, Paterson AH, Stein L (2005) Chromosome evolution in eukaryotes: a multi-kingdom perspective. *Trends Genet.* 21(12): 673-82. Epub 2005 Oct 19.

Darlington CD, Koller PC (1947) The chemical breakage of chromosomes. *Heredity* 1: 187–221

Decottignies A (2013) Alternative end-joining mechanisms: a historical perspective. *Front. Genet.* <http://dx.doi.org/10.3389/fgene.2013.00048>.

Delprat A, Negre B, Puig M and Ruiz A (2009) The transposon *Galileo* generates natural chromosomal inversions in *Drosophila* by ectopic recombination. *PLoS One* 18; 4(11): e7883.

Deriano L, Roth DB (2013) Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage. *Annu. Rev. Genet* 47: 433–55.

Edelmann L, Spiteri E, Koren K, Pulijaal V, Bialer MG, *et al.* (2001) AT-rich palindromes mediate the constitutional t(11;22) translocation. *Am. J. Hum. Genet.* 68, 1–13.

Eichler EE, Sankoff D (2003) Structural dynamics of eukaryotic chromosome evolution. *Science* 8; 301 (5634): 793-797.

Faria R, Navarro A (2010) Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends Ecol Evol* 25:660–669.

Farré M, Micheletti D, Ruiz-Herrera A (2012) Recombination rates and genomic shuffling in human and chimpanzee—a new twist in the chromosomal speciation theory. *Mol Biol Evol.* 30(4): 853-864.

Feder JL, Gejji R, Powell TH, Nosil P (2011) Adaptive chromosomal divergence driven by mixed geographic mode of evolution. *Evolution* 65(8): 2157-2170.

Feschotte C, Pritham EJ (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet* 41:331-68.

Gisselsson D, Pettersson L, Höglund M, Heidenblad M, Gorunova L, *et al.* (2000) Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity. *Proc Natl Acad Sci U S A.* 9;97(10):5357-5362.

Gray YH (2000) It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements. *Trends Genet* 16: 461–468.

Griffiths AJF, Wessler SR, Lewontin RC, Carroll SB (2008) Mutaç o cromossomica I: Mudanç a na Estrutura dos cromossomos. *In: Griffiths AJF (ed) Introduç o   Gen tica* 9th edition. Guanabara Koogan, Rio de Janeiro, pp 490-503.

Guerra M (1988) *Introduç o   Citogen tica Geral*. Guanabara Koogan, 142p.

Gupta PK (2007) *Cytogenetics*. 1st edition. Rastogi Publications. Meerut, India 374pp.

Haber JE (2000) Repairing a double-strand break. *TIG* 16 (6): 259-264.

Hoang ML, Tan FJ, Lai DC, Celniker SE, Hoskins RA, *et al.* (2010) Competitive repair by naturally dispersed repetitive DNA during non-allelic homologous recombination. *PLoS Genet.* 2; 6(12): e1001228. doi: 10.1371/journal.pgen.1001228.

Jackson BC (2011) Recombination-suppression: how many mechanisms for chromosomal speciation? *Genetica* 139(3): 393-402.

Koshland D, Kasperek TR, Humphrey TC (2011) DNA double-strand break repair pathways, chromosomal rearrangements and cancer. *Seminars in Cell & Developmental Biology* 22: 886–897.

Krishnaswamy L, Zhang J, Peterson T (2008) Reversed end *Ds* element: a novel tool for chromosome engineering in *Arabidopsis*. *Plant Mol Biol* 68: 399–411.

Lam I, Keeney S (2014) Mechanism and regulation of meiotic recombination initiation. *Cold Spring Harb Perspect Biol.*16; 7(1): a016634.

Lee JA, Carvalho CM, Lupski JR (2007) A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell*. 28; 131(7): 1235-47.

Lieber MR (2010) The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End Joining Pathway. *Annu Rev Biochem* 79: 181–211.

Liu P, Carvalho CM, Hastings PJ, Lupski JR (2012) Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev.* 22(3): 211-20. doi: 10.1016/j.gde.2012.02.012.

McClintock B (1941) The stability of broken ends of chromosomes of *Zea mays*. *Genetics*; 23: 234–82.

Macgregor HC (1993) *An Introduction to Animal Cytogenetics*. 1st edition. Chapman & Hall. London, 229pp.

Mandrioli M, Manicardi GC (2012) Unlocking Holocentric Chromosomes: New Perspectives from Comparative and Functional Genomics? *Curr Genomics*. 13(5): 343–349.

Muller HJ (1927) Artificial transmutation of the gene. *Science* 46: 84-87.

Neale JM, Keeney S (2006) Clarifying the mechanics of DNA strand exchange in meiotic recombination. *Nature* 442: 153-158.

Ohno S (1970) *Evolution by gene duplication*. Springer, New York, 160p.

Okuno E (2013) Efeitos biológicos das radiações ionizantes: acidente radiológico de Goiânia. *Estud. av. São Paulo*, v. 27, n. 77, p. 185-200.

Pastwa E, Błasiak J (2003). Non-homologous DNA end joining. *Acta Biochimica Polonica* 50(4): 891-908.

Petes TD (2001) Meiotic recombination hot spots and cold spots. *Nat Rev Genet.* 2(5):360-9.

Schulz-Schaeffer J (1980) *Cytogenetics: Plants, Animals, Humans*. 1st edition. Springer-Verlag, New York-Heilderberg, Berlin, 446pp.

Shrivastav M, De Haro LP, Nickoloff JA (2008) Regulation of DNA double-strand break repair pathway choice Cell Research 18:134-147.

Stankiewicz P, Lupski JR (2002) Genome architecture, rearrangements and genomic disorders. Trends Genet.18(2):74-82.

Stankiewicz P, Lupski JR (2010) Structural Variation in the Human Genome and its Role in Disease. Annu. Rev. Med.v61:437–55.

Sturtevant HÁ (1925) Unequal crossing over at the Bar locus. Genetics 10(2): 117–147.

Weil CF (2009) Too many ends: aberrant transposition. Genes Dev. 1; 23(9): 1032-6.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, *et al.* (2007) A unified classification system for eukaryotic transposable elements. Nat Reviews Genet 8:973-982.

Wolf-Dietrich H, Ehmsen KT, Liu J (2010) Regulation of homologous recombination in eukaryotes. Annu Rev Genet. 44: 113–139.

Xuan YH, Park SJ, Cho JH, Kim TH, Han C (2015) Diverse Chromosomal Rearrangements Induced by *Ac/Ds* at the *OsRLG5* Locus in Regenerated Rice Plants. J. Plant Biol. 58: 318-326.

Yu C, Han F, Zhang J, Birchler J, Peterson T (2012) A transgenic system for generation of transposon *Ac/Ds*-induced chromosome rearrangements in rice. Theor Appl Genet 125:1449–1462.

Zhang F, Khajavi M, Connolly AM, Towne CF, Batish SD, Lupski JR (2009) The DNA replication *FoSTeS/MMBIR* mechanism can generate genomic, genic and exonic complex rearrangements in humans. Nat Genet. 41(7): 849-53.

Zhang J, Peterson T (2004) Transposition of reversed *Ac* element ends generates chromosome rearrangements in maize. Genetics 167(4): 1929-37.

Zhang J, Yu C, Pulletikuri V, Lamb J, Danilova T, *et al.* (2009) Alternative *Ac/Ds* transposition induces major chromosomal rearrangements in maize. Genes Dev 23: 755–76

Anexo II

Tabela Anexo II: Sumário histórico dos estudos sobre cromossomos e polimorfismo cromossômico de *Drosophila willistoni*, e demais espécies crípticas do subgrupo *willistoni* de *Drosophila*.

Assunto	Referência
Descrição do complemento cromossômico da <i>D. willistoni</i>	Metz (1916)
Homologação dos cromossomos de <i>D. willistoni</i> com os de <i>D. melanogaster</i>	Lancefield & Metz (1921)
Análise de cromossomos metafásicos e salivares de várias espécies, incluindo <i>D. willistoni</i>	Wharton (1943)
Referências ao cariótipo de <i>D. willistoni</i>	Sturtevant & Novitski (1941)
	Sturtevant (1942)
	Dobzhansky & Pavan (1943)
	Patterson & Mainland (1944)
	Patterson & Stone (1952)
Descrição detalhada dos cromossomos de <i>D. willistoni</i> : primeiro mapa de referência	Dobzhansky (1950)
Polimorfismo Cromossômico no grupo críptico da <i>D. willistoni</i>	Dobzhansky <i>et al.</i> (1950)
Comparação do polimorfismo cromossômico de espécies crípticas de <i>D. willistoni</i>	Dobzhansky <i>et al.</i> (1950)

Tabela Anexo II: continuação

Adaptabilidade do polimorfismo cromossômico de <i>D. willistoni</i>	Da Cunha <i>et al.</i> (1950)
Genética (polimorfismo e carga genética) de populações marginais de <i>D. willistoni</i>	Townsend (1952)
Inversões nos cromossomos X de <i>D. willistoni</i> e <i>D. paulistorum</i>	Da Cunha (1953)
Polimorfismo cromossômico e variabilidade ambiental em <i>D. willistoni</i>	Da Cunha & Dobzhansky (1954)
Revisão sobre polimorfismo cromossômico em Diptera, com <i>D. willistoni</i> e suas crípticas	Da Cunha (1955)
Polimorfismo cromossômico e fatores ecológicos em <i>D. willistoni</i>	Battaglia & Birch (1956a)
Crossing over em <i>D. willistoni</i>	Battaglia & Birch (1956b)
Adaptação de portadores de diferentes inversões em <i>D. willistoni</i> a diferentes ambientes	Da Cunha (1956a)
Viabilidade diferencial de heterozigotos para inversões em <i>D. willistoni</i>	Da Cunha (1956b)
Diferenças na variabilidade cromossômica de populações experimentais de <i>D. willistoni</i> a diferentes levedos	Da Cunha (1957)
Variabilidade cromossômica em populações <i>D. willistoni</i> de ilhas e continente na América Central e Índias Orientais	Dobzhansky (1957)
Os cromossomos de <i>Drosophila insularis</i>	Dobzhansky <i>et al.</i> (1957)
Heterose e eliminação de homozigotos “fracos” em populações naturais de <i>D. willistoni</i>	Pavan <i>et al.</i> (1957)
Considerações sobre o polimorfismo cromossômico de <i>D. willistoni</i> e <i>D. paulistorum</i>	Dobzhansky & Pavlovsky (1958)
Polimorfismo cromossômico de populações de <i>D. willistoni</i> das Ilhas do Caribe	Townsend (1958)
Alteração experimental do equilíbrio do polimorfismo cromossômico de populações naturais de <i>D. willistoni</i> do sul do Brasil por hibridização com populações do norte do Brasil	Cordeiro <i>et al.</i> (1958)

Tabela Anexo II: continuação

Dados suplementares sobre polimorfismo cromossômico da <i>D. willistoni</i> em relação à variabilidade ambiental	Da Cunha <i>et al.</i> (1959)
Hibridização interracial em populações naturais de <i>D. willistoni</i> , usando o polimorfismo cromossômico como marcador	Cordeiro <i>et al.</i> (1960)
Decréscimo no polimorfismo cromossômico da <i>D. willistoni</i> devido à radiação gama em populações naturais	Cordeiro (1961)
Os cromossomos de <i>Drosophila pavlovskiana</i>	Dobzhansky & Pavlovsky (1962)
Cromossomos de híbridos interespecíficos do grupo críptico da <i>D. willistoni</i>	Cordeiro & Winge (1964)
Polimorfismo cromossômico no complexo de espécies incipientes da <i>D. paulistorum</i>	Kastritsis (1966)
Estudo comparativo dos cromossomos das espécies incipientes do complexo da <i>D. paulistorum</i>	Kastritsis (1967)
Comparação dos cromossomos de <i>D. pavlovskiana</i> e <i>D. insularis</i>	Kastritsis & Dobzhansky (1967)
Recombinação em machos de <i>D. willistoni</i>	França <i>et al.</i> (1968)
Cromossomos de híbridos interespecíficos no grupo críptico da <i>D. willistoni</i>	Cordeiro (1968)
Polimorfismos cromossômico e enzimático em populações de ilhas e continente de <i>D. willistoni</i>	Ayala <i>et al.</i> (1971)
Resposta diferencial de populações de <i>D. willistoni</i> monomórficas e polimórficas para inversões a uma situação de estresse	Tabachnik & Powell (1977)
Descrição de novas inversões de populações do sul do Brasil	Valente & Morales (1985)
Adaptabilidade de portadores de diferentes inversões de <i>D. willistoni</i> à escolha de breeding e feeding sites	Valente & Araújo (1985)
Variação temporal de polimorfismo cromossômico de populações naturais de <i>D. willistoni</i>	Valente & Araújo (1986)

Tabela Anexo II: continuação

Mapeamento <i>in situ</i> do gene de estresse <i>hsp70</i> e puffs de estresse em espécies do subgrupo <i>willistoni</i>	Bonorino <i>et al.</i> (1993a)
	Bonorino <i>et al.</i> (1993b)
Cromossomos de híbridos interespecíficos no grupo da <i>D. willistoni</i>	Cordeiro & Winge (1995)
Primeiro fotomapa e associação de pontos de quebras de inversões e elementos <i>P</i>	Regner <i>et al.</i> (1996)
Mapeamento <i>in situ</i> dos genes <i>Cu/Zn</i> e <i>Sod</i>	Rohde <i>et al.</i> (1994)
Mapeamento <i>in situ</i> dos gene <i>Adh</i>	Rohde <i>et al.</i> (1995)
Fusão dos elementos de Muller E e F em <i>D. willistoni</i>	Papaceit & Juan (1998)
Polimorfismo cromossômico em populações uruguaias de <i>D. willistoni</i>	Valente <i>et al.</i> (2001)
	Valente <i>et al.</i> (2003)
Cromossomos meióticos em <i>D. willistoni</i>	Colares <i>et al.</i> (2003)
Detecção citológica de recombinação em machos	Colares <i>et al.</i> (2004a)
Cromossomos meióticos de fêmeas	Colares <i>et al.</i> (2004b)
Inversão pericêntrica em <i>D. willistoni</i>	Rohde <i>et al.</i> (2005)
Evolução cromossômica no grupo <i>willistoni</i> . Cromossomo IIR	Rohde <i>et al.</i> (2006)
Polimorfismo de <i>D. willistoni</i> em populações mexicanas	Salceda (2006)
Genoma de <i>Drosophila willistoni</i>	<i>Drosophila</i> 12 Genomes Consortium (2007)
Primeiro posicionamento dos <i>scaffolds</i> ao mapa cromossômico. Fotomapa aprimorado de <i>D. willistoni</i>	Schaeffer <i>et al.</i> (2008)

Tabela Anexo II: continuação

Polimorfismo de <i>D. willistoni</i> em populações mexicanas	Salceda (2010)
Mapa de referência aprimorado e de 50 inversões em <i>D. willistoni</i>	Rohde & Valente (2012)
Elemento de Muller F em <i>D. insularis</i>	Powell <i>et al.</i> (2011)
Elemento de Muller F no grupo <i>willistoni</i> de <i>Drosophila</i>	Pita <i>et al.</i> (2014)
Redefinição da posição e orientação dos <i>scaffolds</i> do cromossomo II de <i>D. willistoni</i> – Presente Tese	Garcia <i>et al.</i> (2015)

Referências Bibliográficas Anexo II

Ayala FJ, Powell JR, Dobzhansky T (1971) Polymorphism in continental and island populations of *Drosophila willistoni*. PNAS USA 68: 2480-2483.

Battaglia B, Birch LC (1956a) Polimorfismo cromossômico e fattori ecologici della selezione in *Drosophila willistoni*. Archivio Zoologico Italiano 41: 281-315.

Battaglia B, Birch LC (1956b) Crossing over in *Drosophila willistoni*. Nature 178: 1005.

Bonorino CBC, Silva T, Abdelhay E, Valente VLS (1993a) Heat Shock genes in the *willistoni* group of *Drosophila*: induced puffs and proteins. Cytobios, 73, p. 49-64.

Bonorino CBC, Pereira M, Alonso CEV, Valente VLS, Abdelhay E (1993b). *In situ* mapping of the Hsp 70 locus in seven species of the *willistoni* group of *Drosophila*. Revista Brasileira de Genética. 16(3): 561-571, 1993.

Colares MCS, Valente VLS, Goñi B (2003) The meiotic chromosomes of male *Drosophila willistoni*. Caryologia 56 (4): 431-437.

Colares MCS, Degrandi TH, Valente VLS (2004a). Cytological detection of male recombination in *Drosophila willistoni*. Cytologia 68(4): 359-365.

Colares MCS, Valente VLS (2004b). Meiosis in *Drosophila willistoni* and *D. paulistorum* females. Hereditas 141: 89-93.

Cordeiro AR, Salzano FM, Marques V, Petersen JA (1958) Experimental shift from genetic balance of natural populations of *Drosophila willistoni*. Proc.X Int. Congr.Genet.2: 59.

Cordeiro AR, Salzano FM, Marques VB (1960) An interracial hybridization experiment in natural populations of *Drosophila willistoni*. Heredity 15: 35-45.

Cordeiro AR (1961) Chromosomal polymorphism decrease due to gamma radiation in natural populations of *Drosophila willistoni*. Experientia 17 (405): 1-3.

Cordeiro AR, Winge H (1964) The chromosomes of interspecific hybrids of the *willistoni* cryptic group of *Drosophila*. DIS 39: 100-101.

Cordeiro AR (1968) Chromosomal pairing variability of interspecific hybrids of *Drosophila willistoni* cryptic group. Proc.XII Intern.Congr. Genet. Tokyo 1: 191.

Cordeiro AR, Winge H (1995) Levels of evolutionary divergence of *Drosophila willistoni* sibling species. *In: Genetics of natural populations. The continuing importance of Theodosius Dobzhansky* L. Levine (Ed). Columbia University Press, New York. pp. 262-297.

Da Cunha AB (1953) Chromosomal inversions with sex limited effects. *Nature* 172 (4383): 815-816.

Da Cunha AB (1955) Chromosomal polymorphism in the Diptera. *Advances in Genetics*, 7: 93-138.

Da Cunha AB (1956a) Adaptation of carriers of different chromosomal types in *Drosophila willistoni* to a variety of environments. *Rev.Bras.Biol.* 16(3): 264-272.

Da Cunha AB (1956b) Differential viability favoring inversion heterozygotes in *Drosophila willistoni*. *Evolution* X(3): 231-234.

Da Cunha AB (1957) Contribuição ao estudo da adaptação das populações de *Drosophila* (Diptera) a diferentes levedos. *Bol.Fac.Fil.Cien e Letr. Univ. São Paulo*, 220 *Biol. Geral* (10): 7-56.

Da Cunha AB, Dobzhansky T (1954) A further study of chromosomal polymorphism in *Drosophila willistoni* in its relation to the environment. *Evolution* 8: 119-134.

Dobzhansky T, Pavlovsky O (1958) Interracial hybridization and break down of coadapted gene complexes in *Drosophila willistoni*. *PNAS, USA* 44: 622-629.

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 in *Drosophila willistoni* in its relation to the environment. *Evolution* 13: 389-404.

Dobzhansky T (1950) The chromosomes of *Drosophila willistoni*. *J. Hered.* 41: 156-158.

Dobzhansky T, Pavan C (1943) Studies on Brazilian species of *Drosophila*. *Bol.Fac.Fil. Cien.Letr.Univ.São Paulo. Biologia Geral* 36 (4) :7-72.

Dobzhansky T, Pavlovsky O (1962) A comparative study of the chromosomes in the incipient species of *Drosophila paulistorum* complex. *Chromosoma* 13: 196-218.

Dobzhansky T, Burla H, Da Cunha AB (1950) A comparative study of chromosomal polymorphism in sibling species of the *willistoni* group of *Drosophila*. Amer Natur 84: 229-246.

Dobzhansky T (1957) Genetics of natural populations. XXVI. Chromosomal variability in island and continental populations of *Drosophila willistoni* from Central America and the West Indies. Evolution 11: 280-293.

Dobzhansky T, Erhmann L, Pavlovsky O (1957) Genetics of *Drosophila*. IX. *Drosophila insularis*, a new sibling species of the *willistoni* group. Univ.Texas Publ. 5721: 39-47.

Drosophila 12 Genomes Consortium (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. Nature 480: 2003-218.

França ZM, Da Cunha AB (1968a) Crossing over between heterozygous inversions and its relation with polymorphism in *Drosophila willistoni*. Rev Bras Biol. 28(4): 495-497.

França ZM, Da Cunha AB, Garrido MC. (1968b) Recombination in *Drosophila willistoni*. Heredity 23: 199-204.

Garcia CF, Delprat A, Ruiz A, Valente VLS (2015) Reassignment of *Drosophila willistoni* genome scaffolds to chromosome II arms. G3 (Bethesda) pii: g3.115.021311. doi: 10.1534/g3.115.021311

Kastritsis CD (1966) A comparative chromosome study in the incipient species of the *Drosophila paulistorum* complex. Chromosoma 19: 208-222.

Kastritsis CD (1967) A comparative study of the chromosomal polymorphs in the incipient species of the *Drosophila paulistorum* complex. Chromosoma 23: 180-202.

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

Lancefield RC, Metz CW (1921) Non-disjunction and the chromosome relationships of *Drosophila willistoni*. PNAS USA 7: 225-229.

Metz CW (1916) Chromosome studies in the Diptera. III. Additional types of chromosome groups in the Drosophilidae. Amer.Natur 50: 587-599.

Papacit M, Juan E (1998) Fate of dot chromosome genes in *Drosophila willistoni* and *Scaptodrosophila lebanonensis* determined by *in situ* hybridization. Chromosome Res. 6: 49-54.

Patterson JT, Stone WS (1952) Evolution in the Genus *Drosophila*. The MacMillan Company, N.York, 610p.

Patterson JT, Mainland GB (1944) Studies in the genetics of *Drosophila*. IV. Papers dealing with taxonomy, nutrition, cytology and interspecific hybridization in *Drosophila*. I. The Drosophilidae of Mexico. Univ. Texas Publ. 4445:9-101.

Pavan C, Dobzhansky T, Da Cunha AB (1957) Heterosis and the elimination of weak homozygotes in natural populations of three related species of *Drosophila*. PNAS USA 43: 226-234.

Pita S, Panzera Y, Valente VLS, Melo ZGS, Garcia C, et al. (2014) Cytogenetic mapping of the F element genes in *Drosophila willistoni* group. Genetica 142(5): 397-403.

Powell JR, Dion K, Papaceit M, Aguadé M, Vicario S, Garrick RC (2011) Nonrecombining genes in a recombination environment: the *Drosophila* "dot" chromosome. Mol Biol Evol 28: 825-833.

Rohde C, Valiati VH, Pinto Jr H, Schranck A, Valente VLS (1994). Localization of the Cu/Zn Superoxide Dismutase gene in the *Drosophila willistoni* species group by *in situ* hybridization. Cytobios 80: 193-198

Rohde C, Abdelhay E, Pinto Jr H, Schrank A, Valente VLS (1995) Analysis and *in situ* mapping of the Adh locus in species of the *willistoni* group of *Drosophila*. Cytobios 81: 37-47.

Rohde C, De Toni DC, Degrandi, TH, Valente VLS (2005) *Drosophila willistoni* polytene chromosomes. I. Pericentric inversion on X chromosome. Caryologia 58 (3): 249-254.

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

Rohde C, Valente VLS (2012) Three decades of studies on chromosomal polymorphism of *Drosophila willistoni* and description of fifthly different rearrangements. Genetics and Molecular Biology 35:966-979.

Salceda VM (2006) Chromosomal polymorphism in a natural population of *Drosophila willistoni* (Diptera: Drosophilidae) from Laguna Verde, Vera Cruz, Mexico: eight year record. The Southwestern Naturalist 51 (3): 320-325.

Salceda VM (2010) A prospective study of inversion polymorphism in natural populations of two *Drosophila* species from Eastern Mexico. Genetika 42(3): 407-414.

Schaeffer SW, Bhutkar A, Ma Allister BF 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.

Sturtevant AH, Novitski E (1941) The homologies of the chromosome elements in the Genus *Drosophila*. *Genetics* 26: 517-541.

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

Tabachnick WJ, Powell JR (1977) Adaptive flexibility of “marginal” versus “central” populations of *Drosophila willistoni*. *Evolution* 31(3): 692-695.

Townsend JI (1952) Genetics of marginal populations of *Drosophila willistoni*. *Evolution* 6: 428-442.

Townsend JI (1958) Chromosomal polymorphism in Caribbean Island populations of *Drosophila willistoni*. *PNAS USA* 44: 38-42.

Valente VLS, Morales NB (1985) New inversions and qualitative description of inversion heterozygotes in natural populations of *Drosophila willistoni*. *Revista Brasileira de Genética VIII(1)*: 167-173.

Valente VLS, Araújo AM (1985) Observations on the chromosomal polymorphism of natural populations of *Drosophila willistoni* and its association with the choice of breeding and feeding sites *Revista Brasileira de Genética VIII(2)*: 271-284.

Valente VLS, Araújo AM (1986) Chromosomal polymorphism, climatic factors and variation in population size of *Drosophila willistoni*. *Heredity.* 57(2):149-160.

Valente VLS, Ruzszyk A, Santos RA (1993) Chromosomal polymorphism in urban *Drosophila willistoni*. *Revista Brasileira de Genética* 16(2): 307-319.

Valente VLS, Rohde C, Valiati V, Morales NB, Goñi B (2001) Chromosome inversions occurring in Uruguayan populations of *Drosophila willistoni*. *DIS* 84: 55-59.

Valente VLS, Goñi B, Valiati VH, Rohde C, Morales NB (2003) Chromosomal polymorphism in *Drosophila willistoni* populations from Uruguay. *Genetics and Molecular Biology* 26(2): 163-173.

Wharton LT (1943) Studies in the genetics of *Drosophila*. The Drosophilidae of Southwest. III. Analysis of the metaphase and salivary chromosome morphology within the Genus *Drosophila*. *Univ.Texas Publ* 4313: 282-319.

Winge H (1971) Níveis de divergência evolutiva no grupo críptico da *Drosophila willistoni*. Tese de Doutorado. Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.

Anexo III

Evolução cromossômica em eucariotos – Rearranjos cromossômicos e elementos de transposição

Claudia Marcia Aparecida Carareto
Claudia Barros Monteiro-Vitorello
Marie-Anne Van Sluys
Organizadoras

Elementos de Transposição

Diversidade, evolução, aplicações e
impacto nos genomas dos seres vivos



FUNDAÇÃO OSWALDO CRUZ

Presidente Paulo Gadelha
 Vice-Presidente de Ensino, Informação e Comunicação Nísia Trindade Lima

EDITORA FIOCRUZ

Diretora Nísia Trindade Lima
 Editor Executivo João Carlos Canossa Mendes
 Editores Científicos Carlos Machado de Freitas
 Gilberto Hochman
 Conselho Editorial Claudia Nunes Duarte dos Santos
 Jane Russo
 Lígia Maria Vieira da Silva
 Maria Cecília de Souza Minayo
 Marília Santini de Oliveira
 Moisés Goldbaum
 Pedro Paulo Chieffi
 Ricardo Lourenço de Oliveira
 Ricardo Ventura Santos
 Soraya Vargas Córtes



Editora FioCruz
 Avenida Brasil, 4036, sala 112 | Manguinhos
 CEP 21040-361 | Rio de Janeiro, RJ
 Tels.: (21) 3882-9039 e 3882-9007
 Fax: (21) 3882-9006
 editora@fiocruz.br
 www.fiocruz.br/editora

COMISSÃO EDITORIAL SOCIEDADE BRASILEIRA DE GENÉTICA

Editor Élgion Lúcio Silva Loreto
 Universidade Federal de Santa Maria
 Comissão Editorial Carlos Frederico Martins Menck
 Universidade de São Paulo
 Louis Bernard Klaczko
 Universidade Estadual de Campinas
 Marcio de Castro Silva-Filho
 Universidade de São Paulo
 Maria Cátira Bortolini
 Universidade Federal do Rio Grande do Sul
 Marcelo dos Santos Guerra Filho
 Universidade Federal de Pernambuco
 Pedro Manoel Galetti Junior
 Universidade Federal de São Carlos



Sociedade Brasileira de Genética
 Rua Capitão Adélmio Norberto da Silva, 736 | Alto da Boa Vista
 CEP 14025-670 | Ribeirão Preto, SP
 Tels.: (16) 3621-8540
 contato@sbg.org.br
 www.sbg.org.br

Claudia Marcia Aparecida Carareto
 Claudia Barros Monteiro-Vitorello
 Marie-Anne Van Sluys
 Organizadoras

Elementos de Transposição

Diversidade, evolução, aplicações e
 impacto nos genomas dos seres vivos



Copyright © 2015 dos autores
 Todos os direitos reservados à Sociedade Brasileira de Genética e à Fundação Oswaldo Cruz/ Editora

Catálogo na fonte
 Instituto de Comunicação e Informação Científica e Tecnológica
 Biblioteca de Saúde Pública

C261e Carareto, Claudia Marcia Aparecida (Org.)
 Elementos de transposição: diversidade, evolução,
 aplicações e impacto nos genomas dos seres vivos. /
 organizado por Claudia Marcia Aparecida Carareto,
 Claudia Barros Monteiro-Vitorello e Marie-Anne Van
 Sluys. — Rio de Janeiro : Sociedade Brasileira de
 Genética/Editora FIOCRUZ, 2015.
 196 p. : il. ; tab.
 ISBN 978-85-7541-462-0
 ISBN 978-85-89265-20-1

1. Elementos de DNA Transponíveis. 2. Genoma
 Humano. 3. Cromossomos. 4. Epigênese Genética.
 5. Transferência Genética Horizontal. 6. Marcadores
 Genéticos. 7. Biotecnologia. I. Vitorello, Claudia Barros
 Monteiro (Org.). II. Sluys, Marie-Anne Van (Org.).
 III. Título.

CDD - 22.ed. – 611.018166

Impressão
 Imos Gráfica e Editora Ltda.

Tiragem
 1 mil exemplares

Revisão, Capa, Projeto Gráfico e Diagramação

editora  cubo
 soluções para o universo acadêmico

Sumário

Apresentação	7
Claudia Marcia Aparecida Carareto, Claudia Barros Monteiro-Vitorello, Marie-Anne Van Sluys	
Capítulo 1 Elementos de Transposição: Classificação e Mecanismos de Mobilização	11
Alessandro de Mello Varani, Leonardo Castelo Branco Carvalho, Marcelo Marques Zerillo, Claudia Barros Monteiro-Vitorello	
Capítulo 2 Evolução cromossômica em eucariotos - Rearranjos cromossômicos e elementos de transposição	43
Carolina Flores Garcia, Vera Lúcia S. Valente Gaiesky	
Capítulo 3 O Impacto dos Elementos de Transposição na Evolução de Sequências Reguladoras e Codificantes	65
Elaine Silva Dias, Claudia Marcia Aparecida Carareto	
Capítulo 4 Regulação de Elementos de Transposição: Mecanismos Epigenéticos de Silenciamento, Autorregulação e Ativação por Estresse	91
Adriana Granzotto, Guilherme Marcello Queiroga Cruz	
Capítulo 5 Transferência Horizontal de Elementos de Transposição	115
Mauro de Freitas Ortiz, Gabriel Luz Wallau	
Capítulo 6 Instabilidade Genética e Doenças Humanas	141
Rita Daniela Fernández Medina, Claudia Marcia Aparecida Carareto	
Capítulo 7 Elementos de Transposição, Marcadores Moleculares e Aplicabilidade Biotecnológica	169
Mayra Akemi Kuroki, Nathalia de Setta	
Epílogo	195

1. INTRODUÇÃO

Uma das características marcantes dos cariótipos de eucariotos é a ocorrência de grande quantidade de rearranjos cromossômicos. A existência de polimorfismos intraespecíficos e a adaptabilidade diferencial de portadores de diferentes rearranjos cromossômicos têm sido registradas em organismos tão diversos como invertebrados e vertebrados. Inicialmente atribuídos exclusivamente a quebras ao acaso e/ou em resposta a mutagênicos físicos e químicos, com os avanços das técnicas moleculares e o sequenciamento dos genomas de diferentes espécies foram obtidas evidências diretas do envolvimento dos elementos de transposição na reorganização física dos cromossomos. O presente capítulo abordará, de maneira sucinta, o envolvimento dos elementos de transposição na gênese de diferentes rearranjos cromossômicos no organismo modelo *Drosophila*, em primatas e em plantas.

2. REARRANJOS CROMOSSÔMICOS

Os rearranjos cromossômicos resultam de quebras cromossômicas, as quais podem ocorrer em um único cromossomo e ou em um conjunto de cromossomos, e da subsequente reunião da(s) região(ões) envolvida(s) de maneira reordenada. De acordo com revisões de Lacadena (1996) e de Griffiths et al. (2008), essas reordenações, resumidamente, envolvem:

a) Deleção: Também denominada Deficiência, envolve perda de um segmento cromossômico e, conseqüentemente, a informação genética incluída nele. Origina-se por quebras aleatórias ou induzidas por agentes físicos e químicos; por meio de recombinação desigual entre cromossomos homólogos, entre regiões homólogas repetidas ao longo do cromossomo (figura 1) e da recombinação entre cromossomos, portando ou inversão ou uma translocação. Classificam-se em Deleção terminal, que envolve apenas uma quebra e perda da porção terminal de um cromossomo; e Deleção intersticial, que ocorre em qualquer região entre as duas extremidades do cromossomo. Alelos recessivos deletérios ou até mesmo letais, os quais são compensados por alelos selvagens, encontram uma oportunidade de se expressarem fenotipicamente pela ação da deleção, sendo esse fenômeno denominado Pseudodominância. Junto com a Pseudodominância, a deleção pode afetar genes essenciais aos

organismos. Isso a torna um rearranjo muito deletério, principalmente para os organismos diploides, sendo tanto mais nociva quanto maior for a porção perdida.

b) Duplicação: Repetição de um determinado segmento cromossômico. Esse rearranjo compreende desde a duplicação de um segmento gênico até a duplicação de cromossomos completos (polissomias) e de todo o genoma (poliploidia). Podem ser encontradas in tandem (uma ao lado da outra no mesmo cromossomo) ou deslocadas (em diferentes regiões em um mesmo cromossomo ou em cromossomos distintos). A duplicação de um segmento cromossômico origina-se, como a deleção, por meio da recombinação desigual entre cromossomos homólogos ou entre regiões homólogas repetidas ao longo do cromossomo (figura 1) e da recombinação entre cromossomos portando ou inversão ou translocação. Recombinação desigual envolvendo uma duplicação fornece matéria-prima para a geração de outras duplicações de maior magnitude.

A duplicação gênica tem sido implicada como uma das principais fontes de atuação da evolução na modulação dos genomas. A cópia duplicada muitas vezes não sofre pressão seletiva, e assim sofre mutações de maneira mais rápida do que as demais regiões essenciais do genoma. Isto pode resultar em novas funções gênicas, as quais são consideradas um dos resultados mais importantes dos eventos de duplicação (Ohno, 1970).

c) Inversão: Essa variante origina-se de duas quebras e um posterior giro de 180° da região cromossômica envolvida. Essas quebras podem ser aleatórias ou induzidas por diferentes agentes químicos ou físicos, bem como ser oriunda de recombinação entre regiões homólogas repetidas ao longo do cromossomo (Figura 1). Ao contrário da deleção e da duplicação, a inversão geralmente não altera a quantidade geral do material genético. Portanto, a inversão tende a ser menos deletéria (a não ser que ocorra dentro de um gene com função essencial ou que acarrete também deleções ou duplicações a ela associadas) e a não apresentar anomalia fenotípica. Ela pode ser visualizada em estado heterozigoto (quando somente um dos cromossomos homólogos abriga a inversão, gerando uma alça para propiciar o paramento correto na meiose) ou homozigoto (quando ambos os cromossomos homólogos portam a

inversão, pareando normalmente). As inversões podem ser de dois tipos: paracêntrica (não envolve o centrômero na sua formação, ocorrendo no mesmo braço cromossômico) e pericêntrica (a qual envolve o centrômero e mais de um braço cromossômico).

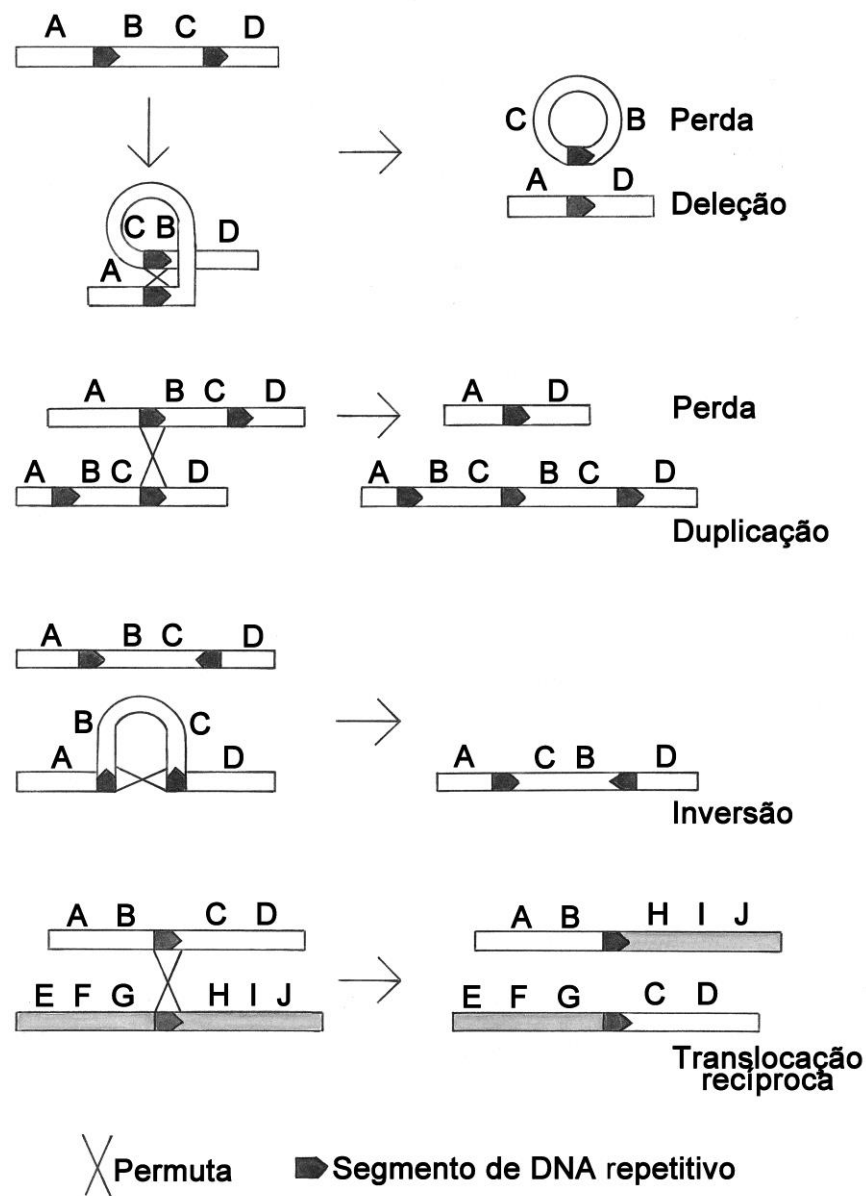


Figura 1. Origem dos diferentes rearranjos cromossômicos através de recombinação entre regiões homólogas repetidas ao longo de um cromossomo, ou cromossomos distintos.

Fonte. Modificado de Griffithis et al. (2008).

A ocorrência de recombinação dentro da alça de uma inversão heterozigota paracêntrica leva à formação de um cromossomo dicêntrico (presença de dois centrômeros) e um fragmento acêntrico (ausência do centrômero), os quais são caracterizados por deleções e duplicações. Já a ocorrência de recombinação em inversões pericêntricas dá origem a deleções e duplicações. Em decorrência disto, a inversão é considerada uma das alterações cromossômicas estruturais com maior poder para restringir recombinação (juntamente com a translocação) e, assim, pode manter grupos de ligação gênica e combinações alélicas por muitas gerações.

d) Translocação (fusões e fissões): Compreende a troca de fragmentos entre dois cromossomos, podendo ser simples (inserção de um segmento cromossômico em outro distinto), recíproca (troca mútua de fragmentos entre dois cromossomos) e “robertsoniana” (fusão entre os braços longos de cromossomos acrocêntricos). Origina-se das quebras aleatórias ou induzidas, bem como da recombinação entre regiões homólogas repetidas ao longo de cromossomos não homólogos (figura 1). Podem acarretar o estabelecimento de novos grupos de ligação gênica e mudanças drásticas na morfologia dos cromossomos.

3. ELEMENTOS DE TRANSPOSIÇÃO E REARRANJOS CROMOSSÔMICOS

Estudos clássicos desde a década de 1940 utilizavam termos como genes instáveis, elementos controladores, variegação, mosaicismo, locos hipermutáveis e efeito de posição para designar o que atualmente sabe-se serem algumas das consequências do envolvimento dos elementos de transposição (do Inglês, TEs) para os genomas hospedeiros. Não foi diferente com o entendimento da gênese de diferentes rearranjos cromossômicos. Sendo assim, esses fenômenos foram sendo revelados por diferentes estudos, antes mesmo da descrição molecular dos TEs. Os estudos iniciais, destacando-se os de Barbara McClintock com milho, por meio de cruzamentos programados e análises citológicas, forneceram a base para o conhecimento da existência de sequências específicas que favorecem as quebras

cromossômicas e assim promovem variabilidade genômica (Veja Capítulo I e Tópico 7 deste capítulo).

O crescente conhecimento molecular acerca de diferentes TEs aponta dois possíveis mecanismos pelos quais rearranjos cromossômicos associados a esses podem originar-se. De acordo com as análises em Gray *et al.* (1996) e Gray (2000), um desses mecanismos ocorre por via indireta de ação do TE, por meio de recombinação homóloga não alélica (do Inglês, Nonallelic Homologous Recombination – NAHR), também denominada recombinação ectópica. Esse mecanismo promove rearranjos cromossômicos devido à presença no genoma de múltiplas cópias muito semelhantes de um TE, se não idênticas, entre as quais podem ocorrer recombinações. A recombinação ectópica entre diferentes cópias de TEs também origina diferentes alterações cromossômicas devido à reparação defeituosa das quebras em fitas duplas formadas durante a excisão do TE, que utiliza sequências molde do próprio cromossomo homólogo para a reparação.

Já o outro mecanismo origina rearranjos cromossômicos por via direta de ação, em decorrência da transposição do TE, muitas vezes acarretada por agentes físicos e químicos. Trata-se de um mecanismo no qual as regiões terminais dos TEs envolvidos são provenientes de elementos distintos (dando origem a um elemento híbrido), ao invés de pertencerem à mesma cópia de um elemento, como na transposição tradicional. Esse mecanismo também é conhecido como transposição aberrante, não convencional ou alternativa, e até agora tem sido descrito para os elementos de Classe II (os quais utilizam um mecanismo corta-e-cola para se transpor, como abordado no capítulo 1 deste livro). Quando os elementos envolvidos se encontram em cromátides irmãs ou em cromossomos homólogos, o mecanismo denomina-se Inserção do Elemento Híbrido (do Inglês, *Hybrid Element Insertion – HEI*), nos estudos em *Drosophila*; e Transposição entre Cromátides-Irmãs (do Inglês, *Sister-Chromatid Transposition – SCT*) nos estudos com milho; já quando o mecanismo envolve elementos que se encontram na mesma cromátide, denomina-se Transposição dos Terminais Invertidos (do Inglês, *Reversed-Ends Transposition – RET*) (Gray *et al.*, 1996; Gray, 2000; Delprat *et al.*, 2009).

Sendo assim, dois TEs podem participar de um evento de transposição alternativa, onde as regiões terminais de TEs separados sinapsam devido à sua complementaridade. O elemento híbrido formado (figura 2) pode se transpor para um novo sítio cromossômico, o qual pode estar até mesmo dentro de outro TE. Os tipos de rearranjos resultantes desse processo dependem do tipo de molécula de DNA envolvida (circular ou linear), da localização do sítio-alvo e da orientação dos TEs envolvidos nos cromossomos parentais. A visualização do resultado do evento de transposição alternativa depende da viabilidade do organismo examinado e de sua capacidade de suportar alterações cromossômicas estruturais (para melhor visualização e compreensão, o estudo de Gray *et al.* (1996) envolvendo *HEI* do elemento *P* em *Drosophila melanogaster* fornece uma ampla explicação sobre esse mecanismo e os possíveis rearranjos cromossômicos resultantes).

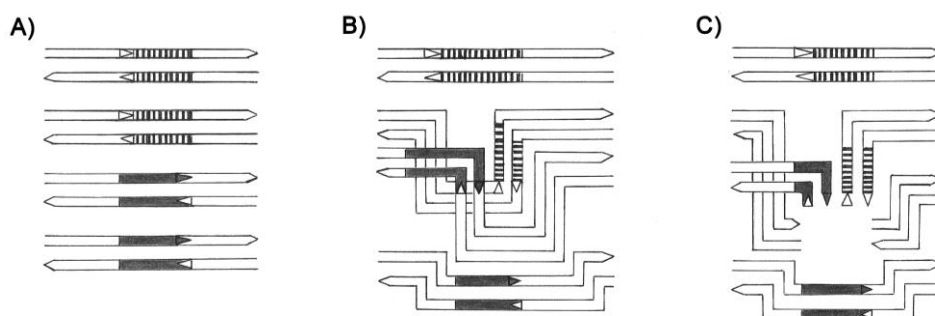


Figura 2. Pareamento da porção terminal de dois TEs distintos, originando um elemento híbrido; a) Cromossomos homólogos duplicados pareados em mitose, antes da definição das células germinativas; b) Sinapse da porção terminal de TEs separados, nesse caso presentes nos cromossomos homólogos; c) Quebra das fitas duplas. A partir daí a transposição alternativa para outro sítio-alvo pode suceder, dando origem a diferentes rearranjos cromossômicos. Setas ao final dos cromossomos indicam região 3' terminal.

Fonte. Modificado de Gray *et al.* (1996).

4. ELEMENTOS DE TRANSPOSIÇÃO E ORIGEM DAS INVERSÕES EM *DROSOPHILA*

O estudo dos diferentes tipos de alterações cromossômicas é bastante consolidado no gênero *Drosophila*, em grande parte devido à facilidade de análise oferecida pelos cromossomos politênicos. Esses cromossomos estão presentes nos núcleos interfásicos das glândulas salivares das larvas, nas células nutridoras (do Inglês, nurse cells) dos oócitos e nos Tubos de Malphigi. Eles são o resultado de um evento de endopoliploidia em cromossomos interfásicos, no qual dois cromossomos homólogos de cada parental permanecem unidos como na meiose e o material genético sofre sucessivos ciclos de replicação sem que ocorra a divisão celular. Isso resulta em uma grande estrutura com bandamento natural (o qual é espécie específico, como um código de barras), formado pela sinapse precisa dos cromômeros paralelos entre cada cromátide irmã.

A estrutura física dos cromossomos politênicos permite a análise precisa das diferentes alterações cromossômicas em *Drosophila*, consagrando suas espécies como organismos modelos para estudos populacionais e evolutivos para os diferentes tipos de rearranjos. A variante cromossômica mais extensivamente estudada nesse gênero é a inversão. Mais da metade das espécies de *Drosophila* estudadas em detalhe são naturalmente polimórficas para inversões paracêntricas em um ou mais braços cromossômicos (Powell, 1997).

A hipótese de Dobzhansky (1970) para explicar o alto polimorfismo cromossômico encontrado no gênero *Drosophila* aborda o maior efeito da heterozigosidade nos genomas, que é o fato de a recombinação ser efetivamente suprimida nos cromossomos heterozigotos para uma inversão. Desse modo, complexos de genes ligados em uma inversão nos diferentes tipos cromossômicos são herdados como blocos e raramente são corrompidos por crossing over; dando base para a seleção natural agir de modo a preservar os complexos de alelos de genes favoráveis de acordo com as necessidades impostas. A adaptabilidade das inversões no gênero *Drosophila* é altamente

estudada e documentada para diversos fatores como: clima, altitude, nicho (Powell, 1997).

A associação dos TEs com diferentes rearranjos cromossômicos no gênero *Drosophila* teve seu início principalmente em estudos feitos em linhagens que apresentavam disgenesia híbrida, síndrome na qual ocorrem altas frequências de prole inviável, recombinação em machos, mutações, alterações cromossômicas e distorção da razão sexual. A causa da disgenesia híbrida tem sido reportada à ativação de diversas famílias de TEs, incluindo elementos *P*, *I* e *hobo* em *Drosophila melanogaster* e *Penelope*, *Ulysses*, *Helena*, *Paris* e *Telemac* em *Drosophila virilis* (revisão em Kidwell, 2005).

Um estudo clássico envolvendo mobilização de *TE* em uma linhagem disgênica foi o de Engels e Preston (1984). Neste, a análise citológica e por hibridação *in situ* de 746 rearranjos gerados no cromossomo X, obtidos por cruzamentos entre linhagens π^2 de *D. melanogaster*, demonstrou que o processo de formação de rearranjos envolve em diferentes casos o ganho, a perda ou a conservação dos elementos *P* nos pontos de quebra.

Já o envolvimento do *TE hobo* com pontos de quebra para inversões foi estudado por Lim (1988) e por Lyttle e Haymer (1992). No primeiro estudo, o autor encontrou o elemento *hobo* em 29 pontos de quebra de 16 diferentes alterações detectadas no cromossomo X designado *Uc* de *D. melanogaster*. Esse cromossomo apresenta instabilidades sítio-específicas peculiares e também a capacidade de formar rearranjos idênticos de maneira recorrente. O autor inferiu que todos os rearranjos podem ser explicados por pareamento intracromossômico e posterior recombinação entre as cópias do elemento *hobo*. Também que as cópias preexistentes do elemento em orientação direta mediam deleções, enquanto que as cópias em orientação invertida mediam inversões.

O segundo trabalho analisou quatro inversões paracêntricas endêmicas de *D. melanogaster* das ilhas Hawai e Oahu. Nas inversões *In(2R)Df891*, *In(2R)Df892* e *In(2R)Ko877*, o elemento *hobo* estava presente nos, ou muito próximos aos, pontos de quebra. Já a inversão *SDKon871* apresentou sítio de inserção de *hobo* em apenas um de seus pontos de quebra. Eventuais inversões cosmopolitas também presentes nas linhagens dessas ilhas não

apresentaram associação entre o elemento *hobo* e seus pontos de quebra. Esse trabalho foi de grande importância, pois foi o primeiro a buscar associação de TEs com pontos de quebra de inversões ocorrendo na natureza.

Nesse contexto, o trabalho de Regner *et al.* (1996) foi também pioneiro nas análises de populações naturais. Nesse estudo, os autores analisaram a coincidência dos sítios de inserção do elemento P com pontos de quebra de inversões em *Drosophila willistoni* (uma das espécies mais polimórficas do gênero), encontrando associação desse elemento com 10 pontos de quebra de inversões paracêntricas segregantes na espécie.

Estudos com as espécies do grupo da *D. virilis* (Evgen'ev *et al.*, 2000) mostraram que a associação dos elementos *Penelope* e *Ulysses* com os pontos de quebra ocorrentes nas espécies do grupo era estatisticamente significativa, com vários sítios preferenciais (do Inglês, *hotspots*) para a inserção desses elementos. Também foram analisados cruzamentos com linhagens disgênicas de *D. virilis* e dos 32 pontos de quebra originados por esse processo, 10 estavam associados ao elemento *Penelope* e quatro ao elemento *Ulysses*.

Apesar das crescentes evidências do envolvimento dos TEs com pontos de quebra de inversões no gênero *Drosophila*, o verdadeiro papel que esses elementos exercem na geração desses polimorfismos permanece ainda por esclarecer, sendo o conhecimento na área baseado em evidências indiretas. Esse quadro começou a mudar após os trabalhos de Cáceres *et al.* (1999), Casals *et al.* (2003) e Delprat *et al.* (2009), os quais mostraram que a origem de diferentes inversões em *Drosophila buzzatii* foi mediada por um TE isolado nessa espécie, pertencente à superfamília P de transposons de Classe II (Marzo *et al.*, 2008), cuja estrutura apresenta repetições terminais invertidas (do Inglês, *Terminal Inverted Repeats* – TIRs muito longas). Esse TE foi denominado *Galileo* pelos autores (Cáceres *et al.*, 1999).

5. TRANSPOSON *GALILEO*: AGENTE MEDIADOR DE REARRANJOS CROMOSSÔMICOS EM *DROSOPHILA*

O trabalho de Cáceres *et al.* (1999) apresentou uma prova sem precedentes sobre a geração de uma inversão devido à mediação de um TE em *D. buzzatii*. Dois arranjos cromossômicos foram molecularmente analisados nessa espécie: o arranjo padrão *2st* (*standard*), e o arranjo *2j*, que se originou de *2st* pela inversão *2j*. O sequenciamento dos pontos de quebra, tanto na linhagem *2st* (denominados *AB* – pontos de quebra proximal e *CD* – pontos de quebra distal) quanto na linhagem com a inversão *2j* (denominados *AC* e *BD* – proximal e distal, respectivamente), evidenciou a presença de sequências (inserções) em ambos os pontos de quebra da inversão *2j* que não estão presentes em cromossomos com arranjo *2st*. Os autores concluíram que essas sequências apresentavam uma região com homologia da ordem de 274 pb, sendo essa região a responsável pela ocorrência de recombinação ectópica entre elas, dando origem à quebra e posterior rearranjo *2j*.

As sequências que estavam presentes nos pontos de quebra da inversão *2j* apresentavam uma duplicação de 7 pb flanqueando cada inserção, bem como a presença de *TIRs* e um padrão de hibridação *in situ* moderadamente repetitivo. Esses achados levaram os autores a propor que se tratava de um TE de Classe II, denominando-o *Galileo*.

Outro exemplo do TE *Galileo* como agente promotor de rearranjos cromossômicos em *D. buzzatii* foi registrado por Casals *et al.* (2003). Nesse estudo, os autores analisaram os pontos de quebra da inversão *2q*, a qual é oriunda de um arranjo *2j*. O sequenciamento dos pontos de quebra tanto da linhagem *2j* (sem a inversão) quanto da linhagem com a inversão *2q⁷* fixada (linhagem *jq⁷-4*) apontou a presença de grandes inserções nos pontos de quebra *AC* (proximal) e *BD* (distal) do cromossomo *2q⁷*. Essas grandes inserções novamente propiciaram um evento de recombinação ectópica entre as cópias do transposon *Galileo* presentes nessas regiões.

A implicação do TE *Galileo* na origem de inversões cromossômicas em populações naturais de *D. buzzatti* foi mostrada mais uma vez no estudo de Delprat *et al.* (2009), no qual os autores isolaram e caracterizaram os pontos de

quebra de outra inversão polimórfica na espécie, $2z3$. Essa inversão originou-se de um cromossomo portando a inversão $2j$, dando origem ao arranjo $2z$. O segmento $2z^3$ abrange cerca de um terço do cromossomo 2 (11 Mpb) e se sobrepõe à inversão $2j$ de modo que as duas inversões não podem ser separadas por recombinação. Seguindo as análises anteriores de pontos de quebra de inversões feitas pelo grupo (Cáceres *et al.*, 1999; Casals *et al.*, 2003), os pontos de quebra distal e proximal foram denominados *AB* e *CD* em cromossomos não invertidos e *AC* e *BD* em cromossomos invertidos. Essa análise apontou a presença de inserções relativamente grandes em *AC* (2.870 pb) e *BD* (4.786 pb) que não estavam presentes nas junções *AB* e *CD* dos cromossomos não invertidos. Novamente, por meio de uma análise elegante, a recombinação entre duas cópias homólogas nas regiões centrais do TE *GalileoN* [neste trabalho TEs *Galileo* foram classificados como uma família com três subfamílias de TEs em *D. buzzatti*: *GalileoG* (*Galileo*), *GalileoN* (*Newton*) e *GalileoK* (*Kepler*)] presentes em ambos os pontos de quebra *AC* e *BD* da inversão cromossômica $2jz3$ foi apontada como a geradora da inversão.

Marzo *et al.* (2008) realizaram buscas *in silico* de sequências similares ao TE *Galileo* no genoma das 12 espécies de *Drosophila* disponíveis (*Drosophila* 12 Genomes Consortium, 2007) à época e encontraram um padrão bastante amplo de distribuição no gênero. A análise da distribuição dos elementos *Galileo* nos genomas sequenciados de *Drosophila* constatou a presença de sequências similares ao TE *Galileo* também em *D. ananassae*, *D. pseudoobscura*, *D. persimilis*, *D. willistoni*, *D. virilis* e *D. mojavensis*.

O envolvimento de elementos *Galileo* na gênese de três inversões polimórficas em *D. buzzatii* forneceu, pois, novas perspectivas para o estudo das origens de inversões em *Drosophila* e, por extensão, em outros organismos. A detecção de cópias quiméricas, conforme encontrado por Marzo *et al.* (2008), no genoma de *D. mojavensis*, *D. pseudoobscura* e *D. willistoni* indica que o elemento *Galileo* também pode estar envolvido na origem de inversões em outras espécies do gênero *Drosophila*.

A disponibilização do genoma sequenciado de diferentes espécies de *Drosophila* (*Drosophila* 12 Genomes, 2007) tem favorecido os estudos dos pontos de quebra de inversões, tanto fixadas quanto polimórficas; intra e

interespecíficas. Atualmente, tem-se conhecimento do envolvimento dos TEs na origem das inversões através de recombinação ectópica para outras espécies do gênero, como visto no estudo dos pontos de quebra da inversão *In(4)a* em *D. americana*. Presente no cromossomo 4 dessa espécie, que por mecanismo de fusão com o cromossomo X originou um braço neo-X, essa inversão tem sua origem na recombinação entre um elemento do tipo MITE denominado *DAIBAM* (Evans *et al.*, 2007). Esse mesmo elemento também está envolvido na gênese da inversão polimórfica *In(5)a* em *D. americana*, e também na da inversão fixada *In(X)a* que diferencia o cromossomo X entre *D. americana* e *D. virillis* (Fonseca *et al.*, 2012).

Outro caso envolve o elemento do tipo MITE *BuT5*, o qual está implicado na origem da inversão fixada *2s* em *D. mojavensis*. Através de análises *in silico*, os autores constataram que a inversão e o TE podem estar relacionados ao aumento do nível de expressão do gene *DmojlCG10375*, um inferido membro da família *Hsp40*, oferecendo maior termotolerância a essa espécie habitante do deserto de Sonora, localizado na fronteira entre os Estados Unidos e o México (Guillén e Ruiz, 2012).

6. RETROELEMENTOS *LINES*, *SINES* E *SVA* NA ORIGEM DE REARRANJOS CROMOSSÔMICOS EM PRIMATAS

Diversos estudos de comparações genômicas *in silico* entre humanos e chimpanzés têm demonstrado uma forte relação entre rearranjos cromossômicos e DNA repetitivo, esses variando desde aglomerados de genes de RNAr e RNAt até diversos TEs.

Os retroelementos *L1* (do tipo *LINEs* – do Inglês, *Long Interspersed Nuclear Elements*, com aproximadamente 6 kpb) e *Alu* (do tipo *SINE* – do Inglês, *Short Interspersed Nuclear Element*, com aproximadamente 300 pb) são anteriores à radiação humano-chimpanzé (entre 4 a 8 milhões de anos), já existindo no genoma de mamíferos desde aproximadamente 100 milhões de anos (Batzler e Deininger, 2002).

Retroelementos *L1* constituem aproximadamente 17% do genoma humano. Embora muitos *L1* sejam inativos devido ao acúmulo de mutações,

alguns elementos completos permanecem ativos nas células germinativas, e no decorrer do tempo evolutivo esse TE não somente garantiu a sua propagação nos genomas, mas também a de outros elementos não autônomos, tais como *Alu* e *SVA* (Lee *et al.*, 2012).

Um caso bastante peculiar envolvendo o TE *L1* [analisado por Schwartz *et al.* (1998) antes mesmo da disponibilidade do genoma humano] é o do bloco genômico no cromossomo X denominado Xq21, o qual é homólogo com o braço cromossômico Yp. Esse bloco encontra-se na região estritamente ligada ao sexo, a qual não sofre recombinação na meiose, e se originou por meio da transposição de sequências do cromossomo X para o Y, após a radiação chimpanzé-humano, porém anteriormente ao surgimento do *Homo sapiens*. Entretanto, enquanto as sequências Xq21 homólogas ao cromossomo Y formam um único bloco contínuo, as sequências correspondentes em Yp são divididas em dois segmentos, sendo um intervalo distal com aproximadamente 4 kpb (denominado bloco1-2-3) e um proximal com várias centenas de kpb (denominado intervalo 4A), as quais são separadas por vários megabases de DNA específicos do cromossomo Y. Uma questão-chave nessa análise, além da repartição desse segmento no cromossomo Y, é a grande diferença na ordem de segmentos homólogos. Em Yp, a ordem do segmento (proximal para distal) é de 4A-3-21, enquanto que em Xq21 a ordem do segmento homólogo (proximal para distal) é 4A-1-2-3, um rearranjo explicado por um único evento de transposição do segmento homólogo via X-Y, seguida de uma inversão que tenha separado o segmento 4A do bloco 1-2-3.

Análises de relógio molecular de 27 STSs (do Inglês, *Sequence-Tagged Sites*) inferiram níveis similares de divergência entre as sequências de Xq21 e Yp, o que corrobora um único evento de transposição. Já, o sequenciamento da junção entre a região 1 do bloco 1-2-3 e da junção do bloco 4A em Y, bem como os limites do bloco homólogo em X e a comparação do mapa de restrição dessas sequências permitiram a determinação dos pontos de quebra da inversão no cromossomo Y. O alinhamento dessas três sequências prontamente apontou o compartilhamento de uma repetição de 9 pb (CATTATTCT). A região 5' da repetição (CATTATTCT) era quase idêntica entre

os cromossomos X e Y, e o mesmo para a região 3'. Essas repetições correspondem a segmentos do TE *L1*.

As análises mostraram que a inversão em Yp foi mediada pela recombinação ectópica entre cópias idênticas ou muito similares dessas repetições presentes nos limites distal e proximal em Yp (Schwartz *et al.*, 1998).

No que concerne ao genoma humano, sabe-se que os elementos *Alu* perfazem mais de 10% dele. Esses consistem de um monômero esquerdo e um monômero direito, os quais supostamente evoluíram do 7SL-RNA de maneira independente, fusionando-se subsequentemente em um dímero na linhagem que originou os primatas. Esses TEs proporcionam muitas oportunidades para recombinação homóloga, tanto intracromossômica, resultando na supressão ou duplicação de éxons em um gene, quanto intercromossômica, causando anormalidades cromossômicas mais complexas. Por não possuírem a capacidade de se transpor (com a exceção de algumas subfamílias), dependem da maquinaria enzimática de elementos autônomos, geralmente da família de retrotransposons *L1* (Chisholm *et al.*, 2012).

O genoma humano contém numerosos blocos de sequências homólogas denominadas duplicações segmentares (do Inglês, *Segmental Duplications* – *SDs*). Essas proporcionam substrato para recombinação e recorrente formação de rearranjos cromossômicos. Devido a isso, esses segmentos têm sido implicados não apenas na reorganização do genoma humano, mas também dos genomas de outros primatas (Stankiewicz *et al.*, 2012).

Análises *in silico* dos pontos de quebra das *SDs* no genoma humano total disponível evidenciaram um enriquecimento estatisticamente significativo do retrotransposon *Alu* próximo ou entre essas quebras. Esses resultados são consistentes com a recombinação homóloga *Alu-Alu* envolvida na origem e aumento das *SDs* intra e intercromossômicas, embora outros dados apontem que o aumento do número de cópias de *SDs* por meio desse mecanismo ocorre somente em 8% dos casos (Kehrer-Sawatzki e Cooper, 2008).

Eventos de recombinações ectópicas *L1-L1* e *Alu-Alu* que originam inversões (do Inglês, *Retrotransposon Recombination-Mediated Inversion* –

RRMI) têm sido implicados na origem de 49 inversões entre cromossomos humanos e de chimpanzés (sendo 27 inversões específicas de humanos e 22 de chimpanzés), de um total de 252 inversões claramente caracterizadas entre esses organismos. Dessas 49, 28 estavam relacionadas com a presença do TE *Alu* em ambos os pontos de quebra da inversão, e 21 com o TE *L1*. Interessantemente, em 63 das inversões analisadas ocorria a presença de *Alu* (22) e de *L1* (41) em somente um dos pontos de quebra do rearranjo, sendo que a recombinação pode ter ocorrido com outra sequência muito similar do próprio genoma e não com outra cópia de TE. Junto com esses resultados foi encontrado que 26 dessas inversões ocorriam em regiões gênicas e 23 em regiões intergênicas, destacando-se a mediação dos TEs na origem da variação e na evolução dos genomas (Lee *et al.*, 2008).

Os mesmos retroelementos *L1* e *Alu* originam deleções no genoma de seus hospedeiros por meio dos seus mecanismos de transposição e também via recombinação ectópica intra e intercromossômica. Deleções desse tipo são denominadas *L1RAD* (do Inglês, *L1 Recombination-Associated Deletion*) e *ARMD* (do Inglês, *Alu Recombination-Mediated Deletion*), respectivamente. Deleções oriundas desses mecanismos são bastante frequentes em diversos tipos de cânceres e desordens genéticas. Entretanto, análises comparativas *in silico* entre os genomas de humanos e chimpanzés apontam a contribuição dessas deleções também para a diferenciação dos genomas desses primatas (Han *et al.*, 2005).

O envolvimento de eventos *L1RAD* na geração de rearranjos cromossômicos em humanos foi analisado em 73 casos obtidos pela comparação *in silico* dos genomas de humanos e chimpanzés. Desses 73 eventos de *L1RAD*, 55 se originaram por recombinação ectópica *L1-L1*. Os demais eventos de *L1RAD* originaram-se por *NHEJ* (do Inglês, *Nonhomologous End Joining*), o qual consiste de um mecanismo de reparo da quebra da fita dupla que utiliza micro-homologias nas regiões flangeadoras da quebra para uni-las e reparar o dano. Essas micro-homologias foram geradas pelos elementos *L1* nessas regiões, implicando na atuação desse TE na manutenção do genoma. A porção do genoma deletada em humanos por *L1RAD* desde a sua divergência do chimpanzé foi de aproximadamente 447.567 pb. As

deleções tiveram uma média de 6.132 pb e uma mediana de 3.239 pb, caracterizando-se como deleções grandes. Dentre os elementos *L1* envolvidos nos eventos de *L1RAD*, somente 22 correspondiam a cópias completas dele, com o restante sendo defectivas. Eventos *L1RAD*, de uma maneira geral, caracterizaram-se por ocorrerem em regiões pobres de conteúdo gênico em humanos (Han *et al.*, 2005).

Além das deleções no genoma de primatas ocasionadas por intermédio de *L1RADs*, elementos *L1* também estão implicados na origem deste rearranjo cromossômico por meio do próprio evento de transposição, um mecanismo denominado *L1IMD* (do Inglês, *L1 Insertion-Mediated Deletions*). Por meio desse mecanismo estima-se que aproximadamente 17.671 pb foram deletados do genoma humano e aproximadamente 14.921 pb do genoma do chimpanzé desde a sua divergência (Han *et al.*, 2005).

A análise de 492 eventos de *ARMD* que se sucederam após a radiação chimpanzé-humano estimou a deleção de aproximadamente 396.420 pb no genoma humano por esse processo. A maioria dos casos envolveu recombinação *Alu-Alu* com elementos na mesma orientação e em cromossomos homólogos. A média das *ARMDs* foi 806 pb, mas houve um desvio (75% das deleções) para eventos envolvendo *ARMDs* com a mediana de 468 pb sendo a melhor representante do tamanho das deleções. Cerca de 45% dos eventos de *ARMD* estavam localizados em regiões gênicas, de acordo com a base de dados de genes humanos *RefSeq*, e 15% em regiões intergênicas, deletando aproximadamente 229.205 pb de sequências intrônicas do genoma humano. Essa peculiaridade levou à análise das sequências pré-*ARMD* presentes no genoma de chimpanzés, nas quais se verificou em três casos a deleção de um éxon inteiro dos genes *LOC471177* e *LOC452742* de chimpanzés, os quais são ortólogos dos genes *CHRNA9*, *LOC440141* e estão relacionados, em humanos, à audição e à proteína *MRPS31* (do Inglês, *Mitochondrial Ribosomal Protein S31*), respectivamente, e *LOC471116*, uma proteína hipotética em chimpanzés. O estabelecimento de uma sequência consenso dos elementos *Alu* envolvidos em *ARMD* permitiu traçar a distribuição da frequência dos pontos de quebra originados por esse mecanismo em diferentes posições ao longo da sequência consenso, o que

revelou a presença de um *hotspot* para recombinação na porção de 21 a 48 (27 pb) do elemento (Sen *et al.*, 2006).

Da mesma forma, no genoma de chimpanzés comparado ao de humanos, 663 eventos de deleção foram atribuídos à *ARMD*. Isso acarretou a perda de aproximadamente 771.497 pb do genoma, porém considerando a média de divergência por deleções/inserções (*indels*) entre chimpanzés e humanos, a quantidade precisa de DNA deletado por meio do mecanismo *ARMD* no genoma de chimpanzés seria entre 733 e 811 kpb. O tamanho médio das deleções é de 1.164 pb, maior que a média das deleções no genoma humano e, como nesse, também foi encontrado um desvio para deleções com menos de 1 kpb (68% dos casos). Um achado notório foi o de que 10 eventos de *ARMDs* específicos de chimpanzés deletaram cada um mais de 7,3 kpb, sendo que o maior evento envolveu a deleção de 31.861 pb. Através da análise dos genes ortólogos pôde-se verificar que em 46% dos casos de *ARMD* as deleções ocorreram em genes (supostos ou conhecidos, de acordo com a base *RefSeq*) e cinco eventos geraram 13 deleções exônicas em seis genes possivelmente funcionais no genoma humano (Han *et al.*, 2007).

Dentre os 663 eventos detectados no estudo anterior, foram encontrados quatro tipos de recombinação *Alu-Alu*. O primeiro tipo consiste na recombinação entre o mesmo monômero de dois elementos *Alu* separados. O segundo tipo envolve dois elementos *Alu*, sendo que um elemento inseriu-se previamente no meio do outro elemento. O terceiro tipo consiste da recombinação entre um monômero esquerdo e um monômero direito de elementos *Alu* diferentes. Já o último tipo consiste na recombinação entre elementos *Alu* em orientação oposta. A análise dos dois primeiros tipos permitiu a inferência de uma sequência consenso de 22 pb (5'-TGTAATCCCAGCACTTTGGGAGG-3'), a qual seria um *hotspot* para o evento de recombinação *Alu-Alu* e subsequente quebra cromossômica. Esse *hotspot* se apresentou como a região mais conservada entre os elementos *Alu* envolvidos em *ARMDs* no genoma de chimpanzés (Han *et al.*, 2007).

Com relação ao mecanismo *ARD* (do Inglês, *Alu Retrotransposition-Mediated Deletion*), o qual origina deleções resultantes do próprio evento de transposição, o estudo feito por Callinan *et al.* (2005) estimou através de uma

ampla abordagem computacional e molecular do genomas de chimpanzés e humanos a ocorrência de aproximadamente 3 mil eventos *ARD* ao longo da evolução entre esses primatas, resultando na deleção de aproximadamente 900 kpb.

O número de cópias de retroelementos *Alu* é maior do que o número de retroelementos *L1*. Sendo assim, a distância média entre cada retroelemento *Alu* é menor (uma inserção a cada aproximadamente 3 kpb) do que entre cada retroelemento *L1* (uma inserção a cada 6 kpb). Isso se reflete na maior prevalência de eventos *ARMDs* no genoma humano, associado ao fato de que retroelementos *Alu* são dímeros, o que resulta numa gama maior de recombinações possíveis. Apesar de os eventos *ARMDs* serem prevalentes no genoma humano, eventos *L1RAD* são responsáveis pela deleção de uma parcela genômica maior (aproximadamente 450 kpb) do que os demais eventos *L1IMD*, *ARMDs* e *ARD* juntos (aproximadamente 427 kpb). Com isto, o quadro que emerge da análise destes genomas é que eventos *L1RAD* abrangem porções extensas do genoma, o que resulta em menos eventos de recombinação e maiores deleções quando elas ocorrem (Callinan *et al.*, 2005; Han *et al.*, 2005; 2007; Sen *et al.*, 2006).

De maneira geral, o que se verificou nas três análises acima é que as frações deletadas do genoma são, principalmente, sequências repetitivas. Sendo assim, retroelementos *L1* e *Alu* estariam envolvidos na expansão dos genomas, bem como na manutenção do seu tamanho e do próprio número de cópias.

Por sua vez, elementos *SVA* são atualmente ativos, porém, assim como os elementos *Alu*, também dependem da maquinaria enzimática dos elementos *L1* para a transposição. Esse retrotransposon é composto por sequências *SINE-R* (um proposto retrotransposon derivado do retrovírus endógeno humano *HERV-K10*), *VNTR* (do Inglês, *Variable Number Tandem Repeat*) e *Alu* (sequência *Alu-like*). Comparado aos retroelementos *L1* e *Alu*, *SVA* ocorre em menor quantidade no genoma humano, sendo caracterizado como a mais nova família de retroelementos entre primatas (20 milhões de anos) com a ocorrência de eventos recentes de retrotransposição no genoma humano (Wang *et al.*, 2005). Com base nessas características, análises de eventos

SRADs (do Inglês, *SVA Recombination-Associated Deletions*) e *SIMDs* (do Inglês, *SVA Insertion-Mediated Deletions*) foram feitas a fim de se verificar o quanto esse elemento tem contribuído para a reestruturação do genoma humano. Com isso foram inferidos 13 eventos de *SRADs* e 13 eventos de *SIMDs*, acarretando a deleção de aproximadamente 46.537 pb no genoma humano desde a divergência com chimpanzé. Dentre os eventos de *SRADs*, somente um foi devido à recombinação entre a cauda poli-A de elementos *SVA*, os demais envolveram *NHEJ* (Lee *et al.*, 2012). Esse resultado foi atribuído à maior divergência entre as cópias de elementos *SVA* do genoma, bem como a sua baixa densidade, o que implica cópias mais distantes umas das outras para recombinarem-se. Interessantemente, um evento de *SIMD* estava envolvido na deleção de dois éxons do gene *tMDC II* em humanos, tornando-os não funcionais em comparação ao seu órtologo em chimpanzés. Essa alteração resultou em modificações quanto à fertilidade em humanos, uma vez que esse gene está envolvido na interação do espermatozoide com o óvulo antes da fecundação ocorrer (Lee *et al.*, 2012).

7. REARRANJOS CROMOSSÔMICOS EM MILHO (ZEA MAYS) INDUZIDOS PELO SISTEMA AC/DS

Os estudos dos locos *Ac/Ds* na década de 1940 por Bárbara McClintock foram de uma importância imensurável para a ciência da vida, pois foi a partir deles que os TEs e sua capacidade de se mover no genoma foram descobertos (veja capítulo 1).

Uma das primeiras inferências de McClintock para a origem da variação da cor do grão no milho envolvia a quebra cromossômica que ocorria repetidas vezes em um loco no braço curto do cromossomo 9, eventualmente se translocando para outro sítio do genoma. Esse loco foi denominado *Ds* (do Inglês, *Dissociation*), o qual se encontrava no local da quebra; entretanto ele necessitava da presença de outro elemento genético não ligado, denominado *Ac* (do Inglês, *Activator*) o qual “ativava” o elemento *Ds* para quebrar a dupla fita (capítulo 1; revisão em Kass e Chomet, 2009).

Subsequente à quebra ocorria o ciclo de quebra-fusão-ponte (do Inglês,

Breakage-Fusion-Bridge Cycle). Esse ciclo se origina com um cromossomo dicêntrico formando uma ponte, na qual cada centrômero é puxado em direção aos pólos distintos de uma só vez durante a anáfase. Isso resulta na quebra desse cromossomo dicêntrico, e as extremidades quebradas permanecem adesivas, fusionando-se subsequentemente à duplicação na prófase mitótica das células filhas. O resultado é o surgimento de outro cromossomo dicêntrico na anáfase, dando continuidade ao ciclo, com os cromossomos quebrando novamente a cada mitose. Uma vez que a cada ciclo a quebra provavelmente ocorre em um local diferente da quebra anterior, acarretará em um reagrupamento dos locos genéticos com deleções e duplicações gênicas (McClintock, 1941).

Eventos de transposição do loco *Ds* estavam associados intrinsecamente à geração de rearranjos cromossômicos (de acordo com McClintock, *gross chromosomal rearrangements*). Foi só a partir da década de 1980 que as bases moleculares dos TEs *Ac/Ds* e do mecanismo de transposição envolvido na origem desses rearranjos cromossômicos começaram a ser elucidados.

O elemento *Ac* pertence à família de transposons de DNA *hAT* (*hobo*, *Ac* e *Tam3*) e compreende 4,6 kpb, com uma *TIR* de 11 pb imperfeita e *TSD* de 8 pb. Este TE de Classe II gera um único mRNA de 3,5 kpb que codifica uma transposase de 102 kD (Zhang *et al.*, 2011).

Já elementos *Ds* apresentam alta heterogeneidade na sua estrutura, sendo que na maior parte dos casos constituem elementos derivados de deleções internas de *Ac*; porém, em outros casos sua sequência pode ser virtualmente não correlacionada com *Ac*, exceto pela presença de *TIRs* de 11 pb conservadas, as quais são comuns a todos os elementos *Ds*. Outra característica comum aos elementos *Ac* e *Ds* é a presença de múltiplas cópias de motivos de 6 pb (AAACGG ou similares) localizados na região de 250 pb terminal do elemento (Zhang *et al.*, 2011; Lazarow *et al.*, 2013).

Nem todos os elementos *Ds*, os quais necessitam da maquinaria de *Ac* para transpor, são agentes causadores de quebras cromossômicas. Diversos estudos têm identificado que os elementos *Ds* que originam quebra cromossômica são do tipo duplo *Ds*, isto é, o elemento é constituído por duas

cópias idênticas de um elemento *Ds* com 2.041 pb (gerado por uma deleção interna do elemento *Ac*). Uma dessas cópias é inserida, em orientação invertida, dentro da outra cópia. O duplo *Ds* tem quatro repetições das *TIRs*, em duplas com orientação direta e invertida. Todas as demais sequências do elemento ocorrem duas vezes em orientação invertida uma em relação à outra. Duplo *Ds* são também denominados *Ds* de Classe I, enquanto que elementos não duplo *Ds* são de Classe II (Zhang *et al.*, 2011).

Dentre os locos analisados por McClintock, o alelo *sh-m5933* foi abordado molecularmente. Esse alelo é um tipo recessivo instável do gene *Sh* (do Inglês, *Shrunken*) envolvido na origem da enzima *sucrose synthase*, localizado no braço curto do cromossomo 9, o qual apresenta uma estrutura complexa. Mutações nesse gene conferem o desenvolvimento de um endosperma defeituoso e o grão apresenta um aspecto enrugado e pequeno. Análises comparativas dos alelos mutantes *sh-m5933* e dos alelos com a mutação revertida apontaram que o alelo mutante *sh-m5933* contém uma inserção de 30 kpb que interrompe a região codificante do gene *Sh*. Essa inserção consiste de uma duplicação invertida e apresenta suas porções terminais semelhantes em uma região de 8 kpb, a qual foi inferida como *TIRs* imperfeitas. Além disso, ocorre uma duplicação que abrange sequências da inserção e da região 5' do próprio loco no mesmo cromossomo (Courage-Tebbe *et al.*, 1983). A inserção contém, adjacente à região 3' do gene *Sh*, um elemento duplo *Ds*, e adjacente à região 5' do gene *Sh*, um elemento *Ds* defectivo, sem a metade de um dos seus elementos *Ds* constituintes. Já o loco *Sh* nas linhagens com a mutação revertida se apresentava estruturalmente normal, sem a presença da inserção, porém mantinha a duplicação, com um *Ds* defectivo (Fedoroff, 1989).

A origem da duplicação invertida do alelo *sh-m5933* foi inferida por Fedoroff (1989) como devida à fusão das cromátides irmãs, durante a replicação. Esse evento envolveu a sinapse das porções terminais dos elementos *Ds*, os quais estavam na mesma orientação em cada cromátide irmã, seguida da transposição alternativa e *HEI* (Peterson e Zhang, 2013; figura 2).

Diversos estudos têm mostrado o envolvimento do sistema *Ac/Ds* na origem dos diferentes rearranjos cromossômicos e os efeitos fenotípicos resultantes. Dentre esses, destacam-se os que envolvem o gene *p1* (do Inglês, *pericarp color 1*). Esse gene codifica um fator de transcrição Myb e regula a síntese de pigmentos flavonoides vermelhos em diversos órgãos florais do milho, incluindo o sabugo e o pericarpo. Distintos alelos mutantes desse gene têm sido caracterizados, dentre eles os alelos *P1-rr* (pericarpo e sabugo vermelhos), *P1-wr* (pericarpo branco e sabugo vermelho), *P1-ww* (pericarpo e sabugo brancos), *P1-ovov* (pericarpo e sabugo laranja variegado), *P1-oo* (pericarpo e sabugo laranja) e *p1-vv* (pericarpo e sabugo variegados) (Yu *et al.*, 2011).

O isolamento de diferentes alelos oriundos de *P1-rr11* resultou na obtenção de 10 novos alelos *P1ww* e dois novos alelos *p1vv*, os quais são fenotipicamente instáveis. O alelo *P1-rr11* caracteriza-se pela presença de um elemento *Ac* completo, estando a aproximadamente 9 kpb antes da região transcrita do gene e de um elemento *Ac* terminalmente defeutivo (do Inglês, *fractured Ac - fAc*), o qual é um elemento *Ds*, no íntron II do gene *p1*, separados por um segmento de 13 kpb (do Inglês, *Intertransposon Segment – ITS*) e em orientação invertida. As análises por meio de enzimas de restrição desses 12 novos alelos obtidos mostraram que em 10 alelos *P1ww* a sequência de 13 kpb entre os elementos *Ac* estava deletada, já os dois novos alelos *p1vv* apresentavam bandas de tamanho alterado. A análise molecular resultou na caracterização de três alelos distintos genotipicamente com relação a *P1-rr11:p1-ww21* com o fragmento de 13 kpb deletado; *p1-vv22* com inversões e permutação da sequência em *ITS*; e *p1-vv30* com permutação das sequências em *ITS* distintas de *p1-ww21*. Posteriormente, uma nova categoria de rearranjo foi elucidada, a qual envolvia permutas, inversões e deleções ocorrentes na região *ITS*, denominadas rearranjos *ITS*. O modelo proposto para a origem dos rearranjos comentados acima foi a transposição alternativa envolvendo as porções terminais 5' de *Ac* e 3' de *fAc*, que estão na mesma cromátide, o que resulta em deleções (com a possível geração de um cromossomo em anel ou permanência como um fragmento acêntrico) ou rearranjo da região promotora, éxon 1 e éxon 2 do gene *p1*, dependendo da

orientação da inserção das porções terminais dos TEs e, com isso, originando setores sem coloração. Uma característica que reforça esse mecanismo na origem desses rearranjos é a presença de *TSDs* nos pontos de quebra dos rearranjos formados (Yu *et al.*, 2011).

O gene *p2*, um parálogo do gene *p1*, está localizado, aproximadamente, 60 kpb mais proximalmente que *p1*, o qual não é expresso em células do pericarpo e assim não contribui para a sua coloração. Análises moleculares de quatro alelos com pericarpo e sabugo laranja isolados da linhagem *P1-rr11* (*P-oo32*, *P-oo1062*, *P-oo1067* e *P-oo1068*) mostraram que a origem desse fenótipo deve-se a eventos de transposição alternativa envolvendo as porções terminais de *Ac* e *fAc* no gene *p1*. A inserção do elemento híbrido no íntron 2 do gene *p2* resulta na junção das regiões promotoras éxon 1 e éxon 2 do gene *p2* com o éxon 3 do gene *p1*. Como resultado ocorre a deleção dos éxons 3 do gene *p2* e 1 e 2 do gene *p1*. A partir desse rearranjo cromossômico, o padrão de expressão de ambos os genes é alterado, resultando em um novo fenótipo para a pigmentação do pericarpo. Um ponto muito instigante dessa análise é que o evento exposto acima, o qual envolve a transposição alternativa de elementos na mesma cromátide, é análogo ao rearranjo somático V(D)J que ocorre nos genes da imunoglobulina em vertebrados, uma vez que ocorre o “embaralhamento” de sequências gênicas intracromossômicas, dando origem a novas configurações e fenótipos (Zhang *et al.*, 2006).

Transposição alternativa envolvendo a porção terminal de elementos que se encontram na mesma cromátide fornecem explicações para a origem de deleções e inversões e também translocações recíprocas, que são passadas para as próximas gerações, como inferido por Zhang *et al.* (2011). Nesse caso porção 5' de *Ac* fundiona-se com a porção distal do sítio-alvo e a região 5' de *fAc*, com a porção proximal do sítio-alvo, gerando uma translocação recíproca. Elementos *Ac/Ds* apresentam um viés para inserções em regiões gênicas e em regiões próximas ao sítio doador, dando origem a aglomerados desses elementos. Essa configuração em si já favorece a ocorrência de transposição alternativa e a reorganização de sequências codificantes e regulatórias, que muito podem contribuir para a plasticidade e para a evolução dos genomas (Dooner *et al.*, 1994).

Embora os rearranjos cromossômicos possam ser mediados por diferentes fatores, a descoberta do envolvimento direto dos Elementos de Transposição em suas origens proporcionou uma nova visão do seu papel na evolução dos genomas. De simples “DNA lixo” até a condição de agente da reorganização dos cariótipos, sem dúvida, foi um grande salto.

8. REFERÊNCIAS BIBLIOGRÁFICAS ANEXO III

Batzer MA, Deininger PL (2002) *Alu* repeats and human genomic diversity Nat Rev Genet 3: 370-379.

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.

Callinan PA, Wang J, Herke SW, Garber RK, Liang P, Batzer MA (2005) *Alu* retrotransposition-mediated deletion. J Mol Biol 348:791–800.

Casals F, Cáceres M, Ruiz A (2003) The *Foldback-like* transposon *Galileo* is involved in the generation of two different natural chromosomal inversions of *Drosophila buzzatii*. Mol Biol Evol 20: 674-685.

Chisholm KM, Aubert SD, Freese KP, Zakian VA, King MC, Welch PL (2012) A genomewide screen for suppressors of *Alu* -mediated rearrangements reveals a role for *PIF1*. PLoS One 7:e30748.

Courage-Tebbe U, Döring HP, Fedoroff N, Starlinger P (1983) The controlling element *Ds* at the Shrunken locus in *Zea mays*: structure of the unstable *sh-m5933* allele and several revertants. Cell 34: 383-93.

Delprat A, Negre B, Puig M, Ruiz A (2009) The transposon *Galileo* generates natural chromosomal inversions in *Drosophila* by ectopic recombination. PLoS One 18;4: e7883.

Dobzhansky T (1970) Balancing Selection and Chromosomal Polymorphism. In Dobzhansky T (ed) Genetics of the Evolutionary Process. Columbia University, New York, pp 433-484.

Dooner HK, Belachew A, Burgess D, Harding S, Ralston M, Ralston E (1994) Distribution of unlinked receptor sites for transposed *Ac* elements from the *bz-m2(Ac)* allele in maize. Genetics 136:261–279.

Drosophila 12 Genomes Consortium. (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 480: 2003-218.

Engels WR, Preston CR (1984) Formation of chromosome rearrangements by *P* factors in *Drosophila*. *Genetics* 107: 657-678.

Evans AL, Mena PA, McAllister BF (2007) Positive selection near an inversion breakpoint on the neo-X chromosome of *Drosophila americana*. *Genetics* 177(3): 1303-19.

Evgen'ev M, Zelentsova H, Mnjoian L, Poluectova H, Kidwell MG (2000) Invasion of *Drosophila virilis* by the *Penelope* transposable element. *Chromosoma* 109:350-357.

Fedoroff NV (1989) Maize TE. In: Berg DE and Howe MM (eds) *Mobile DNA*. American Society for Microbiology, USA, pp 375- 412.

Fonseca NA, Vieira CP, Schlötterer C, Vieira J (2012) The *DAIBAM* MITE element is involved in the origin of one fixed and two polymorphic *Drosophila virilis* phylad inversions. *Fly (Austin)* 6(2): 71-4.

Gray YH, Tanaka MM, Sved JA (1996) *P*-element-induced recombination in *Drosophila melanogaster*: hybrid element insertion. *Genetics* 144:1601-10.

Gray YH (2000) It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements. *Trends Genet* 16: 461–468.

Griffiths AJF, Wessler SR, Lewontin RC, Carroll SB (2008) Mudança na Estrutura dos cromossomos. In: Griffiths AJF (ed) *Introdução à Genética* 9th edition. Guanabara Koogan, Rio de Janeiro, pp 490-503.

Guillén Y, Ruiz A (2012) Gene alterations at *Drosophila* inversion breakpoints provide prima facie evidence for natural selection as an explanation for rapid chromosomal evolution. *BMC Genomics* 1: 13-53.

Han K, Sen SK, Wang J, Callinan PA, Lee J, Cordaux R, Liang P, Batzer MA (2005) Genomic rearrangements by *LINE -1* insertion-mediated deletion in the human and chimpanzee lineages. *Nucleic Acids Res* 33: 4040–4052.

Han K, Lee J, Meyer TJ, Wang J, Sen SK, Srikanta D, Liang P, Batzer MA. (2007) *Alu* recombination-mediated structural deletions in the chimpanzee genome. *PLoS One Genetics* 3:1939–1949.

Kass LB, Chomet P (2009) Barbara McClintock. In Bennetzen JL and Hake S (eds). Handbook of Maize: Genetics and Genomics. Springer New York, New York, pp 17-52.

Kehrer-Sawatzki H, Cooper DN (2008) Molecular mechanisms of chromosomal rearrangement during primate evolution. *Chromosome Res* 16: 41-56.

Kidwell MG (2005) TE and the Evolution of Dipterian Genomes. In Yeates DK and Wiegmann BM (eds). The Evolutionary Biology of the Flies. Columbia University Press, New York, pp 145-172.

Lacadena RJ (1996) Variaciones cromosómicas estructurales I: Deleciones y Duplicaciones; II: Inversiones, y III: Translocaciones. In Lacadena RJ (ed) Citogenética. Complutense, Madrid, pp 511-570.

Lazarow K, Doll ML, Kunze R (2013) Molecular biology of maize *Ac/Ds* elements: an overview. *Methods Mol Biol* 1057: 59-82.

Lee J, Han K, Meyer TJ, Kim HS, Batzer MA (2008) Chromosomal inversions between human and chimpanzee lineages caused by retrotransposons. *PLoS One* 3: e4047.

Lee J, Ha J, Son SY, Han K (2012) Human Genomic Deletions Generated by SVA-Associated Events. *Comp Funct Genomics*: ID807270.

Lim JK (1988) Intrachromosomal rearrangements mediated by *hobo* transposons in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 85: 9153-9157.

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

Marzo M, Puig M, 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.

McClintock B (1941) The Stability of Broken Ends of Chromosomes in *Zea Mays*. *Genetics* 26(2): 234-82.

Ohno S (1970) Evolution by gene duplication. Springer, New York, 160p.

Peterson T, Zhang J (2013) The Mechanism of *Ac/Ds* Transposition. In Fedoroff NV (ed.). *Plant Transposons and Genome Dynamics in Evolution*. Wiley- Blackwell, USA, pp 41-56.

Powell JR (1997) Population Genetics: Inversions. In: Powell JR (ed) *Progress and prospects in evolutionary biology: the Drosophila model*. Oxford University, New York, pp: 48-80.

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

Schwartz A, Chan DC, Brown LG, Alagappan R, Pettay D, Disteche C, McGillivray B, de la Chapelle A, Page DC (1998) Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through *LINE-LINE* recombination. *Hum Mol Genet* 7: 1–11.

Sen SK, Han K, Wang J, Lee J, Wang H, Callinan PA, Dyer M, Cordaux R, Liang P, Batzer MA (2006) Human genomic deletions mediated by recombination between *Alu* elements. *Am J Hum Genet* 79: 41–53.

Stankiewicz P, Kulkarni S, Dharmadhikari AV, Sampath S, Bhatt SS, Shaikh TH, Xia Z, Pursley AN, Cooper ML, Shinawi M *et al.* (2012) Recurrent deletions and reciprocal duplications of 10q11.21q11.23 including *CHAT* and *SLC18A3* are likely mediated by complex low-copy repeats. *Hum Mutat* 33: 165-79.

Wang H, Xing J, Grover D, Hedges DJ, Han K, Walker JA, Batzer MA (2005) *SVA* elements: a hominid-specific retroposon family. *J Mol Biol* 354: 994-1007.

Yu C, Zhang J, Peterson T (2011) Genome rearrangements in maize induced by alternative transposition of reversed *Ac/Ds* termini. *Genetics* 188:59-67.

Zhang J, Zhang F, Peterson T (2006) Transposition of reversed *Ac* element ends generates novel chimeric genes in maize. *PLoS Genetics* 6;2(10):e164.

Zhang J, Yu C, Krishnaswamy L, Peterson T (2011) TE as Catalysts for Chromosome Rearrangements. *Methods Mol Biol* 701:315-26.