
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS
(Biologia Celular e Molecular)**

**CARACTERIZAÇÃO CROMOSSÔMICA DE *DIATRAEA SACCHARALIS*
(CRAMBIDAE, LEPIDOPTERA) COM ÊNFASE NOS CROMOSSOMOS SEXUAIS
E DNAS REPETITIVOS**

ANA ELISA GASPAROTTO DA SILVA

Dissertação apresentada ao Instituto de Biociências do Campus de Rio Claro, Universidade Estadual Paulista "Júlio de Mesquita Filho", como parte dos requisitos para obtenção do título de Mestre no programa de pós-graduação em Ciências Biológicas (Biologia Celular e Molecular).

Novembro – 2021

ANA ELISA GASPAROTTO DA SILVA

**CARACTERIZAÇÃO CROMOSSÔMICA DE *Diatraea saccharalis*
(CRAMBIDAE, LEPIDOPTERA) COM ÊNFASE NOS CROMOSSOMOS
SEXUAIS E DNAS REPETITIVOS**

Orientador: Prof. Dr. Diogo Cavalcanti Cabral de Mello

Coorientadora: Dra. Vanessa Bellini Bardella

Dissertação de Mestrado apresentada ao Instituto de Biociências do Campus de Rio Claro, Universidade Estadual Paulista “Júlio de Mesquita Filho”, como parte dos requisitos para obtenção do título de Mestre no programa de pós-graduação em Ciências Biológicas (Biologia Celular e Molecular).

Rio Claro

2021

S586c Silva, Ana Elisa Gasparotto da
Caracterização cromossômica de *Diatraea saccharalis* (Crambidae, Lepidoptera) com ênfase nos cromossomos sexuais e DNAs repetitivos / Ana Elisa Gasparotto da Silva. -- Rio Claro, 2021
85 f. : tabs., fotos

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências, Rio Claro
Orientador: Diogo Cavalcanti Cabral de Mello
Coorientadora: Vanessa Bellini Bardella

1. fusão cromossômica. 2. FISH. 3. cromossomo holocêntrico. 4. DNAs repetitivos. 5. cromatina sexual. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências, Rio Claro. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Caracterização cromossômica de *Diatraea saccharalis* (Crambidae, Lepidoptera) com ênfase nos cromossomos sexuais e DNAs repetitivos

AUTORA: ANA ELISA GASPAROTTO DA SILVA

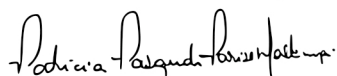
ORIENTADOR: DIOGO CAVALCANTI CABRAL DE MELLO

COORIENTADORA: VANESSA BELLINI BARDELLA

Aprovada como parte das exigências para obtenção do Título de Mestra em CIÊNCIAS BIOLÓGICAS (BIOLOGIA CELULAR E MOLECULAR), pela Comissão Examinadora:



Prof. Dr. DIOGO CAVALCANTI CABRAL DE MELLO (Participação Virtual)
Departamento de Biologia Geral e Aplicada / Unesp - Instituto de Biociências de Rio Claro



Profa. Dra. PATRICIA PASQUALI PARISE MALTEMPI (Participação Virtual)
Departamento de Biologia Geral e Aplicada / Unesp - Instituto de Biociências - Rio Claro



Profa. Dra. DENILCE MENESES LOPES (Participação Virtual)
Universidade Federal de Viçosa

Rio Claro, 08 de novembro de 2021

Agradecimentos

Ao meu pai, Luis Carlos, meu maior exemplo de dedicação, que sempre me incentivou a buscar meus sonhos e não desistir nem um minuto do que eu almejava ser. À minha mãe Rosângela, meu porto seguro, que sempre esteve de braços abertos para me acalmar e confortar meus medos. Ao meu irmão Fernando, pelo apoio, por me aconselhar e estar do meu lado. A minha conquista devo a vocês por me darem asas e forças para buscar o meu melhor caminho.

Agradeço imensamente ao meu orientador Professor Diogo C. Cabral De Mello, por acreditar, confiar e me incentivar em um novo desafio no estudo de citogenética de Lepidoptera. Obrigada por sempre estar disposto a me ensinar e estar de prontidão para resolver qualquer dificuldade minha. Agradeço por ser um grande exemplo de orientador, pesquisador, professor e amigo.

Agradeço a minha coorientadora Vavá pelas conversas e conselhos, por estar presente em cada manhã preparando FISH e por me auxiliar na pesquisa. Agradeço aos amigos do grupo Insect Cytogenomics Group, Ana Beatriz, Dioguinho, Emiliano e Lucas, por me ajudarem nessa caminhada, pelas risadas e por fazerem os dias de laboratório tornarem-se mais leves e divertidos. Agradeço aos nossos queridos agregados, Ger, Tati, Bernardo e Roque, obrigada pela amizade, pelos encontros e por torcerem por mim. Agradeço também à Professora Patrícia P. Parise Maltempì e seu grupo pela troca de conhecimento, pela amizade e por nosso convívio diário.

Agradeço aos meus amigos que conquistei nessa jornada e ao longo da vida, que mesmo não estando perto diariamente, se fazem tão presente compartilhando conquistas e alegrias. Em especial a Michelle, Rafael, Mileni, Guilherme, Samantha e Leonardo, obrigada pelos momentos divididos.

E, finalmente, agradeço ao Programa de Pós-Graduação em Biologia Celular e Molecular, no Instituto de Biociências da UNESP – campus Rio Claro/SP.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

“The very essence of instinct is that it's followed independently of reason.”

Charles Darwin

RESUMO

A família de mariposas Crambidae incluem representantes que causam prejuízos econômicos às lavouras agrícolas, como a broca da cana-de-açúcar *Diatraea saccharalis*. Embora espécies de Crambidae apresentem extensa radiação adaptativa e importância econômica, pouco se sabe sobre sua arquitetura genômica e evolução cromossômica. Neste trabalho, caracterizamos os cromossomos e DNAs repetitivos de *D. saccharalis*, através do sequenciamento genômico e associação de métodos citogenéticos e bioinformáticos, com ênfase nos cromossomos sexuais. O número diploide foi diferente entre os sexos, i.e., $2n = 33$ (fêmeas) e $2n = 34$ (machos), além disso, foi observado a ocorrência de um sistema sexual múltiplo $WZ_1Z_2/Z_1Z_1Z_2Z_2$. A sonda telomérica revelou um forte sinal intersticial no cromossomo W e este cromossomo foi associado a dois cromossomos Z (Z_1 e Z_2), confirmando a presença do trivalente sexual WZ_1Z_2 como resultado de uma fusão cromossômica. Acerca dos DNAs repetitivos, os Elementos de Transposição (TEs) representaram cerca de 39,18% (machos) a 41,35% (fêmeas) enquanto que os DNAs satélites (DNAsat) representaram apenas 0,214% do genoma de machos e 0,215% de fêmeas. O mapeamento por FISH revelou distinta organização cromossômica para os DNAsat, como único sinal clusterizado, sinais dispersos e *repeats* não clusterizados. Os dois TEs mapeados apresentaram sinais espalhados. Embora nota-se um pequeno enriquecimento de alguns DNAsat no genoma de fêmeas, eles não foram diferencialmente enriquecidos no cromossomo W. Entretanto, notamos o enriquecimento dos sinais de TEs no cromossomo W, sugerindo seu envolvimento na degeneração e diferenciação deste cromossomos. Esses resultados revelam o dinamismo da estrutura geral do cariótipo e DNAs repetitivos em *D. saccharalis*, principalmente como resultado de fusão cromossômica e disseminação ou amplificação local de DNAs repetitivos. Por fim, nossos dados contribuem para o entendimento geral do genoma das espécies que podem auxiliar em futuros estudos genômicos.

Palavras-chave: fusão cromossômica, FISH, cromossomo holocêntrico, DNAs repetitivos, cromatina sexual

ABSTRACT

Crambidae lepidopterans include pests that causes economic losses to agricultural crops, as the sugarcane borer *Diatraea saccharalis*. Although the crambid moths have extensive adaptive radiation and economic importance, little is known about their genome architecture and chromosomal evolution. Here we characterized the chromosomes and repetitive DNAs of *D. saccharalis* by association of low pass genome sequencing, bioinformatic and cytogenetic methods, with an emphasis on sex chromosomes. Diploid numbers differed between sexes, i.e., $2n=33$ females and $2n=34$ males, and occurrence of a multiple sex system $WZ_1Z_2/Z_1Z_1Z_2Z_2$ was noticed. Telomeric probe revealed one strong interstitial signal on the W chromosome, and this chromosome was associated to two chromosomes Z (Z_1 and Z_2), confirming the occurrence of a sex trivalent WZ_1Z_2 , as result of a chromosomal fusion. Concerning the repetitive DNAs, the TEs represented about 39,18% (males) to 41,35% (females), while satDNAs comprised only 0,214% of male genome and 0,215% of female genome. FISH mapping revealed distinct chromosomal organization for satDNAs, as single located clusters, spread signals and non-clustered repeats. The two TEs mapped by FISH were scattered. Although we noticed slightly enrichment of some satDNAs on female genome, they were not differentially enriched on the W chromosome. On the other hand, we noticed enrichment of FISH signals on the W chromosome for TEs, suggesting their involvement on W chromosome degeneration and differentiation. These data reveal dynamism of general karyotype structure and repetitive DNAs of *D. saccharalis*, mainly as result of chromosomal fusions and spread or local amplification of repetitive DNAs. Finally, our data can help in future structural and functional genomic analysis that could be also useful in pest control programs.

Key-words: chromosome fusion, FISH, holocentric chromosome, repetitive DNA, W-chromatin.

SUMÁRIO

1. INTRODUÇÃO	7
1.1. DNAs repetitivos: estrutura e evolução	7
1.2. Cromossomos holocêntricos	10
1.3. Cromossomos sexuais em Lepidoptera	12
1.3.1. Estudos cromossômicos em Crambidae	16
2. OBJETIVOS	20
2.1. Geral.....	20
2.2. Específicos.....	20
3. MATERIAL E MÉTODOS	21
3.1. Material biológico e análises cromossômicas convencionais.....	21
3.2. Preparação de núcleos poliploides e microdissecção da cromatina sexual.....	22
3.3. Caracterização do <i>repeatome</i> através do sequenciamento genômico.....	22
3.4. Caracterização genômica e mapeamento cromossômico por Hibridização <i>in situ</i> fluorescente (FISH)	24
4. RESULTADOS E DISCUSSÃO	27
4.1. Manuscrito.....	28
5. CONCLUSÕES FINAIS.....	73
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	74

1. INTRODUÇÃO

1.1. DNAs repetitivos: estrutura e evolução

Os DNAs repetitivos representam um importante componente organizacional e funcional dos genomas eucarióticos. Esses elementos estão representados nos genomas por famílias de diferentes classes com variável grau de repetição, encontrando-se organizadas em tandem (uma repetição após a outra) ou dispersas pelo genoma. Dentre as sequências repetitivas estão incluídas as famílias multigênicas, os elementos de transposição (TEs) e os DNAs satélites (DNAsat) (CHARLESWORTH *et al.* 1994; LÓPEZ-FLORES e GARRIDO-RAMOS, 2012). Estas sequências estão envolvidas na organização e funcionalidade dos centrômeros, telômeros, segregação cromossômica, reparo e replicação do DNA, assim como na diferenciação dos genomas, especialmente de alguns tipos de cromossomos, tais como os cromossomos sexuais (MESSIER *et al.* 1996; MARTINS 2007; LÓPEZ-FLORES e GARRIDO-RAMOS 2012; BISCOTTI *et al.* 2015; GARRIDO-RAMOS 2017). Adicionalmente, os DNAs repetitivos estão envolvidos na formação e manutenção da heterocromatina e em processos de regulação gênica (SHAPIRO e STERNBERG 2005; BIÉMONT e VIEIRA 2006; FESCHOTTE e PRITHAM 2007; POMBO e DILLON 2015).

O DNA satélite é caracterizado por sua organização genômica em longos arranjos de unidades em tandem, formados por múltiplos monômeros idênticos ou com grande similaridade. Essas sequências ocorrem nos genomas em centenas a milhares de cópias organizadas de uma unidade de repetição, que apresenta tamanhos com 5pb até mais de 1.000pb (pares de base) e, em geral, são altamente dinâmicos em sua composição de bases da sequência de nucleotídeos, comprimento de monômeros e localização cromossômica (MACAS *et al.* 2002; RUIZ-RUANO *et al.* 2016; GARRIDO-RAMOS 2017). Em geral, os DNAsat consistem em uma única sequência repetida diversas vezes, na qual o surgimento e fixação das diferentes famílias nas espécies dependem de mecanismos como

a mutação seguida de evolução em concerto, que envolve a homogeneização dos monômeros (CHARLESWORTH *et al.* 1994; PLOHL 2008; LÓPEZ-FLORES e GARRIDO-RAMOS 2012).

As sequências de DNA satélite estão enriquecidas na heterocromatina constitutiva e, em sua maioria, encontram-se principalmente nas regiões centroméricas e teloméricas dos cromossomos, embora muitos casos de DNAsat eucromáticos tenham sido descritos recentemente (CHARLESWORTH *et al.* 1994; LÓPEZ-FLORES e GARRIDO-RAMOS 2012; GARRIDO-RAMOS 2017; PALACIOS-GIMENEZ *et al.* 2020). Além de estarem clusterizados em regiões heterocromáticas, os DNAsat são componentes fundamentais dos cromossomos sexuais, como descrito em planta (*Silene latifolia*), mosca de fruta (*Drosophila melanogaster*); grilo (*Eneoptera surinamensis*), gafanhoto (*Ronderosia bergii*), mariposa (*Plodia interpunctella*), dentre outros (CHARLESWORTH *et al.* 1994; KEJNOVSKY *et al.* 2013; HOBZA *et al.* 2015; DALIKOVÁ *et al.* 2017b, PALACIOS-GIMENEZ *et al.* 2017; FERRETTI *et al.* 2020). E embora a maioria dessas sequências esteja presente na heterocromatina, região considerada transcritionalmente silenciosa e inativa, estudos mostram evidência de transcrição desses elementos em insetos, plantas e vertebrados (UGARKOVÍČ 2005; PEZER *et al.* 2012; BISCOTTI *et al.* 2015). Os transcritos de DNAsat estão envolvidos, em geral, na formação e manutenção da heterocromatina, na funcionalidade dos centrômeros e no alongamento e replicação de telômeros (BISCOTTI *et al.* 2015).

Além disso, os DNAsat compreendem um dos componentes mais dinâmicos do genoma, sofrendo rápidas mudanças na composição de arranjos e sequências em um curto período de tempo. A expansão e encolhimento de DNAsat contribuem significativamente para geração de polimorfismo (PLOHL *et al.* 2014; GARRIDO-RAMOS 2015). Adicionalmente, a rápida evolução destas sequências repetitivas pode causar barreiras

reprodutivas entre os organismos, promovendo a especiação (MALIK e HENIKOFF 2009; PLOHL *et al.* 2014; FELICIELLO *et al.* 2014).

Outro grupo de sequências repetitivas são as famílias multigênicas que representam um conjunto de genes que possuem sequências de DNA com notável similaridade estrutural e funcional. Essas sequências podem ocorrer em múltiplas cópias no genoma, e dentre as mais analisadas em estudos cromossômicos estão os genes de RNAs ribossomais (rRNA), snRNA (*small nuclear RNA*) e genes codificadores de proteínas histônicas (NEI e ROONEY 2005). A evolução em concerto é um mecanismo evolutivo proposto para as famílias multigênicas, no qual todos os genes são considerados como um conjunto e, uma mutação ocorrendo em uma repetição tende a se espalhar através dos genes membros inteiros por ocorrência repetida de *crossing-over* desigual ou conversão gênica. Outra forma de evolução das famílias multigênicas é por meio do nascimento e morte, a qual novos genes duplicados são mantidos no genoma por um longo tempo, enquanto outros são deletados ou inativos por mutações deletérias (NEY e ROONEY 2005; EIRÍN-LOPEZ *et al.* 2012). Mais recentemente, a ocorrência de um modelo misto envolvendo evolução em concerto e nascimento e morte tem sido relatado para as famílias multigênicas em diversos grupos (FREIRE *et al.* 2010; MERLO *et al.* 2013; BUENO *et al.* 2016; LIU *et al.* 2018; BARDELLA e CABRAL-DE-MELLO 2018; MARTÍ *et al.* 2021).

Finalmente, os TEs são sequências de DNA que se diferem pela capacidade de alterar sua posição no genoma, muitas vezes se duplicando durante o processo. Como resultado de sua diversificação contínua, os TEs apresentam uma grande variedade de estrutura e organização (KIDWELL 2005; BOURQUE *et al.* 2018). A diversidade desses elementos é dividida em duas classes principais de acordo com seu mecanismo de transposição. A classe I, denominada retrotransposons, no qual um intermediário de RNA

é transcrito reversamente em uma cópia de cDNA que é integrada em local distinto no genoma; e a classe II, conhecida como transposons de DNA, que se mobilizam por meio de transposição de cópias de DNA (BOEKE *et al.* 1985; KIDWELL 2005; WICKER *et al.* 2007; CRAIG *et al.* 2015).

A inserção dos TEs pode interromper sequências codificadoras ou regulatórias de genes causando efeitos deletérios pela modificação ou eliminação de um gene (LYNCH 2007; FESCHOTTE 2008; BOURQUE *et al.* 2018; SERRATO-CAPUCHINA e MATUTE 2018). Em *Arabidopsis sp.* (Brassicaceae) foi observado a inserção de um transposon “*Mutator-like*” no locus de floração C (FLC) resultando na interrupção do gene e levando a um atraso no tempo de floração (LIU *et al.* 2004). Além disso, TEs podem desempenhar um papel significativo na diferenciação e evolução de cromossomos sexuais, isso porque esses elementos podem amplificar as regiões não homólogas dos cromossomos Y ou W por meio de inserções e permitindo inversões (CHARLESWORTH *et al.* 2005; HUA-VAN *et al.* 2005; BACHTROG *et al.* 2008; SLIWINSKA *et al.* 2016). O acúmulo de TEs em cromossomos sexuais foi observado em espécies como na mosca de fruta, *Drosophila miranda*, e na planta *Silene latifolia* (BACHTROG 2005; KEJNOVSKY *et al.* 2013; BACHTROG *et al.* 2013). Deste modo, os TEs são componentes distintos e difundidos nos genomas eucarióticos com potencial de impactar a diversificação e evolução (SERRATO-CAPUCHINA e MATUTE 2018).

1.2. Cromossomos holocêntricos

O centrômero é a região essencial no cromossomo pois, é neste local que as fibras do fuso se ligam em uma única placa proteica, o cinetócoro, no processo de segregação cromossômica durante a divisão celular e caracteriza os cromossomos conhecidos como monocêntricos (DERNBURG 2001; HOUBEN e SCHUBERT 2003). Esta região é

enriquecida por distintas sequências de DNAs repetitivos, principalmente por DNAsat (HOUBEN e SCHUBERT 2003). Entretanto, alguns grupos como os artrópodes (exemplos, Odonata, Lepidoptera, Hemiptera) e plantas (Cyperaceae, Juncaceae), apresentam a placa cinetocórica estendendo-se ao longo de todo o cromossomo, em uma distribuição típica dos cromossomos denominados holocêntricos (HAIZEL *et al.* 2005; ZEDEK e BURES 2016; MELTERS *et al.* 2012; ESCUDERO *et al.* 2016). Esses cromossomos também são chamados de holocinéticos, com potencial de atividade cinetocórica ao longo de toda a extensão dos cromossomos e, durante a divisão celular, as cromátides separam-se em paralelo e não no formato de “V” característico de cromossomos monocêntricos (COOPER 1972; WRENSCH *et al.* 1994; MANDRIOLI e MANICARDI 2020). Além da grande diferença em relação a estrutura e segregação entre os cromossomos holocêntricos e monocêntricos, a organização de sequências de DNA (eucromatina e heterocromatina) também é diferencial entre os mesmos. De modo geral, os DNAsat estão enriquecidos no centrômero dos cromossomos monocêntricos, enquanto que para os holocêntricos, esta classe de DNA repetitivo pode estar mais dispersa pelo cromossomo (HECKMANN *et al.* 2012; MARQUES *et al.* 2015). Essa variação na organização de DNAs repetitivos foi evidenciada na espécie de hemíptera, *Chariesterus armatus*, em que se encontra distintos arranjos de DNA ribossomal 5S ao longo de todo cromossomo, disperso em regiões eucromáticas do genoma (BARDELLA e CABRAL-DE-MELLLO 2018), assim como para DNAsat em mariposas (CABRAL-DE-MELLO *et al.* 2021), percevejos (BARDELLA *et al.* 2020) e plantas (MARQUES *et al.* 2015).

Os cromossomos holocêntricos surgiram de forma independente treze vezes durante a evolução das plantas e animais, sendo nove em animais e quatro em plantas (MELTERS *et al.* 2012). Em Insecta, as ordens Lepidoptera e Hemiptera representam os maiores grupos de espécies com cromossomos holocêntricos. Embora a origem holocêntrica

desses cromossomos seja independente entre os dois grupos, ambos compartilham de modo específico a ausência da proteína cinetocórica (CenH3), a qual é conservada no nucleossomo responsável pela segregação nos eucariotos (MELTERS *et al.* 2012; DRINNENBERG *et al.* 2014).

A natureza holocêntrica, como consequência de seu cinetócoro difuso, pode estabilizar fragmentos cromossômicos, favorecendo os rearranjos do cariótipo (MONTI *et al.* 2012; MANDRIOLI e MANICARDI 2020). Entre os animais, os lepidópteros exibem notável variabilidade cariotípica inter e intraespecífica. Lepidoptera apresenta tal tolerância a variações cromossômicas em vista de seus cromossomos holocêntricos facilitarem a herança bem sucedida de novos fragmentos de fissão ou fusão e, como resultado, os lepidópteros evitam os consequentes eventos deletérios da fissão e fusão cromossômica em larga escala (MAREC *et al.* 2001; HILL *et al.* 2019; MANDRIOLI e MANDICARDI 2020). Contudo, o cromossomo holocêntrico pode apresentar limitações no processo de *crossing-over*, o que causa uma reestruturação das divisões meióticas resultando em uma meiose invertida (LUKHTANOV *et al.* 2018).

1.3. Cromossomos sexuais em Lepidoptera

Os cromossomos sexuais (X e Y ou Z e W) representam um componente especial do genoma, sendo em diversos grupos os elementos mais notáveis de um complemento cromossômico em termos de tamanho, forma, heterocromatinização e comportamento na meiose. Frequentemente, são esses cromossomos que fornecem o sinal primário na via de determinação do sexo, diferindo macho e fêmea (CHARLESWORTH *et al.* 1994; KEJNOVSKY *et al.* 2009; SAHARA *et al.* 2012; HOBZA *et al.* 2015). Os cromossomos sexuais originaram-se múltiplas vezes independentemente a partir de um par comum de autossomos após a aquisição de um *locus* principal de determinação do sexo por um dos

dois homólogos autossômicos (CHARLESWORTH *et al.* 2005; WRIGHT *et al.* 2016; ABBOTT *et al.* 2017; FURMAN *et al.* 2020). A diferenciação dos cromossomos sexuais ocorre pela supressão de recombinação, que proporciona o acúmulo de diferenças na composição molecular, entre os cromossomos X e Y ou Z e W. Esse evento pode resultar na degeneração molecular dos cromossomos Y/W levando a uma perda de sequências de codificação (pseudogenização) e ao acúmulo de DNAs repetitivos. Enquanto que o cromossomo X ou Z continua a se recombinar no sexo homomórfico e, assim, os genes ligados a esse cromossomo persistem gerando variabilidade por recombinação ao longo das gerações (CHARLESWORTH *et al.* 2005; BACHTROG 2013; NGUYEN e CARABAJAL PALADINO 2016).

Na ordem Lepidoptera (mariposas e borboletas) o número cromossômico é alto, $2n=62$ (VIRKKI 1963), com o cariótipo formado por cromossomos holocêntricos e, geralmente, com tamanho e forma uniformes, além de 1-2 pares grandes de cromossomos, correspondente ao bivalente sexual (ROBINSON 1971; NGUYEN e CARABAJAL PALADINO 2016). A característica mais marcante dos lepidópteros, e que os difere da grande maioria dos outros grupos de insetos, é o mecanismo de determinação sexual com a fêmea heterogamética representada pelo sistema sexual WZ/ZZ (fêmea/macho). Com base em estudos da árvore filogenética de Lepidoptera, foi evidenciado que o sistema sexual Z/ZZ é comum nas linhagens basais de lepidópteros, indicando que o cromossomo W foi uma aquisição posterior. O W surgiu no ancestral comum dos clados-irmãos Ditrysia e Tischerioidea (TRAUT e MAREC 1996; LUKHTANOV 2000). Uma das hipóteses sobre a origem evolutiva do W postula que ele se derivou de um autossomo, tornando-se feminino por uma fusão de seu homólogo com o cromossomo Z ancestral. Outra hipótese propõe o surgimento do W a partir de um cromossomo supranumerário (cromossomo B) que foi incorporado no genoma, adquirindo combinação com o

cromossomo Z (LUKHTANOV 2000; MAREC *et al.* 2010; SAHARA *et al.* 2012; DALÍKOVÁ *et al.* 2017a). Curiosamente, alguns grupos recentes de lepidópteros, ocasionalmente e independentemente, apresentam o sistema sexual Z/ZZ, devido à perda secundária do cromossomo W. Enquanto que ainda, em outras espécies, o sistema sexual WZ/ZZ é substituído por um sistema sexual múltiplo a partir da fissão ou fusão cromossômica do W ou Z com um autossomo, resultando nos cromossomos neo-sexuais (exemplo, W_1W_2Z/ZZ e $WZ_1Z_2/Z_1Z_1Z_2Z_2$) (MAREC *et al.* 2010; SAHARA *et al.* 2012; SICHOVÁ *et al.* 2015). Um caso interessante de organização do sistema sexual foi observado na espécie selvagem de traça-da-seda, *Samia cynthia*, em que três diferentes sistemas foram observados em três populações diferentes. A população do Vietnã apresentou sistema sexual Z/ZZ, a população de Sapporo com um sistema WZ/ZZ e a população de Nagano com um sistema sexual múltiplo $WZ_1Z_2/Z_1Z_1Z_2Z_2$ (YOSHIDO *et al.* 2005; YOSHIDO *et al.* 2011).

O cromossomo W é heterocromático e, assim como o Y, sofre degeneração molecular resultando em pseudogenização e considerável acúmulo de DNAs repetitivos (MAREC *et al.* 2010) como, por exemplo, elementos transponíveis (non-LTRs e SINEs) identificados no W da espécie de mariposa *Bombyx mori* (OSANAI-FUTAHASHI *et al.* 2008) e, também, o DNA satélite (PiSAT1) caracterizado na espécie de traça- indiana, *Plodia interpunctella* (DALÍKOVÁ *et al.* 2017b). Nas espécies de Crambidae, *Ostrinia nubilalis* e *Cydalima perspectalis*, o cromossomo W está enriquecido com três satélites, contribuindo para sua diferenciação dos cromossomos Z (CABRAL-DE-MELLO *et al.* 2021). Além disso, o W é responsável por uma particularidade no núcleo interfásico dos lepidópteros. Devido a sua natureza heterocromática, o W forma um corpúsculo denso conhecido como cromatina sexual, a qual é específica nas fêmeas. A cromatina sexual W é altamente condensada e especialmente visível em tecidos poliploides dos túbulos de

Malpighi e das glândulas de seda (MAREC *et al.* 2010). Uma aplicabilidade para este corpúsculo é através da microdissecção da cromatina W para ser utilizada como sonda na pintura cromossômica para identificação do cromossomo W. Esta cromatina sexual, derivada do cromossomo W, é ausente no sistema sexual Z0 (FUKOVÁ *et al.* 2007; VÍTKOVÁ *et al.* 2007; TRAUT *et al.* 2007).

Em contraste ao W, o cromossomo Z é bastante conservado e, de modo geral, sua composição é parecida com a dos autossomos (SAHARA *et al.* 2012). O conteúdo gênico do Z é conservado entre os lepidópteros, sendo muito parecido com o que ocorre no Z das aves (NANDA *et al.* 2008) ou no X dos mamíferos (OHNO 1973). Yasukochi *et al.* 2006 e Pringle *et al.* 2007, identificaram os genes *Tpi* (triose fosfato isomerase) e *apterous* localizados no cromossomo Z de *B. mori* (Bombycoidea) e de uma espécie distantemente relacionada, a borboleta ninfalídea *Heliconius melpomene* (Papilionoidea), indicando a estabilidade evolutiva entre os grupos de Lepidoptera. A sintenia conservada de uma ordem de genes entre o cromossomo 15 de *B. mori* e um cromossomo de *Manduca sexta* (Sphingidae, Bombycoidea) também apoia essa evidência de estabilidade gênica (SAHARA *et al.* 2007; NGUYEN *et al.* 2010). O cromossomo Z parece desempenhar um papel importante no processo de especiação pois, seus genes podem estar relacionados ao controle e reconhecimento de feromônios, favorecendo barreiras reprodutivas (SPERLING 1994; DOPMAN *et al.* 2005; SAHARA *et al.* 2012). Adicionalmente, os genes no cromossomo Z evoluem mais rápido do que aqueles em autossomos pois, encontram-se em um estado hemizigótico e, portanto, estão mais expostos às forças de seleção (MANK *et al.* 2011; SAHARA *et al.* 2012).

A evolução do cromossomo neo-sexual por meio de múltiplas fusões entre cromossomos sexual-autossomo foi descrita em mariposas com cariótipos altamente derivados como a exemplo das espécies *Orgyia antiqua* e *Orgyia thyellina*

(Lymantriidae), e na espécie geográfica *Samia cynthia* (Saturniidae) (YOSHIDO *et al.* 2005). Sugere-se que a seleção natural pode favorecer esses rearranjos cromossômicos, diminuindo a recombinação entre alelos, contribuindo para adaptações locais, que por sua vez, podem aumentar a fixação de mudanças na arquitetura do genoma nas populações (FARIA e NAVARRO 2010; NGUYEN *et al.* 2013). Portanto, os cromossomos neo-sexuais podem desempenhar um papel importante na evolução adaptativa e na especiação com o fluxo gênico (NGUYEN e CARABAJAL PALADINO 2016). Na grande família Tortricidae, por exemplo, o cromossomo neo-Z surgiu pela fusão entre um cromossomo Z ancestral e um autossomo correspondente ao cromossomo 15 no genoma de referência de *Bombyx mori*. A fusão associou dois genes principais que conferem resistência a inseticidas, além de um arranjo de genes envolvidos na desintoxicação de metabólitos secundários de plantas ao cromossomo neo-sexual. Assim, os resultados sugerem que o cromossomo neo-Z de Tortricidae é um elemento relevante de inovação evolutiva que aumentou significativamente o potencial adaptativo das mariposas tortricidas e contribuiu para sua radiação adaptativa e subsequente especiação (NGUYEN *et al.* 2013).

1.3.1. Estudos cromossômicos em Crambidae

A família de mariposas Crambidae, com mais de onze mil espécies descritas, corresponde a uma das famílias mais especiosa entre os grupos de insetos lepidópteros. Além da riqueza de espécies, também inclui importantes pragas agrícolas, sendo mais conhecida a broca de milho europeia, *Ostrinia nubilalis*. Essa espécie nativa da Europa, originalmente infestava variedades de painço, mas, no início dos anos 1900, ao ser introduzida nos Estados Unidos, tornou-se uma praga séria do milho. Além disso, há outras espécies de importantes pragas, como, por exemplo, a broca de milho neotropical, *Diatraea lineolata*, que ocorre no sul do Texas, Caribe, América Central até o Brasil, e a

broca de milho do sudoeste, *Diatraea grandiosella*, encontrada no México e partes do sul da América do Norte. Outra importante espécie, a broca-do-caule do arroz, *Chilo suppressalis*, comum na China, Índia, Japão e sudeste da Ásia, é uma das principais pragas do arroz. Além disso, várias espécies de Crambidae se estabeleceram como principais pragas da cana-de-açúcar. Essas são, por exemplo, a broca da cana (*Chilo infuscatellus*) e a broca manchada (*Chilo sacchariphagus*), ambas nativas do sul e sudeste da Ásia, e a broca da cana-de-açúcar, *Diatraea saccharalis*, que é encontrada no Caribe, América Central e na América do Sul, destacando-se como importante praga nas plantações da cana-de-açúcar do Brasil. Desta forma, a família Crambidae é uma das principais linhagens de lepidópteros em termos de riqueza e importância econômica (ROBINSON 1971).

Contudo, apesar destas mariposas terem grande importância, como demonstrado em estudos sobre sua taxonomia, ecologia, genética de populações e controle de pragas, pouco se sabe sobre sua arquitetura genômica pois, a citogenética de Crambidae é pouco explorada. Isso se aplica particularmente aos cromossomos sexuais que foram estudados em duas espécies do gênero *Ostrinia*, *O. scapularis* (KAGEYMA e TRAUT 2004) e *O. nubilalis* (YASUKOCHI *et al.* 2016), ambas apresentando cariótipo haploide $n=31$, número cromossômico ancestral de Lepidoptera, e sistema sexual ZW (fêmea) e ZZ (macho), comum nesse grupo (TRAUT *et al.* 2007; MAREC *et al.* 2010; SAHARA *et al.* 2012). No entanto, ainda em algumas outras espécies de Crambidae, apenas o número cromossômico é conhecido (ROBINSON 1971), como na espécie *Chilo infuscatellus*, com número reduzido de cromossomo $n=28$, além de 1-2 pares grandes de cromossomos, indicando fusão cromossômica na evolução do cariótipo (ZHANGWU *et al.* 1994). Um grande par de cromossomos também foi encontrado na espécie *O. scapularis*, equivalente ao bivalente sexual WZ (KAGEYMA e TRAUT 2004). Em lepidópteros, grandes pares

de cromossomos sexuais ocorrem com uma frequência muito alta e, foi demonstrado em vários grupos de lepidópteros que eles representam o cromossomo neo-sexual, originado pela fusão de um cromossomo sexual ancestral com autossomo (YOSHIDO *et al.* 2011, NGUYEN *et al.* 2013; CARABAJAL PALADINO 2016). Logo, parece provável que os grandes cromossomos sexuais observados no cariótipo de algumas espécies de Crambidae também representem cromossomos neo-sexuais.

Recentemente, análises genômicas de três importantes espécies praga, a broca de milho europeia (*Ostrinia nubilalis*), a mariposa praga do *Buxus* sp. (*Cydalima perspectalis*) e a broca de cana-de-açúcar da Guatemala (*Diatraea postlineella*), identificaram sete DNAs satélites (DNAsat) presentes no genoma dessas espécies, dos quais três estavam enriquecidos e arranjos ao longo dos cromossomos W de *C. perspectalis* e *O. nubilalis* e, em contraste, no cromossomo W de *D. postlineella*, não foi identificado a presença de satélites. Adicionalmente, o mapeamento por FISH com sonda de DNAr 18S em células de paquítenos de machos, revelou significativa diferença no número e arranjos de *clusters* de DNAr entre as três espécies de Crambidae. Em *C. perspectalis*, dois arranjos foram encontrados, cada um localizado na posição terminal de diferentes bivalentes, enquanto que para *D. postlineella* apenas um *cluster* foi observado na posição intersticial. Em contraposição, quatro arranjos de DNAr foram identificados em diferentes bivalentes na espécie *O. nubilalis*. Esses resultados elucidam a plasticidade no arranjo e composição de sequências repetitivas nos genomas de lepidópteros além de contribuir para o entendimento da arquitetura genômica de Crambidae (CABRAL-DE-MELLO *et al.* 2021).

Embora as espécies de Crambidae apresentem extensa radiação adaptativa e grande importância econômica, com diversas espécies pragas incluindo a broca da cana-de-açúcar (*Diatraea saccharalis*), ainda pouco se sabe sobre sua organização cromossômica,

composição do genoma e evolução de seus sistemas sexuais. Assim, neste trabalho, caracterizamos o cariótipo, detalhando os DNAs repetitivos e o sistema neo-sexual dessa importante praga, a fim de contribuir com o conhecimento sobre a organização genômica e evolução dos cromossomos sexuais em representantes de Crambidae.

2. OBJETIVOS

2.1. Geral

Entender a composição, estrutura e evolução do genoma, com ênfase nos cromossomos sexuais, da mariposa praga *Diatraea saccharalis*, baseado na análise dos DNAs repetitivos.

2.2. Específicos

2.2.1. Analisar detalhadamente o cariótipo e os cromossomos sexuais da espécie *D. saccharalis* utilizando técnicas convencionais e de citogenética molecular;

2.2.2. Elucidar a composição e origem dos cromossomos sexuais em *D. saccharalis* pela prospecção e mapeamento comparativo de DNAs repetitivos entre machos e fêmeas, utilizando abordagem direta como mapeamento físico cromossômico por Hibridização *in situ* fluorescente – FISH e indireta com análises computacionais.

3. MATERIAL E MÉTODOS

3.1. Material biológico e análises cromossômicas convencionais

Indivíduos machos e fêmeas de *Diatraea saccharalis* foram obtidos da Escola Superior de Agricultura Luiz de Queiroz, ESALQ-USP (Piracicaba, São Paulo/Brasil). As lagartas foram alimentadas com dieta artificial e mantidas na Incubadora B.O.D (Demanda Bioquímica de Oxigênio) (Solab, Piracicaba, Brasil) em condições controladas (temperatura $25 \pm 4^{\circ}\text{C}$, 60% de umidade e 12 horas/luz e 12 horas/escuro) (King *et al.* 1975).

Quatro órgãos diferentes foram dissecados no estágio de lagarta no terceiro ou quarto instar para obtenção de preparações cromossômicas, incluindo os testículos, ovários, discos da asa e tubos de Malpighi. Foram estudados, pelo menos, cinco indivíduos de cada sexo para cada uma das análises. As preparações cromossômicas foram feitas de acordo com Mediouni *et al.* (2004) e Sichová *et al.* (2013). Brevemente, os cromossomos em metáfase meiótica e no estágio de paquíteno foram obtidos de testículos no terceiro instar e ovários do último instar e em estágio inicial de impupação. Cromossomos mitóticos foram obtidos de discos da asa no terceiro instar de ambos os sexos. Os testículos e discos da asa foram dissecados em solução fisiológica, hipotonizados por 10 minutos em KCl 75mM e fixados em solução de Carnoy (etanol/clorofórmio/ácido acético, 6:3:1) por 15 minutos. Os ovários dissecados são transferidos diretamente para a solução de Carnoy sem tratamento hipotônico para preservar o padrão heterocromático do cromossomo W. Os tecidos fixados foram macerados em lâmina contendo uma gota de ácido acético 60% e em seguida espalhados sob placa aquecida a 45°C . As lâminas com qualidade suficiente de material foram desidratadas em série alcoólica (70%, 80% e 100% por 30 segundos cada) e armazenadas a -20°C até utilização em diferentes técnicas citogenéticas. Todo o corpo das lagartas

dissecadas foi armazenado em etanol 100% em freezer a -20°C até a extração do DNA genômico (DNAg).

3.2. Preparação de núcleos poliploides e microdissecção da cromatina sexual

Os túbulos de Malpighi de lagartas de último instar de indivíduos macho e fêmea foram usados para obtenção de células poliploides do núcleo interfásico. Após dissecção em solução fisiológica, os tubos de Malpighi montados em lâminas foram brevemente fixados em Carnoy por 1 minuto e em seguida corados com orceína lática-acética a 1,25% por 5 minutos. As lâminas foram observadas em microscópio de luz para análise da presença ou ausência de cromatina sexual, referente ao cromossomo W, específica nas fêmeas. Este cromossomo é facilmente reconhecido nesta fase por se apresentar altamente condensado como um corpúsculo (Traut e Marec 1996, 1997).

Para a microdissecção da cromatina sexual foram usados tubos de Malpighi dissecados de lagartas fêmeas no último instar, previamente hipotonizados por 10 minutos em uma solução hipotônica (KCl 75mM) e fixados em Carnoy por 15 minutos. O material foi dissociado com agulha de tungstênio em uma gota de ácido acético 60% e espalhado sobre lamínulas em placa aquecida a 40°C e, em seguida, corado com Giemsa 5% a 7 minutos. Um total de 15 corpúsculos foram manualmente microdissectados usando um micromanipulador Eppendorf 5171 acoplado com um microscópio invertido Nikot Axiphot.

3.3 Caracterização do *repeatome* através do sequenciamento genômico

O DNA genômico foi extraído no estágio de lagarta de um indivíduo macho e um indivíduo fêmea no terceiro instar utilizando o kit “Wizard® Genomic DNA Purification Kit” (Promega), seguindo o protocolo descrito pelo fabricante. Os genomas de ambos os

sexos de *D. saccharalis* foram sequenciados via plataforma Illumina Hiseq 4000 paired-end utilizando o serviço da Novogene (HK) Co., Ltd. (Hong Kong, China), gerando *reads* pareados de 2 x 150pb (pares de base). Os *reads* obtidos através do sequenciamento foram processados, filtrados e entrelaçados utilizando as ferramentas disponíveis na plataforma Galaxy/RepeatExplorer (Novák *et al.* 2013; Ruiz-Ruano *et al.* 2016). A identificação e caracterização do Satelitoma (DNAs satélites) foi feita pela ferramenta do TAREAN disponível no RepeatExplorer (Novák *et al.* 2017). Além disso, os clusters de DNAsat foram confirmados como arranjos em tandem *via* dotplot e pelo Tandem Repeats Finder – TRF (Benson 1999). Após a prospecção de DNAsat, foi utilizada a ferramenta do RepeatMasker (Smit *et al.* 2017) com o script python “rm_homology.py” (disponível em: https://github.com/fjruizruano/satminer/blob/master/rm_homology.py) (Ruiz-Ruano *et al.* 2016) para análise de similaridade e definição de todas as famílias de DNAsat recuperadas. Adicionalmente, o RepeatMasker foi utilizado para calcular abundância e divergência de cada DNAsat, comparando os genomas de macho e fêmea. Para estimar os parâmetros de Kimura (K2P) foi utilizado o script “calcDivergenceFromAlign.pl” disponível na ferramenta do RepeatMasker (Smit *et al.* 2017), gerando landscapes representados pela abundância dessas sequências repetitivas no eixo Y e a distância K2P do consenso no eixo X.

A identificação e quantificação de Elementos Transponíveis (TEs) foi feita pela linha de comando dnaPipeTE (Goubert *et al.* 2015). Essa ferramenta utiliza o programa Trinity (Grabherr *et al.* 2011) para montagem das sequências mais abundantes, seguida por caracterização automática com a ferramenta Blast (Goubert *et al.* 2015). Para isso, utilizamos uma biblioteca personalizada de TEs descrita em Martí *et al.* (2021). As superfamílias de TEs identificadas foram realinhadas com a biblioteca Illumina de ambos

os sexos utilizando a ferramenta de RepeatMasker (Smit *et al.* 2017), a divergências K2P e a abundância de TE para fêmea e macho foram feitas conforme descrito acima.

3.4. Caracterização genômica e mapeamento cromossômico por Hibridização *in situ* fluorescente (FISH)

Foram selecionadas sondas não-específicas para espécie para a localização de sequências conservadas, incluindo as repetições teloméricas (TTAGG) $_n$ e o DNA ribossomal (DNAr) 18S. Também foram preparadas sondas específicas como a sonda de pintura do cromossomo W, sondas de DNAsat e sondas derivadas de TEs. Para os TEs, selecionamos a superfamília mais representativa no genoma da espécie e a mais diferencialmente enriquecida no genoma da fêmea em comparação ao macho. Enquanto que para os DNAsat, selecionamos as cinco famílias mais abundantes no genoma de *D. saccharalis*, como também, as sequências enriquecidas no genoma da fêmea. As sondas foram isoladas por PCR (Reação em cadeia da polimerase), e em seguida, mapeadas por FISH nos cromossomos da mariposa praga *D. saccharalis*.

A sonda de DNAr 18S foi obtida utilizando os *primers* descritos por Cabral-de-Mello *et al.* (2011). Para amplificação de DNAsat e TEs foram desenhados manualmente *primers* ou utilizando o software Geneious com a sequência consenso de cada DNAsat e cada elemento de TE. A repetição telomérica de inseto foi amplificada seguindo o protocolo de Ijdo *et al.* (1991) por PCR sem molde (*non-template DNA*) usando *primers* auto-complementares (TTAGG) $_5$ e (CCTAA) $_5$. O mix de PCR contém: 10x PCR Rxn Buffer, 0.2mM MgCl $_2$, 0.16mM dNTPs, 2mM de cada *primer*, 1 U de Taq Platinum DNA Polymerase (Invitrogen, San Diego, CA, Estados Unidos) e 50-100ng/ μ L de DNA. As condições de PCR incluem a desnaturação inicial a 94°C por 5 minutos e 30 ciclos a 94°C (30 segundos), 55°C (30 segundos) e 72°C (80 segundos), seguido de uma extensão final

a 72°C por 5 minutos. Os produtos resultantes da PCR são separados em gel de agarose 1%. As sondas foram marcadas com digoxigenina 11-dUTP (Roche, Mannheim, Alemanha) ou biotina 14-dATP (Invitrogen) através da nick-translation. Para sete DNAs satélites utilizamos oligonucleotídeos sintéticos diretamente marcados com biotina-14-dATP na extremidade 5' que foram desenhados em regiões conservadas e exclusivas de cada DNAsat.

O DNA genômico extraído de lagartas macho e fêmea foram marcados por nick-translation com digoxigenina 11-dUTP (Roche) para DNAg derivado de macho ou biotina 14-dATP (Invitrogen) para DNAg derivado de fêmea ou vice-versa. Estas sondas derivadas do DNAg foram utilizadas para realizar o CGH (Hibridização genômica *in situ* comparativa) de acordo com o protocolo descrito por Traut *et al.* (1999), com modificações propostas por Daliková *et al.* (2017a). A pintura cromossômica do cromossomo W foi preparada conforme descrito por Fuková *et al.* (2007). As amostras da cromatina-W microdissectada foram submetidas a amplificação por PCR utilizando “GenomePlex Single Cell Whole Genome Amplification Kit – WGA4” (Sigma-Aldrich), seguido para reamplificação e marcação com digoxigenina 11-dUTP (Roche) ou biotina 14-dATP (Invitrogen) por PCR usando o kit “GenomePlex – WGA3” (Sigma-Aldrich), de acordo com recomendações do fabricante. A sonda de pintura da cromatina-W obtida da espécie-irmã *D. postlineella* foi descrita por Cabral-de-Mello *et al.* (2021). Em combinação com a sonda-W também foi utilizada a sequência telomérica para checagem da presença ou não de sinais teloméricos intersticiais.

Lâminas não coradas de preparações mitóticas e meióticas obtidas de discos da asa, testículos e ovários, foram usadas para FISH simples ou dupla realizada de acordo com o protocolo descrito em Cabral-de-Mello e Marec (2021). As sondas marcadas com digoxigenina 11-dUTP foram detectadas por anti-digoxigenin-Rhodamine (Roche),

assim como as sondas marcadas com biotina 14-dATP foram detectadas usando Alexa fluor 488-conjugated (Invitrogen). Em todos os experimentos de FISH, os cromossomos foram contracorados com DAPI (4',6-diamidino-2-phenylindole, Sigma-Aldrich) e as lâminas montadas em VECTASHIELD (Vector, Burlingame, CA, Estados Unidos). As imagens foram capturadas utilizando uma câmera digital refrigerada DP70 acoplada a um microscópio Olympus BX61 equipado com uma lâmina de fluorescência e filtros apropriados. Posteriormente, as imagens foram pseudocoloridas, sobrepostas (sinais das sondas e cromossomos) e otimizadas com brilho e contraste pelo programa Adobe Photoshop CC6. Para caracterizar os padrões de distribuição de DNAs repetitivos foram analisados pelo menos 10 paquítenos.

4. RESULTADOS E DISCUSSÃO

Os resultados obtidos durante o mestrado serão apresentados e discutidos no manuscrito intitulado “Repeated DNAs populating the sugarcane borer *Diatraea saccharalis* (Crambidae) genome with an emphasis on sex chromosomes: a cytogenomic analysis” a ser submetido para a revista *Insect Molecular Biology*, INSS: 1365-2583.

4.1. Manuscrito

Repeated DNAs populating the sugarcane borer *Diatraea saccharalis* (Crambidae) genome with an emphasis on sex chromosomes: a cytogenomic analysis

Ana E Gasparotto¹, Diogo Milani¹, Emiliano Martí¹, Ana BSM Ferretti¹, Vanessa B Bardella¹, Magda Zrzavá^{2,3}, Frantisek Marec², Diogo C Cabral-de-Mello^{1,*}

¹UNESP - Univ Estadual Paulista, Instituto de Biociências /IB, Departamento de Biologia Geral e Aplicada, Rio Claro, São Paulo, Brazil

²Biology Centre, Czech Academy of Sciences, Institute of Entomology, Branišovská 31, 370 05 České Budějovice, Czech Republic

³Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

**Corresponding author:* UNESP – Univ Estadual Paulista, Instituto de Biociências/IB, Departamento de Biologia Geral e Aplicada – CEP 13506-900, Rio Claro, São Paulo, Brazil

Phone/Fax: +55 19 3526-4152, e-mail: cabral.mello@unesp.br

Abstract

Crambidae lepidopterans include pests that causes economic losses to agricultural crops, as the sugarcane borer *Diatraea saccharalis*. Although the crambid moths have extensive adaptive radiation and economic importance, little is known about their genome architecture and chromosomal evolution. Here we characterized the chromosomes and repetitive DNAs of *D. saccharalis* by association of low pass genome sequencing, bioinformatic and cytogenetic methods, with an emphasis on sex chromosomes. Diploid numbers differed between sexes, i.e., $2n=33$ females and $2n=34$ males, and occurrence of a multiple sex system $WZ_1Z_2/Z_1Z_1Z_2Z_2$ was noticed. Telomeric probe revealed one strong interstitial signal on the W chromosome, and this chromosome was associated to two chromosomes Z (Z_1 and Z_2), confirming the occurrence of a sex trivalent WZ_1Z_2 , as result of a chromosomal fusion. Concerning the repetitive DNAs, the TEs represented about 39,18% (males) to 41,35% (females), while satDNAs comprised only 0,214% of male genome and 0,215% of female genome. FISH mapping revealed distinct chromosomal organization for satDNAs, as single located clusters, spread signals and non-clustered repeats. The two TEs mapped by FISH were scattered. Although we noticed slightly enrichment of some satDNAs on female genome, they were not differentially enriched on the W chromosome. On the other hand, we noticed enrichment of FISH signals on the W chromosome for TEs, suggesting their involvement on W chromosome degeneration and differentiation. These data reveal dynamism of general karyotype structure and repetitive DNAs of *D. saccharalis*, mainly as result of chromosomal fusions and spread or local amplification of repetitive DNAs. Finally, our data can help in future structural and functional genomic analysis that could be also useful in pest control programs.

Key-words: chromosome fusion, FISH, holocentric chromosome, W-chromatin

Introduction

Repetitive DNAs are sequences repeated hundreds or thousands of times, comprising the major proportion of all nuclear DNA in eukaryotic genomes. These sequences may be present in long in tandem arrays of similar motifs i.e., satellite DNAs (satDNAs), and some multigene families, while others, may be dispersed through the genome as the transposable elements (TEs), with a mid to high rate of repetitiveness. All types of repetitive sequences can be enriched and located at specific chromosomes domains, mainly the heterochromatic ones, including centromeres and telomeres and are important organizational and functional components of the genomes (Charlesworth *et al.* 1994; López-Flores and Garrido-Ramos 2012; Biscotti *et al.* 2015). Additionally, the accumulation, expansion and lability of repetitive DNAs contribute to formation and differentiation of the heteromorphic sex chromosomes, as described in multiple species from most eukaryote groups (e.g., Charlesworth *et al.* 1994; Kejnovsky *et al.* 2013; Hobza *et al.* 2015; Chalopin *et al.* 2015; Palacios-Gimenez *et al.* 2017; Daliková *et al.* 2017b; Ferretti *et al.* 2020).

Lepidoptera (moths and butterflies) is one highly diverse order of insects with more than 160,000 valid species (Van Nieuwerkerken *et al.* 2011). At karyotypic point of view, Lepidopteran chromosomes are numerous (with a modal number $2n = 60-62$), small and uniform, without any landmarks, such as centromeres, namely holocentric chromosomes (Robinson 1971; Traut *et al.* 2007; Lukhtanov 2000; Sahara *et al.* 2012). This order represents the largest group with female heterogamety, with most species harboring WZ/ZZ sex chromosome system. This differs from the most other insect groups that presents male heterogamety, except Lepidoptera sister order Trichoptera (caddisflies) in which the female is also heterogametic (Traut *et al.* 2007; Marec *et al.* 2010; Sichová *et al.* 2015). Chromosomal rearrangements, like fusions, fissions, and inversions enabled

dynamic karyotype evolution with occurrence of some variable diploid numbers and sex chromosome system in Lepidoptera (Marec *et al.* 2001; Yasukochi *et al.* 2006; Sahara *et al.* 2007, 2012; Lukhtanov 2015; Sichová *et al.* 2015; Hill *et al.* 2019). This is especially evident in some groups, like in *Agrodiaetus* (Lycaenidae) $n = 10$ to $n = 134$ (Kandul *et al.* 2007); *Godyris* (Nymphalidae) $n = 13$ to $n = 120$ (Brown *et al.* 2004); *Lysandra* (Lycaenidae) $n = 24$ to $n = 93$ (Talavera *et al.* 2013). In *Leptidea* (Pieridae), besides large diploid number variability ($n = 28$ to $n = 103$), it was noticed variable multiple sex chromosome systems, including $W_{1-3}Z_{1-4}/Z_{1-4}Z_{1-4}$, $W_{1-3}Z_{1-3}/Z_{1-3}Z_{1-3}$, $W_{1-4}Z_{1-4}/Z_{1-4}Z_{1-4}$, $W_{1-3}Z_{1-6}/Z_{1-6}Z_{1-6}$ (Sichová *et al.* 2015, 2016; Yoshido *et al.* 2020).

Repetitive DNAs could be involved in genome organization and in its plasticity, as they could be frequent chromosomal rearrangement sites. Despite the large number of chromosomal studies developed in Lepidoptera, their great chromosomal variability and the extensive growing in the genome sequencing in this group, the understanding of some repetitive DNA fraction chromosomal organization remains elusive, mainly for the satDNAs and TEs. Until now, most cytogenetic studies mapped mainly multigene families, including rDNAs and H3 histone genes (Nguyen *et al.* 2010; Provazníková *et al.* 2021) and only 12 satDNAs were identified (Cabral-de-Mello *et al.* 2021). These studies allowed generation of landmarks for comparative analysis, that are useful for understanding genomic organization and changing patterns between species. Specially for sex chromosomes, although it is well known that the W chromosome is largely heterochromatic and enriched of DNA repeats, the specific content of this chromosome was poorly identified, with description of some TEs, satDNAs and in the lesser extent rDNA (Abe *et al.* 2005; Traut *et al.* 2013; Vechtová *et al.* 2016; Dalíková *et al.* 2017b; Zrzavá *et al.* 2018; Cabral-de-Mello *et al.* 2021).

The grass moth family Crambidae includes serious pests species that causes several economic losses to agricultural crops (Robinson 1971). Despite the crambid moths have extensive adaptative radiation and economic importance, very little is known about their genome architecture because only a few cytogenetic studies were performed. Among the only few species studied, the diploid number ranged from $n = 17$ to $n = 31$, the WZ/ZZ sex chromosome system was determined for *Ostrinia scapularis* and a FISH (fluorescence *in situ* hybridization) based map was constructed for *O. nubilalis* (Virkki 1963; Zhangwu *et al.* 1994; Kageyma and Traut 2004; Nguyen *et al.* 2010; Yasukochi *et al.* 2016; Cabral-de-Mello *et al.* 2021). Recently, by combination of cytogenetics and genomic approaches in four Crambidae species Cabral-de-Mello *et al.* (2021) revealed differences in gross genome architecture (diploid chromosome numbers) and high variability for repetitive DNAs, that putatively contributed to the plasticity of sex chromosomes in the group.

Diatraea saccharalis, commonly known as sugarcane borer, is one of the most important crambid pests and it is responsible for considerable crop damage and economic losses in maize and sugarcane (Robinson 1971; Flynn *et al.* 1984). This moth is widely distributed throughout the Americas, including Caribbean, South and Central America and the southeastern United States (Myers 1935; Robinson 1971; Palacio Cortés *et al.* 2010). An example of economic losses associated to *D. saccharalis* is observed in Brazil where this species become major pest of sugarcane, causing losses around US\$ 4,6 billion/year in the ethanol and sugar productions, representing more than 10% of the production (Oliveira *et al.* 2014). Due to this economic importance, *D. saccharalis* has attract attention of scientific community for population, structural and functional genetic studies (e.g., Joyce *et al.* 2014; Lopes *et al.* 2014; Merlin and Cônsoli 2019; Noriega *et al.* 2020; Borges dos Santos *et al.* 2020), but there is a complete lacking about the understanding of the genome architecture at chromosomal level in the species, except for

the chromosome number that is $n = 17$ (Virkki 1963). In this way, we characterized the repeatome (repetitive fraction of the genome) of the sugarcane borer by association of cytogenetic and bioinformatic methods, allowing the detailed description of general karyotype, sex chromosomes and chromosomal architecture of repetitive DNAs. The data contribute to knowledge about genomic organization and evolution of sex chromosomes in representatives of Crambidae and can help in future structural and functional genomic analysis in *D. saccharalis* that could be also useful in pest control programs.

Materials and Methods

Insects and chromosome preparations

Diatraea saccharalis individuals were obtained from lab culture of Luiz de Queiroz College of Agriculture, University of São Paulo – ESALQ/USP (Piracicaba, São Paulo, Brazil). The larvae were feed an artificial diet and maintained in B.O.D (Bio-Oxygen Demand) incubator (Solab, Piracicaba, Brazil) at controlled conditions (temperature $25 \pm 4^{\circ}\text{C}$, 60% humidity and 12 h light/12 h dark regime) (see King *et al.* 1975).

Spread chromosome preparations were made as described by Mediouni *et al.* (2004) and Sichová *et al.* (2013). Meiotic chromosomes in the pachytene stage were obtained from testes of the third larvae instar and ovaries at last larvae instar and in initial stage of impupation. Mitotic chromosomes were obtained from wings imaginal disks of the third instar of both sexes. Summarily, testes and wings imaginal disks were dissected in physiological solution, hypotonized for 10 min in 75mM KCl and fixed in Carnoy's solution (ethanol/chloroform/acetic acid, 6:3:1) for 15 min. Ovaries were transferred to Carnoy's solution without hypotonic treatment to preserve the putative pattern of W chromosome heterochromatin. The fixed tissues were macerated on a slide in a drop of 60% acetic acid and then spread at 45°C using a heating plate. The slides with sufficient

quality were dehydrated in an ethanol series (70%, 80% and 100% for 30 s each) and stored at -20 °C until use. All bodies of dissected larvae were stored in 100% ethanol in the freezer at -20 °C until the extraction of genomic DNA (gDNA).

Preparation of polyploid nuclei and microdissection of the sex chromatin

The Malpighian tubules from last instar larvae of female and male individuals were used to obtain polyploid interphase nuclei. After dissection in physiological solution, the Malpighian tubules were briefly fixed in Carnoy fixative on a slide for about 1 min and then stained with 1.25% lactic acetic orcein for 5 min. The slides were observed in a light microscope for the analysis of presence or absence of female specific sex chromatin (Traut and Marec 1996, 1997).

For sex chromatin microdissection we used Malpighian tubules dissected from female larvae in last instar, swollen for 10 min in a hypotonic solution (75mM KCl) and fixed in Carnoy's fixative for 15 min. The material was dissociated with tungsten needles in a drop of 60% acetic acid and spread on coverslips in a heating plate at 40°C and stained with 5% Giemsa for 7 min. A total of 15 W-chromatin bodies were microdissected manually using an Eppendorf 5171 micromanipulator coupled with a Nikon Axiphot inverted microscope.

Low pass genome sequencing, satDNAs and TEs identification and analysis

The gDNA was extracted from one female larva and one male larva on third instar using the "Wizard Genomic DNA Purification Kit" (Promega), following the manufacturer's protocol. The genomes of both sexes of *D. saccharalis* were sequenced via paired-end sequencing (2 x 150pb) using the Illumina HiSeq 4000 system, performed by Novogene (HK) Co., Ltd. (Hong Kong, China). The reads were processed, filtered and interlaced

using the tools available on the Galaxy/RepeatExplorer platform (Novák *et al.* 2013). Identification of satDNAs were performed separately using 500,000 reads for each sex as input. Succeeding the clustering by RepeatExplorer, the satDNAs identified by TAREAN tool (Novák *et al.* 2017) were selected for subsequent analysis. Furthermore, the tandem arrangement was checked by dotplot and Tandem Repeats Finder, TRF (Benson 1999). RepeatMasker (Smit *et al.* 2017) and the “rm_homology.py” python script (https://github.com/fjruizruano/satminer/blob/master/rm_homology.py) were used to performed all-against-all comparison of monomers of the recovered satDNAs for analysis of similarity and definition of all satDNA families.

We also used RepeatMasker to calculate abundance for each satDNA, comparing the female and male genomes. To estimate the average Kimura 2-parameter distances (K2P) we used the calcDivergenceFromAlign.pl script from RepeatMasker utility tool (Smit *et al.* 2017). For this we randomly selected about 2.65 million previously pre-processed read pairs per library obtained by seqtk (<https://github.com/lh3/seqtk>) and aligned them against dimers of consensus satDNA sequences. For normalization, the abundance of each satDNA family was estimated as genome proportion, calculated by the number of mapped nucleotides in respect to the library size. In addition, we compared the divergence of satDNAs between sexes generating repeat landscapes represent the relative abundance of repeat elements on one axis and the K2P distance from the consensus on the other axis. The satDNAs families were named by abundance considering the mean value of abundance between sexes. The possible abundance bias of satDNAs putatively associated to the W chromosome of females were estimated by the female/male abundance ratio for each satDNA family.

We identified and quantified TEs in both sexes by using dnaPipeTE (Goubert *et al.* 2015). dnaPipeTE relies on Trinity (Grabherr *et al.* 2011) to assembly the most abundant

sequences (i.e., repetitive sequences), followed by a homology-based characterization with Blast (Goubert *et al.* 2015). For the automated homology-based annotation of contigs, we used a custom TEs library as described elsewhere (Martí *et al.* 2021). TEs superfamilies were then aligned back to Illumina libraries from both sexes with RepeatMasker (Smit *et al.* 2017), the K2P divergence and normalization for both sexes were done as described above. To detect TEs superfamilies potentially related to the W chromosome, we estimated a female/male abundance ratio for each TE group.

DNA probes and chromosomal mapping by fluorescence *in situ* hybridization (FISH)

We selected non-species-specific probes to localize conserved sequences, including telomeric repeats and the major ribosomal DNA (18S rDNA). We also prepared specific probes, like W-chromosome painting probe, gDNA probes from male and female, satDNAs probes and TEs-derived probes. For TEs, we selected the most representative superfamily in the genome of the species and the most differentially enriched in female library in comparison to male library. While for satDNAs we selected the five most abundant families in the genome of *D. saccharalis* and also the sequences enriched in female genome. These probes were isolated and mapped by FISH in the chromosomes of *D. saccharalis*.

18S rDNA and satDNAs probes were amplified by PCR from female gDNA and labeled through nick-translation with biotin-14-dATP (Invitrogen, San Diego, CA, USA). The 18S rDNA probe was obtained using the primers described by Cabral-de-Mello *et al.* (2011). For satDNA and TEs amplification we designed primers manually or using Geneious software with the consensus sequence of each satDNA (Supplementary Table 1). PCR were carried out using the mix 10x PCR Rxn Buffer, 0.2mM MgCl₂, 0.16mM

dNTPs, 2mM of each primer, 1 U of *Taq* Platinum DNA Polymerase (Invitrogen) and 50-100 ng/μL of template DNA. The conditions of PCR included an initial denaturation at 94°C for 5 min and 30 cycles at 94°C (30 s), 55°C (30 s) and 72°C (80 s), plus a final extension at 72°C for 5 min. PCR products were separated at 1% electrophoresis agarose gel. For seven of the satDNAs we used synthetic oligonucleotides directly labeled with biotin-14-dATP at 5'end designed in conserved and exclusive regions of the satDNAs (Supplementary Table 1).

Genomic DNA extracted from male and female larvae were labeled by nick-translation with digoxigenin-11-dUTP (Roche, Mannheim, Germany) for male DNA or biotin-14-dATP (Invitrogen) for female DNA or vice versa. These labeled gDNAs probes were used for CGH performed according the protocol describe in Traut *et al.* (1999), with modifications proposed by Daliková *et al.* (2017a). W-chromosome painting probes were prepared as described in Fuková *et al.* (2007). The samples of manually microdissected W-chromatin were subjected to amplification by PCR using GenomePlex Single Cell Whole Genome Amplification Kit, WGA4 (Sigma-Aldrich). The amplified product was followed for re-amplification and labeled with either digoxigenin-11-dUTP (Roche) or biotin-14-dATP (Invitrogen) by PCR using the GenomePlex WGA Reamplification Kit, WGA3 (Sigma-Aldrich). The W chromosome painting probe from *D. postlineella* obtaining is described by Cabral-de-Mello *et al.* (2021). We also combined the W-painting with a (TTAGG)_n telomeric probe for checking the presence or absence of interstitial telomere signals. Insect telomeric probe was synthetized by non-template PCR according to Ijdo *et al.* (1991), using self-complementary primers (TTAGG)₅ and (CCTAA)₅, and labeled with digoxigenin-11-dUTP (Roche) through nick-translation.

Unstained slides obtained from ovaries were used to single or two-color FISH performed according to the protocol described by Cabral-de-Mello and Marec (2021).

Probes labeled with digoxigenin-11-dUTP were detected using anti-digoxigenin-Rhodamine (Roche), as probes labeled with biotin-14-dATP were detected with streptavidin Alexa-fluor 488-conjugated (Invitrogen). In all FISH experiments, chromosomes were counterstained with DAPI (4',6-diamidino-2-phenylindole; Sigma-Aldrich) and slides mounted in VECTASHIELD (Vector, Burlingame, CA, USA). Pictures were captured using a DP70 cooled digital camera in grayscale coupled to the Olympus microscope BX51 equipped with a fluorescence lamp and appropriate filters. The images were pseudocolored, merged (chromosomes and probe signals) and optimized with brightness and contrast using Adobe Photoshop CC6. To characterize the patterns of repetitive DNAs distribution at least 10 pachytene were analyzed.

Results

General characteristics and the multiple sex chromosomes of *Diatraea saccharalis*

Diploid chromosome numbers were determined from analyzing and repeated counts of pachytene and mitotic metaphase prepared from testes, ovaries and wings imaginal disks, besides nurse cells from ovaries. To facilitate the identification of individual chromosomes, we used FISH with telomeric probes. The karyotypes consisted of various-sized chromosomes, as large, medium and small. We found differences in chromosome numbers between sexes, i.e., $2n = 34$ males (Figure 1a) and $2n = 33$ females (Figure 1b). In some mitosis of both sexes, it was possible to recognize two small chromosomes with a constriction, which corresponds to the cluster of 18S rDNA (Figure 1c). None more condensed region was observed in pachytenes, suggesting the low abundance of heterochromatin and the non-enrichment of this repetitive fraction in specific regions of the genome. The exception was one chromosome highlighted by DAPI, that corresponds to the W chromosome, clearly seen in pachytene from oocytes (Figure 1d). Besides the

occurrence of canonical telomeres on terminal regions of chromosomes it was observed in the putative W chromosome clearly the occurrence of one interstitial signal (Figure 1b, d-h). The difference in diploid number among males and females indicated the occurrence of a multiple sex chromosome system in the species, formed by one W chromosome and two Z chromosomes (Z_1 and Z_2), i.e., WZ_1Z_2 system. The multiple sex system and interstitial telomeric site in the W chromosome was confirmed by more detailed analysis (see below).

Preparations of polyploid nuclei of the Malpighian tubules from females showed a single conspicuous sex chromatin body with regular spherical shape, confirming the presence of the W chromosome. As expected, the sex chromatin was absent in the polyploid nuclei of Malpighian tubules from males. Moreover, non-additional condensed blocks of heterochromatin were observed (Supplementary Figure 1).

The CGH used in order to investigate and characterize the gross molecular differentiation of W and Z chromosomes revealed more intense signal on the W chromosome from female gDNA derived probe in comparison to male gDNA derived probe, revealing enrichment of females' sequences on W chromosome. No difference of signal intensity was observed along the W chromosome (Figure 2a-e).

W-chromosome painting probe highlighted intensely the entire W chromosome, allowing easy identification of this sex chromosome and confirming molecular differentiation of the W from other chromosomes, as well Z chromosome (Figure 2f). Furthermore, W-chromosome painting combined with telomeric probe revealed one strong interstitial signal in the W chromosome (coincident with the painted region from W-derived probe). Near to the interstitial telomeric site of W chromosome we observed two additional telomeric signals, correspondent to the end of the two Z chromosomes, confirming the occurrence of a sex trivalent WZ_1Z_2 (Figure 2g-l). In nurse cells it was

possible to see clearly the W chromosome associated to two Z chromosomes, that are slightly different in size (Figure 2f, inset).

The cross-hybridization with W-painting probe derived of the W-chromatin from the sister-species, *Diatraea postlineella* revealed hybridization signals along the entire extension of the W chromosome of *D. saccharalis*. The intensity of signal was similar on entire W chromosome but, as expected, much less intense when compared with signals obtained from W-chromosome painting probe obtained from *D. saccharalis* (Figure 2m-o).

Genomic and chromosomal characterization of satDNAs and TEs

By clustering analysis in RepeatExplorer using the TAREAN report, we identified a total of fourteen putative satDNAs on *D. saccharalis*, including male and female genomic libraries. The searching in the RepBase/GIRI and NCBI databases did not find any relevant similarity with the satDNAs sequences described with other repetitive elements already described. These repetitive DNA sequences represent only a small portion of the male and female genomes, 0.214 to 0.215%, respectively. Monomers of these satDNAs were highly variable in size, ranging from 12 to 1,327 bp, in addition the sequences were mostly A+T rich content, ranging from 39,4 to 72,5% (Table 1). Considering the mean value between sexes the abundance of the distinct satDNA families on *D. saccharalis* genome ranged from 0.0002% to 0.0570% and K2P distance from 3,96% to 17,85%. The satellitome landscape (abundance vs. divergence) revealed a higher peak of abundance with low divergence (around zero) and a smaller peak of abundance between 6-10% of divergence. Repeats with higher divergence tended to decrease in amount from 10% of K2P (Figure 3a).

Among the fourteen satDNAs eight families (DsaSat01-30, DsaSat04-901, DsaSat07-294, DsaSat08-338, DsaSat10-84, DsaSat11-12 and DsaSat13-134) were slightly enriched in female genome (Figure 3b), suggesting enrichment on the W chromosome. Although, through FISH the enrichment of these repeats on W chromosome was discarded, as for the DsaSat01-30 (Figure 3c) and DsaSat04-901 (Figure 3e) one autosomal cluster were observed, terminally and interstitially located, respectively. The other female enriched satDNAs (DsaSat07-294, DsaSat08-338, DsaSat10-84, DsaSat11-12 and DsaSat13-134) did not revealed visible FISH signals. FISH signals were absent also for DsaSat03-32 (result not shown). Finally, the DsaSat02-774 and DsaSat05-1327 revealed scattered signals in all chromosomes with no evidence of enrichment in specific chromosomes or chromosomal areas (Figure 3d,f).

The TEs were much more abundant in the genome of *D. saccharalis* than the satDNAs, corresponding to about 41,35% of female genome and 39,18% of male genome, comprising at least 123 elements, considering the few unknown (not annotated) elements identified. LINEs and Helitron were the most representative TEs (about 10% each) followed by LTR that represented about 6% of *D. saccharalis* genome. On the other hand, the retrotransposons and Rolling Circle (RC) TEs were extremely low represented. The number of elements per distinct types of TEs were correlated with their abundance, i.e., TEs with higher abundance presented higher number of elements and vice versa (Table 2, Supplementary Table 2). For distinct classes of TEs almost no differences for abundance were noted between male and female genomes and number of elements (Table 1). Although, for four elements it was noticed higher enrichment in female genome, with abundance ratio > 2 , i.e., Helitron-2, TcMarStowaway, Maverik, and TcMar-Sagan (Supplementary Table 2). The comparative TEs landscape between female versus male genomes evidences clearly the recent amplification of some TEs on female genome, with

occurrence of some elements with higher abundance with low divergence (Figure 4a). The FISH mapping of the most abundant TE, i.e., Helitron (mean genome abundance of 10,774%) and for the TcMar-Stowaway DNA TE elements, that is enriched in female genome, revealed similar patterns on pachytene chromosomes from oocytes. The both TEs were virtually distributed along the entire extension of all chromosomes as small dots, but it is clearly noticed more intense signals on the W chromosome in comparison to other chromosomes, especially in comparison to the Z chromosome (Figure 4b-e).

Discussion

In this study, we made a progress on the structural knowledge of genome characteristics by physical characterization of the karyotype and repetitive DNAs of *D. saccharalis*, contributing to the genomic understanding of this important pest species and also in the understanding of chromosome evolution in Crambidae moths. Moreover, we present advance in the understanding of repetitive DNA composition of sex chromosomes in Lepidoptera, revealing occurrence of satDNAs and enrichment of TEs on the heterochromatic W chromosome.

While most species of Lepidoptera harbor $2n = 62$ chromosomes (Robinson 1971), on Crambidae representatives some examples of chromosome number reduction, as noticed here for *D. saccharalis* were described, that are putatively resultant of fusion events (Cabral-de-Mello *et al.* 2021). This is especially evident on the two *Diatraea* species studied, that present extensive diploid number reduction, *D. postlineella* $2n = 42$ (Cabral-de-Mello *et al.* 2021) and *D. saccharalis* $2n = 34$ (males) and $2n = 33$ (females) (this work). Although, based in the ancestral Lepidopterans diploid number it is supposed occurrence of chromosomal fusions on these species, interstitial telomeric sites were not observed, suggesting elimination of telomeric repeats. Concerning rDNA, among the six

Pyraloidea (Pyralidae and Crambidae) species in which this sequence was mapped four presented terminal sites (Nguyen *et al.* 2010; Cabral-de-Mello *et al.* 2021; Provazníková *et al.* 2021). The unique exceptions are the two species of *Diatraea* with interstitial site in one small chromosome pair (Cabral-de-Mello *et al.* 2021; this work). Based in this data, we could suggest that the chromosome bearing the rDNA cluster was involved in a fusion in *Diatraea*, causing its placement on interstitial position, although, a simple inversion could not be ruled out, as proposed to other few Lepidoptera species (Nguyen *et al.* 2010). Moreover, ectopic recombination, a mechanism involved in rDNA chromosomal repatterning in insects (Nguyen *et al.* 2010; Cabral-de-Mello *et al.* 2011; Panzera *et al.* 2012; Gunderina *et al.* 2015; Bardella *et al.* 2016; Ferretti *et al.* 2020), could also be occurred.

Concerning the sex chromosomes, the observation of interstitial telomeric site, was noticed for the W chromosome of *D. saccharalis*, resultant from a chromosomal fusion involving an autosomal pair and ancestral sex chromosome, generating a neo-sex system, a mechanism previously noticed in Lepidoptera (Marec *et al.* 2010; Yoshido *et al.* 2011; Sahara *et al.* 2012) and other insects (Blackmon *et al.* 2017). It is parsimonious to claim that diploid number reduction to $2n = 42$ occurred in the common ancestor of *D. postlineella* and more fusions occurred independently on the karyotype of *D. saccharalis*, that could also be shared with other species of the genus, deserving more analysis. For *D. saccharalis*, a trivalent exclusively observed on meiotic cells of oocytes explain the sex-specific difference in chromosome numbers as result of the presence of a multiple sex chromosome system formed by one W chromosome and two Z chromosomes i.e., $W_1Z_1Z_2$ (female)/ $Z_1Z_1Z_2Z_2$ (male). Although, more commonly Lepidoptera species harbor WZ/ZZ or Z0/ZZ sex chromosome system (Traut *et al.* 2007; Marec *et al.* 2010), the occurrence of neo-sex chromosome multiple system was also reported sporadically in multiple

species, belonging to distinct evolutionary branches (Marec *et al.* 2010; Nguyen and Carabajal Paladino 2016). Distinct types of neo-sex chromosomes in Lepidoptera, including the $W_1Z_1Z_2$ (female)/ $Z_1Z_1Z_2Z_2$ (male) observed in *D. saccharalis*, could have an important role in adaptive potential, contribute to ecological specialization and formation of reproductive barriers and ultimately speciation (Yoshido *et al.* 2011, 2020; Nguyen *et al.* 2013; Sichová *et al.* 2015; Carabajal Paladino *et al.* 2019). One exceptional example was described on Tortricidae species, in which the origin of a neo-sex system involved an autosome enriched of genes that confers insecticide and plant metabolite detoxification, linking these genes to sex inheritance and increasing the adaptive potential of leafrollers tortricid moths (Nguyen *et al.* 2013). In *D. saccharalis* it is especially important to check the composition of the autosome involved in the generation of multiple neo-sex chromosome system, to check the possible presence of genes on this chromosome that could be involved with performance of the species as sugar-cane pest.

Our CGH analysis revealed that the W and Z in *D. saccharalis* almost fully differentiated, indicating the gross molecular differences in the DNA composition, revealing a more intense signal on the W from female gDNA derived probe. Enrichment of W chromosome on repeats derived from female genomic probe was evidenced for other Crambidae (Cabral-de-Mello *et al.* 2021) and other Lepidoptera (Yoshido *et al.* 2006; Daliková *et al.* 2017a; Hejníčková *et al.* 2021), including species with neo-sex chromosome system (Yoshido *et al.* 2011; Sichová *et al.* 2016; Carabajal Paladino *et al.* 2019). This mainly reflects an accumulation of repetitive sequences, TEs in the case of *D. saccharalis* (see results) and putatively satDNAs and/or TEs on the W chromosome in other species (Sahara *et al.* 2003; Mediouni *et al.* 2004; Abe *et al.* 2005; Daliková *et al.* 2017b). Furthermore, cross-hybridization experiment revealed that the W chromosome from *D. saccharalis* exhibit conservation at least of some repetitive sequences with its

congeneric species *D. postlineella*. Moreover, the fully painting of W chromosome of *D. saccharalis* with the W-probe from *D. postlineella* indicate that repetitive sequences from ancestral W shared between the two *Diatraea* species colonized the new part of the W chromosome of *D. saccharalis* (ancestor autosome involved in the fusion) during the heterochromatinization process. Certain degree of sequence conservation among Crambidae species were noticed (Cabral-de-Mello *et al.* 2021), as among Pyralidae species (Vítková *et al.* 2007), contrasting with congeneric *Abraxas* (Geometridae) species that showed a high molecular divergence between the W chromosomes (Zrzavá *et al.* 2018), suggesting that the rates of molecular differentiation of W chromosomes is variable between Lepidoptera groups.

The combined analysis of low genome pass sequencing, computational tools and chromosomal mapping, allowed a more precise understanding about the composition of the whole genome and especially the W sex chromosomes of *D. saccharalis*. The occurrence of satDNAs with low abundance in the genome of *D. saccharalis* (total of 0,214% of male genome and 0,215% of female genome) is apparently a common feature on Crambidae (Cabral-de-Mello *et al.* 2021), and it is in concordance with the low heterochromatic content observed in *D. saccharalis* and in most species of Lepidoptera in which higher concentration of heterochromatin mostly on the W chromosome (Traut *et al.* 2007). It could also be attributed to the small genome size of this group (Gregory 2020, Animal Genome Size Database) and to the predominance of interspersed elements, i.e., TEs (Talla *et al.* 2017). This last assumption is corroborated by the high amount of TEs identified in *D. saccharalis* genome, that corresponded to about 40% of the species genome. The amount of interspersed repeats identified here is similar to previously described in the first genome draft of *D. saccharalis* (Borges dos Santos *et al.* 2020).

The chromosome organization of the repeats identified here were variable, suggesting high dynamisms of these sequences at chromosomal level. It is evident by comparison of satDNAs that, (i) as in most Lepidoptera presented scattered signals decorating all chromosomes (Mahendran *et al.* 2006; Vechtová *et al.* 2016; Cabral-de-Mello *et al.* 2021), (ii) where located as single localized cluster as described in *D. postlineella*, *Plodia interpunctella* and *Mamestra brassicae* (Mandrioli *et al.* 2003; Daliková *et al.* 2017b; Cabral-de-Mello *et al.* 2021), and (iii) did not revealed evident signals on chromosomes, i.e., non-clustered pattern, or quite small clusters, bellow FISH resolution. This is the first intragenome variation for chromosomal distribution of satDNAs noticed for Lepidoptera, as the species studied until now, in which more than one satDNA was mapped, presented or exclusively scattered or single located signals (Cabral-de-Mello *et al.* 2021). This data suggests that satDNAs is more dynamic in Lepidoptera genomes, than previously thought, and although in low abundance could be important player in the genome repatterning on this group.

Although, a plenty of information were gathered about the W chromosome on Lepidoptera, giving insights about its evolution (Traut *et al.* 2007; Nguyen and Carabajal Paladino 2016), the specific molecular composition of this chromosome still largely elusive. Our analysis for the sex chromosomes revealed that although the W chromosome of *D. saccharalis* is heterochromatic, as in most Lepidoptera (Traut *et al.* 2007), it is not highly enriched on satDNAs in comparison to other chromosomes. Although we identified eight satDNAs slightly overrepresented on female genome in comparison to male these repeats did not reveal evident enrichment on W chromosome, and the genomic abundance variation observed could represent only interindividual variability, not sex related. On the other hand, it was quite evident the higher enrichment of the two TEs mapped by FISH on the W chromosome of *D. saccharalis* in comparison to other

chromosomes, decorating its entire extension. This reveal involvement of TEs in the degeneration of the W chromosome and its differentiation to the Z chromosomes in the species. Occurrence of TEs on W chromosomes were noticed in the lepidopterans, *Bombyx mori* (Abe *et al.* 2005) and *Ephestia kuehniella* (Traut *et al.* 2013).

Altogether, our data reveal dynamism of general karyotype and repetitive DNAs on the genome of *D. saccharalis*. This dynamism is resultant of chromosomal fusion events and spreading or local amplification of the distinct repetitive DNAs. The high incidence fusions in the karyotype of *D. saccharalis* resulted also in the emergence of the multiple neo-sex chromosome system $W_1Z_1Z_2$ (female)/ $Z_1Z_1Z_2Z_2$ (male), in which the W chromosome degenerated and differentiated to the rest of the genome essentially by the accumulation of TEs, giving new insights about evolution of sex chromosomes on Lepidoptera. Finally, as the repetitive DNAs is a difficult part of the genome to be assembled, our data will help for future improvements in the genome sequencing and assembly of this important pest species.

References

- Abe, H., Mita, K., Yasukochi, Y., Oshiki, T. and Shimada, T. (2005). Retrotransposable elements on the W chromosome of the silkworm, *Bombyx mori*. *Cytogenet. Genome Res* 110, 144-151. doi: 10.1159/000084946
- Bardella, V. B., Fernandes, J. A. M. and Cabral-de-Mello, D. C. (2016). Chromosomal evolutionary dynamics of four multigene families in Coreidae and Pentatomidae (Heteroptera) true bugs. *Mol Genet Genomics* 291, 1919-1925. doi: 10.1007/s00438-016-1229-5
- Benson, G. (1999). Tandem repeats Finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27, 573-580. doi: 10.1093/nar/27.2.573
- Biscotti, M. A., Olmo, E. and Heslop-Harrison, J. S. (2015). Repetitive DNA in eukaryotic genomes. *Chromosome Research* 23(3), 415-420. doi: 10.1007/s10577-015-9499-z
- Blackmon, H., Ross, L. and Bachtrog, D. (2017). Sex determination, sex chromosomes and karyotype evolution in insects. *J Hered.* 108(1), 78-93. doi: 10.1093/jhered/esw047
- Borges dos Santos, L., Viana, J. P. G., Francischini, F. J. B., *et al.* (2020). A first draft genome of the sugarcane borer, *Diatraea saccharalis*. *F1000 Research* 9, 1269. doi: 10.12688/f1000research.26614.1
- Brown, K. S., Von Schoultz, B. and Suomalainen, E. (2004). Chromosome evolution in Neotropical Danainae and Ithomiinae (Lepidoptera). *Hereditas* 141(3), 216-236. doi: 10.1111/j.1601-5223.2004.01868.x
- Cabral-de-Mello, D. C. and Marec, F. (2021). Universal fluorescence *in situ* hybridization (FISH) protocol for mapping repetitive DNAs in insects and other arthropods. *Molecular Genetics and Genomics* 296, 513-526. doi: 10.1007/s00438-021-01765-2
- Cabral-de-Mello, D. C., Oliveira, S. G., de Moura, R. and Martins, C. (2011). Chromosomal organization of the 18S and 5S rRNAs and histone H3 genes in *Scarabaeinae coleopterans*: insights into the evolutionary dynamics of
-

-
- multigene families and heterochromatin. *BMC Genet* 12(88). doi: 10.1186/1471-2156-12-88
- Cabral-de-Mello, D. C., Zrzavá, M., Kibicková, S., *et al.* (2021). The role of satellite DNAs in genomic architecture and sex chromosome evolution in Crambidae moths. *Frontiers in Genetics* 12:661417. doi: 10.3389/fgene.2021.661417
- Carabajal Paladino, L., Provazníková, I., Berger, M. *et al.* (2019). Sex chromosome turnover in moths of the diverse superfamily Gelechioidea. *Genome Biology and Evolution* 11(4), 1307-19. doi: 10.1093/gbe/evz075
- Chalopin, D., Volff, J. N., Galiana, D., *et al.* (2015). Transposable elements and early evolution of sex chromosomes in fish. *Chromosome Research* 23, 545-560. doi: <https://doi.org/10.1007/s10577-015-9490-8>
- Charlesworth, B., Sniegowski, P. and Stephan, W. (1994). The evolutionary dynamics of repetitive DNAs in eukaryotes. *Nature*, v. 371, p. 215-220. doi: 10.1038/371215a0
- Dalíková, M., Zrzavá, M., Hladová, I., *et al.* (2017a). New insights into the evolution of the W chromosome in Lepidoptera. *Journal of Heredity* 108(7), 709-719. doi: 10.1093/jhered/esx063
- Dalíková, M., Zrzavá, M., Kubicková, S. and Marec, F. (2017b). W-enriched satellite sequence in the Indian meal moth, *Plodia interpunctella* (Lepidoptera, Pyralidae). *Chromosome Research* 25, 241-252. doi: 10.1007/s10577-017-9558-8
- Ferretti, A. B. S. M., Milani, D., Palacios-Gimenez, O. M., *et al.* (2020). High dynamism for neo-sex chromosomes: satellite DNA reveal complex evolution in a grasshopper. *Heredity* 125, 124-137. doi: 10.1038/s41437-020-0327-7
- Flynn, J. L., Reagan, T. E. and Ogunwolu, E. O. (1984). Establishment and damage of the sugarcane borer (Lepidoptera: Pyralidae) in corn as influenced by plant development. *Journal of Economic Entomology* 77(3), 691-97. doi: 10.1093/jee/77.3.691
- Fuková, I., Traut, W., Vitková, M., Kubicková, S. and Marec, F. (2007). Probing the W chromosome of the codling moth, *Cydia pomonella*, with sequences from
-

-
- microdissected sex chromatin. *Chromosoma* 116, 135-45. doi: 10.1007/s00412-006-0086-0
- Goubert, C., Modolo, L., Vieira, C., *et al.* (2015). De novo assembly and annotation of the Asian tiger mosquito (*Aedes albopictus*) repeatome with dnaPipeTE from raw genomic reads and comparative analysis with the yellow fever mosquito (*Aedes aegypti*). *Genome Biology and Evolution* 7(4), 1192-1205. doi: 10.1093/gbe/evv050
- Grabherr, M. G., Haas, B. J., Yassour, M., *et al.* (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29(7), 644-52. doi: 10.1038/nbt.1883
- Gregory, T. R. (2020). *Animal genome size database*. Available online at: <http://www.genomesize.com>
- Gunderina, L., Golygina, V. and Broshkov, A. (2015). Chromosomal organization of the ribosomal RNA genes in the genus *Chironomus* (Diptera, Chironomidae). *Comparative cytogenetics* 9(2), 201-220. doi: 10.3897/CompCytogen.v9i2.9055
- Hejníčková, M., Dalíková, M., Potocký, P. *et al.* (2021). Degenerated, undifferentiated, rearranged, lost: high variability of sex chromosomes in Geometridae (Lepidoptera) indentified by sex chromatin. *Cells* 10(9), 2230. doi: 10.3390/cells10092230
- Hill, J., Rastas, P., Hornett, E. A., *et al.* (2019). Unprecedented reorganization of holocentric chromosomes provides insights into the enigma of lepidopteran chromosome evolution. *Science advances* 5(6), eaau3648. doi: 10.1126/sciadv.aau3648
- Hobza, R., Kubat, Z., Cegan, R., *et al.* (2015). Impact of repetitive DNA on sex chromosomes in plants. *Chromosome Research* 23:561. doi: 10.1007/s10577-015-9496-2
- Ijdo, J. M., Wells, R. A., Baldini, A. and Reeders, S. T. (1991). Improved telomere detection using a telomere repeat probe (TTAGGG)_n generated by PCR. *Nucleic Acids Research* 19(17), 4780. doi: 10.1093/nar/19.17.4780
-

-
- Joyce, A. L., White, W. H., Nuessly, G. S., *et al.* (2014). Geographic population structure of the sugarcane borer, *Diatraea saccharalis* (F.) (Lepidoptera: Crambidae), in the Southern United States. *PloS One* 9(10):e110036. doi: 10.1371/journal.pone.0110036
- Kageyama, D. and Traut, W. (2004). Opposite sex-specific effects of *Wolbachia* and interference with the sex determination of its host *Ostrinia scapularis*. *Proc. Biological sciences* 271(1536), 251-58. doi: 10.1098/rspb.2003.2604
- Kandul, N. P., Lukhtanov, V. A. and Pierce, N. E. (2007). Karyotypic diversity and speciation in *Agrodiaetus* butterflies. *Evolution* 61(3), 546-559. doi: 10.1111/j.1558-5646.2007.00046.x
- Kejnovsky, E., Michalovova, M., Steflöva, P., *et al.* (2013). Expansion of microsatellites on evolutionary young Y chromosome. *PloS One* 8(1): e45519. doi: <https://doi.org/10.1371/journal.pone.0045519>
- King, E. G., Brewer, F. D. and Martin, D. F. (1975). Development of *Diatraea saccharalis* [Lep.: Pyralidae] at constant temperatures. *Entomophaga* 20, 301-6. doi: 10.1007/BF02371955
- Lopes, D. A., Cantagalli, L. B., Stuchi, A. L. P. B., Mangolin, C. A. and Ruvolo-Takasusuki, M. C. C. (2014). Population genetics of the sugarcane borer *Diatraea saccharalis* (Fabr.) (Lepidoptera: Crambidae). *Acta Scientiarum Agronomy* 36(2). doi: 10.4025/actasciagron.v36i2.16211
- López-Flores, I. and Garrido-Ramos, M. A. (2012) The repetitive DNA content of eukaryotic genomes. *Genome Dyn* 7, 1-28. doi: 10.1159/000337118
- Lukhtanov, V. A. (2000). Sex chromatin and sex chromosome systems in nonditrysian Lepidoptera (Insecta). *J. Zool. Syst. Evol. Res* 38(2), 73-79. doi: 10.1046/j.1439-0469.2000.382130.x
- Lukhtanov, V. A. (2015). The blue butterfly *Polyommatus* (Plebicula) *atlanticus* (Lepidoptera, Lycaenidae) holds the record of the highest number of chromosomes in the non-polyploid eukaryotic organisms. *Comparative Cytogenetics* 9(4), 683-690. doi: 10.3897/CompCytogen.v9i4.5760
-

- Mahendran, B., Acharya, C., Dash, R., Ghosh, S. K. and Kundu, S. C. (2006). Repetitive DNA in tropical tasar silkworm *Antheraea mylitta*. *Gene* 370, 51-57. doi: 10.1016/j.gene.2005.11.010
- Mandrioli, M. Manicardi, G. C. and Marec, F. (2003). Cytogenetic and molecular characterization of the MBSAT1 satellite DNA in holokinetic chromosomes of the cabbage moth, *Mamestra brassicae* (Lepidoptera). *Chromosome Res.* 11, 51-56. doi: 10.1023/A:1022058032217
- Marec, F., Sahara, K. and Traut, W. (2010). Rise and fallo f the W chromosome in Lepidoptera. *In book: Molecular Biology and Genetics of the Lepidoptera*, CRC Press, 49-63. doi: 10.1201/9781420060201-c3
- Marec, F., Tothová, A., Sahara, K., *et al.* (2001). Meiotic pairing of sex chromosome fragments and its relation to atypical transmisión of a sex-linked marker in *Ephestia kuehniella* (Insecta: Lepidoptera). *Heredity* 87, 659-671. doi: 10.1046/j.1365-2540.2001.00958.x.
- Martí, E., Milani, D., Bardella, V. B., *et al.* (2021). Cytogenomic analysis unveils mixed molecular evolution and recurrent chromosome rearrangements shaping the multigene families on *Schistocerca* grasshopper genomes. *Evolution* 75(8), 2027-41. doi: 10.1111/evo.14287
- Mediouni, J., Fuková, I., Frydrychová, R., Dhouibi, M. H. and Marec, F. (2004). Karyotype, sex chromatin and sex chromosome differentiation in the carob moth, *Ectomyelois ceratoniae* (Lepidoptera: Pyralidae). *Caryologia* 57, 184-94. doi: 10.1080/00087114.2004.10589391
- Merlin, B. L. and Cônsoli, F. L. (2019). Regulation of the larval transcriptome of *Diatraea saccharalis* (Lepidoptera: Crambidae) by maternal and other factors of the parasitoid *Cotesia flavipes* (Hymenoptera: Braconidae). *Frontiers in Physiology* 10:1106. doi: 10.3389/fphys.2019.01106
- Myers, J. (1935). The ecological distribution of some South American Grass and sugar-cane borers (*Diatraea* spp., Lep., Pyralidae). *Bulletin of Entomological Research* 26(3), 335-42. doi: 10.1017/S0007485300036646
-

-
- Nguyen, P. and Carabajal Paladino, L. (2016). On the neo-sex chromosomes of Lepidoptera. In book: *Evolutionary Biology*. 171-185. doi: 10.1007/978-3-319-41324-2_11
- Nguyen, P., Sahara, K., Yoshido, A and Marec, F. (2010). Evolutionary dynamics of rDNA clusters on chromosomes of moths and butterflies (Lepidoptera). *Genetica* 138, 343-354. doi: 10.1007/s10709-009-9424-5
- Nguyen, P., Sýkorová, M., Sichová, J. *et al.* (2013). Neo-sex chromosomes and adaptive potential in tortricid pests. *Proc Natl Acad Sci USA* 110(17), 6931-36. doi: 10.1073/pnas.1220372110
- Noriega, D. D., Arraes, F. B. M., Antonino, J. D., *et al.* (2020). Comparative gut transcriptome analysis of *Diatraea saccharalis* in response to the dietary source. *PloS One* 15(8): e0235575. doi: 10.1371/journal.pone.0235575
- Novák, P., Neumann, P., Pech, J., Steinhaisl, J. and Macas, J. (2013). RepeatExplorer: a Galaxy-based web server for genoma-wide characterization of eukaryotic repetitive elements from next generation sequence reads. *Bioinformatics* 29, 792-3. doi: 10.1093/bioinformatics/btt054
- Novák, P., Robledillo, L. A., Koblizková, A., *et al.* (2017). TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Res* 45(12), e111. doi: 10.1093/nar/gkx257
- Oliveira, C. M., Auad, A. M., Mendes, S. M. and Frizzas, M. R. (2014). Grop losses and the economic impact of insect pests on Brazilian agriculture. *Crop Protection* 56, 50-54. doi: 10.1016/j.cropro.2013.10.022
- Palacio Cortés, A. M., Zarbin, P. H. G., Takiya, D. M., *et al.* (2010). Geographic variation of sex pheromone and mitochondrial DNA in *Diatraea saccharalis* (Fab., 1974) (Lepidoptera: Crambridae). *Journal of insect physiology* 56(11), 1624-30. doi: 10.1016/j.jinsphys.2010.06.005
- Palacios-Gimenez, O. M., Bardella, V. B., Lemos, B. and Cabral-de-Mello, D. C. (2017). Satellite DNAs are conserved and differentially transcribed among *Gryllus* cricket species. *DNA Res* 25(2), 137-147. doi: 10.1093/dnares/dsx044
-

-
- Panzer, Y., Pita, S., Ferreiro, M. J. *et al.* (2012). High dynamics of rDNA cluster location in kissing bug holocentric chromosomes (Triatominae, Heteroptera). *Cytogenet Genome Res.* 138(1), 56-67. doi: 10.1159/000341888
- Provazníková, I., Hejníčková, M., Visser, S., *et al.* (2021). Large-scale comparative analysis of cytogenetic markers across Lepidoptera. *Scientific reports* 11(1):12214. doi: 10.1038/s41598-021-91665-7
- Robinson, R. (1971). *Lepidoptera Genetics*. Pergamon Press Oxford, 1ed.
- Sahara, K., Yoshido, A. and Traut, W. (2012). Sex chromosome evolution in moths and butterflies. *Chromosome Research* 20, 83-94. doi: 10.1007/s10577-011-9262-z
- Sahara, K., Yoshido, A., Kawamura, N. *et al.* (2003). W-derived BAC probes as a new tool for identification of the W chromosome and its aberrations in *Bombyx mori*. *Chromosoma* 112, 48-55. doi: 10.1007/s00412-003-0245-5
- Sahara, K., Yoshido, A., Marec, F., *et al.* (2007). Conserved synteny of genes between chromosome 15 of *Bombyx mori* and a chromosome of *Manduca sexta* shown by five-color BAC-FISH. *Genome* 50(11), 1061-65. doi: 10.1139/g07-082
- Sichová, J., Nguyen, P., Dalíková, M. and Marec, F. (2013). Chromosomal evolution in tortricid moths: conserved karyotypes with diverged features. *PloS One* 8: e64520. doi: 10.1371/journal.pone.0064520
- Sichová, J., Ohno, M., Dinca, V., *et al.* (2016). Fissions, fusions and translocations shaped the karyotype and multiple sex chromosome constitution of the northeast-Asian wood-white butterfly, *Leptidea amurensis*. *Biol. J. Linnean Soc* 118, 457-471. doi: 10.1111/bij.12756
- Sichová, J., Voleníková, A., Dinca, V., *et al.* (2015). Dynamic karyotype evolution and unique sex determination systems in *Leptidea* wood white butterflies. *BMC Evol. Biol* 15:89. doi: 10.1186/s12862-015-0375-4
- Smit, A. F. A., Hubley, R. and Green, P. (2017) RepeatMasker Open-4.0. Available at <http://www.repeatmasker.org>
- Talavera, G., Lukhtanov, V. A., Rieppel, L., *et al.* (2013). In the shadow of phylogenetic uncertainty: the recent diversification of *Lysandra* butterflies through
-

-
- chromosomal change. *Mol. Phylogenetics and Evolution* 69(3), 469-478. doi: 10.1016/j.ympev.2013.08.004
- Talla, V., Suh, A., Kalsoom, F. *et al.* (2017). Rapid increase in genome size as a consequence of transposable element hyperactivity in wood-white (Leptidea) butterflies. *Genome Biol Evol.* 9(10), 2491-2505. doi: 10.1093/gbe/evx163
- Traut, W. and Marec, F. (1996). Sex chromatin in Lepidoptera. *Quar. Rev. Biol* 71, 239-256. doi: 10.1086/419371
- Traut, W. and Marec, F. (1997). Sex chromosome differentiation in some species of Lepidoptera (Insecta). *Chromosome Research* 5, 283-91. doi: 10.1023/B:CHRO.0000038758.08263.c3
- Traut, W., Sahara, K. and Marec, F. (2007). Sex chromosomes and sex determination in Lepidoptera. *Sex Dev* 1(6), 332-346. doi: 10.1159/000111765
- Traut, W., Sahara, K., Otto, T. D. and Marec, F. (1999). Molecular differentiation of sex chromosomes probed by comparative genomic hybridization. *Chromosoma* 108, 173-80. doi: 10.1007/s00412005036
- Traut, W., Vogel, H., Glöckner, G., Hartmann, E. and Heckel, D. G. (2013). High-throughput sequencing of a single chromosome: a moth W chromosome. *Chromosome Research* 110, 491-505. doi: 10.1007/s10577-013-9376-6
- Van Nieuwerkerken, E. J., Kaila, L., Kitching, I. J., *et al.* (2011). “Order Lepidoptera Linnaeus, 1758” in *Animal Biodiversity: An Outline of Higher-Level Classification and Survey Taxonomic Richness*, ed. Z. Q. Zhang (Auckland: Magnolia Press). doi: 10.11646/ZOOTAXA.3148.1.41
- Vechtová, P., Daliková, M., Sykorová, M., *et al.* (2016). CpSAT-1, a transcribed satellite sequence from the coding moth, *Cydia pomonella*. *Genetica* 144, 385-95. doi: 10.1007/s10709-016-9907-0
- Virkki, N. (1963). Gametogenesis in the sugarcane borer moth, *Diatraea saccharalis* (f.) (Crambidae). *J. Agricult. University Puerto Rico* 47, 102-37. doi: 10.46429/jaupr. v47i2.12944
-

- Vítková, M. Fuková, I., Kubícková, S. and Marec, F. (2007). Molecular divergence of the W chromosomes in pyralid moths (Lepidoptera). *Chromosome Res.* 15(7), 917-30. doi: 10.1007/s10577-007-1173-7
- Yasukochi, Y., Ashakumary, L. A., Baba, K., *et al.* (2006). A second-generation integrated map of the silkworm reveals synteny and conserved gene order between lepidopteran insects. *Genetics* 173(3), 1319-1328. doi: 10.1534/genetics.106.055541
- Yasukochi, Y., Ohno, M., Shibata, F., *et al.* (2016). A FISH-based chromosome map for the European corn borer yields insights into ancient chromosomal fusions in the silkworm. *Heredity* 116, 75-83. doi: 10.1038/hdy.2015.72
- Yoshido, A., Sahara, K., Marec, F. and Matsuba, Y. (2011). Step-by-step evolution of neo-sex chromosome in geographical populations of wild silkmoths, *Samia cynthia* spp. *Heredity* 106(4), 614-624. doi: 10.1038/hdy.2010.94
- Yoshido, A., Sichová, J., Pospíšilová, K., *et al.* (2020). Evolution of multiple sex-chromosomes associated with dynamic genome reshuffling in *Leptidea* wood-white butterflies. *Heredity* 125, 138-154. doi: 10.1038/s41437-020-0325-9
- Yoshido, A., Yamada, Y. and Sahara, K. (2006). The W chromosomes detection in several lepidopteran species by Genomic *in situ* hybridization (GISH). *J Insect Biotec and Sericology* 75, 147-151. doi: 10.11416/jibs.75.147
- Zhangwu, Z., *et al.* (1994). Chromosome studies of the Millet stem borer, *Chilo infuscatellus*, Snellen (Lepidoptera, Pyralidae). *Journal of Shanxi Agricultural University* 14, 394-96. doi: 10.13842/j.cnki.issn1671-8151.1994.04.017
- Zrzavá, M., Hladová, I., Daliková, M., *et al.* (2018). Sex chromosomes of the iconic moth *Abraxas grossulariata* (Lepidoptera, Geometridae) and its congener *A. sylvata*. *Genes* 9:279. doi: 10.3390/genes9060279
-

Figures

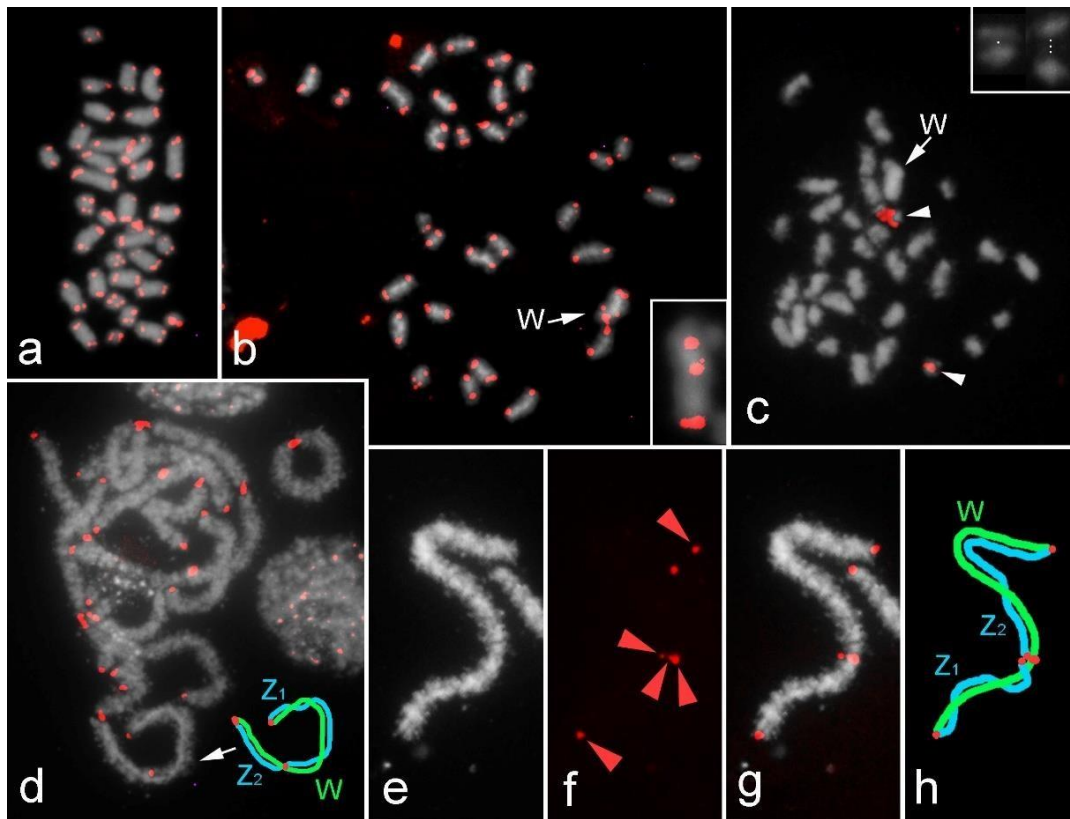


Figure 1. Analysis of general chromosomal characteristics of *Diatraea saccharalis* karyotype. Spreads of mitotic chromosomes from wing discs of male **(a)** and female **(b)** larvae labeled for telomeric probe with $2 = 34$ and $2n = 33$, respectively. In **(b)** the chromosome W is indicated and highlighted in the inset. Note the presence of an interstitial telomeric site on this chromosome. **(c)** mitotic metaphase spread from ovary probed for 18S rDNA. The inset show the small autosome that harbor the rDNA cluster containing an evident constriction (white dots). The W chromosome is indicated. **(d)** Female pachytene with identification of telomeres. In the sex chromosome it is possible to observe interstitial signal for telomeric probe, indicated in the schematic representation in the right side. Observe the W chromosome with deeply staining for DAPI. **(e-h)** Selected sex chromosomes showing the distribution of the telomeric signals; **(e)** DAPI, **(f)** telomeric signals indicated by red arrowheads, **(g)** merge, **(h)** schematic representation.

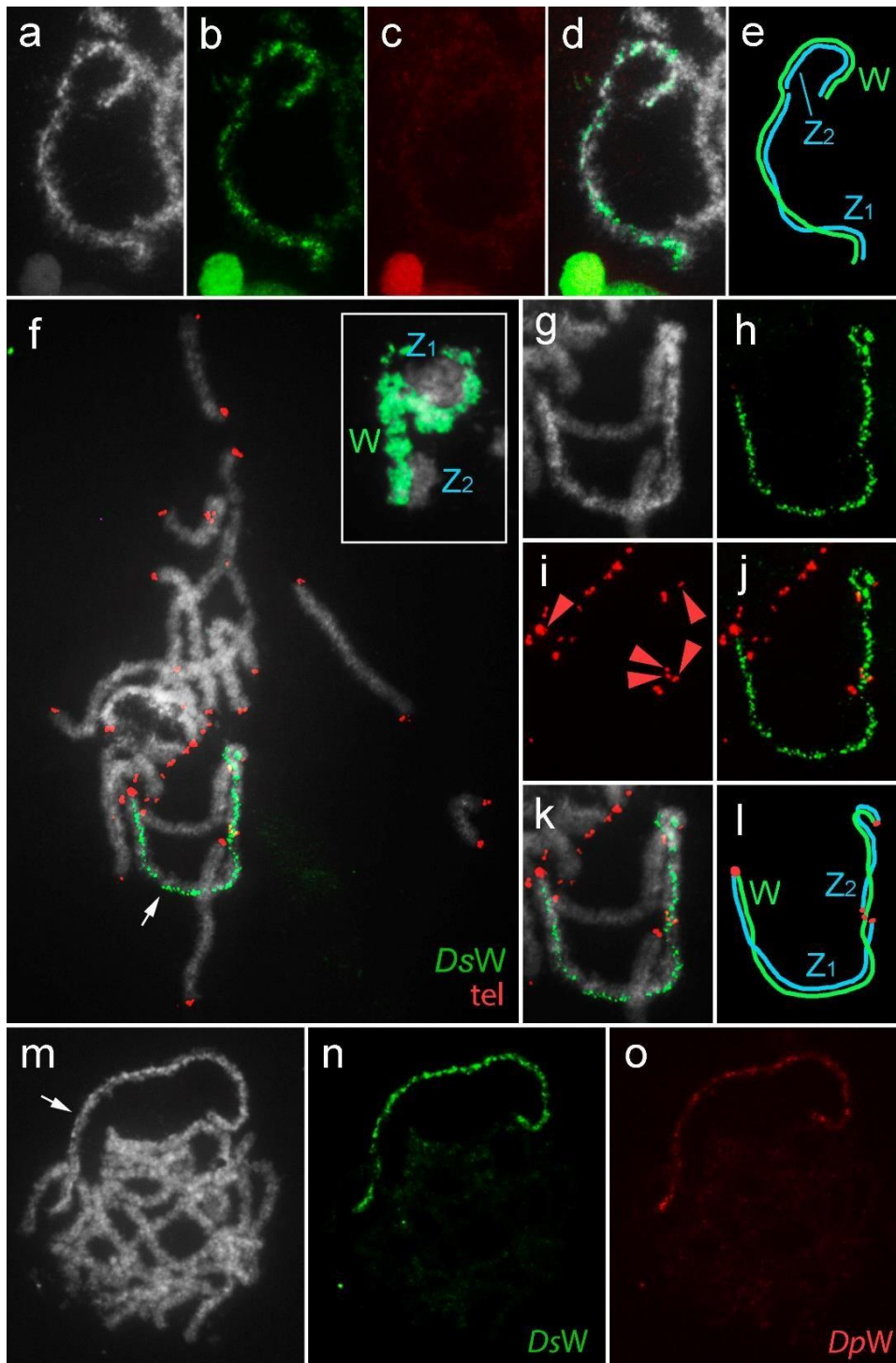


Figure 2. Sex chromosomes characterization of *Diatraea saccharalis*. **(a-e)** Comparative genomic hybridization in pachytene from oocytes, the sex chromosomes were selected. **(a)** DAPI staining, **(b)** female-derived genomic probe, **(c)** male-derived genomic probe, **(d)** merged probed and DAPI, **(e)** schematic representation of sex chromosomes. **(f-l)** two color FISH for W-derived probe (green) and telomeric probe (red) in pachytene oocyte. **(f)** the complete pachytene, **(g)** selected sex chromosomes stained with DAPI, **(h)** W-

derived probe hybridization, **(i)** telomeric probe hybridization (the signals corresponding to sex chromosomes are indicated by red arrowheads), **(j)** merged image for W-painting probe and telomere signals, **(k)** merged image for probed and DAPI counterstaining, **(l)** schematic representation of the sex trivalent. In **(f)** the inset shows a selected sex trivalent WZ_1Z_2 from a nurse cell of a female ovary. **(m-o)** Zoo-FISH of W-chromosome from *D. postlineella* in female pachytene of *D. saccharalis*. **(m)** Chromosomes counterstained in DAPI, **(n)** W-derived probe from *D. saccharalis* and **(o)** W-chromosome probe from *D. postlineella*. Note signals of W-painting probe from *D. postlineella* spread along the entire extension of the W chromosome of *D. saccharalis*, but with faint signal.

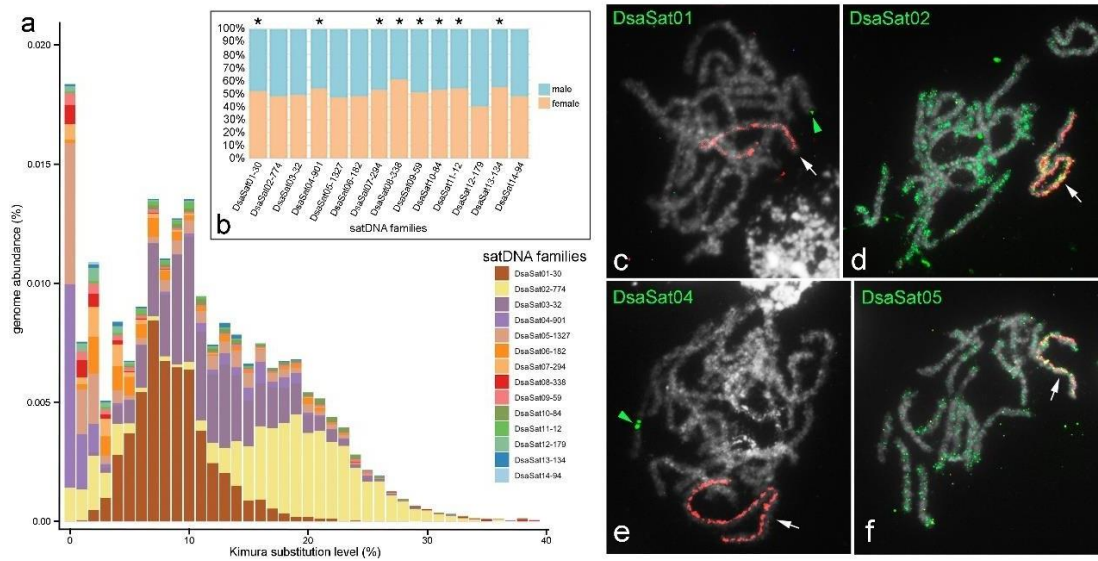


Figure 3. SatDNAs on the genome of *Diatraea saccharalis*, genomic (a,b) and chromosomal (c-f) characterizations. (a) satDNAs landscape (mean of abundance versus K2P divergence of male and female genomes). Note two peaks of abundance, one with K2P divergence of 0 and the second between 6-10. (b) Comparative abundance of each satDNA family in male and female genomes. Observe the slightly higher abundance of eight of the satDNA families in female genome. (c-f) FISH mapping of four satDNAs (green) and W-painting probe (red) on female pachytene. The satDNAs mapped are indicated directly in each panel. The white arrows point the W sex chromosome and the green arrowheads in (c,e) the single cluster for two of satDNAs. Observe in (d,f) the two satDNAs virtually distributed homogeneously in all chromosomes.

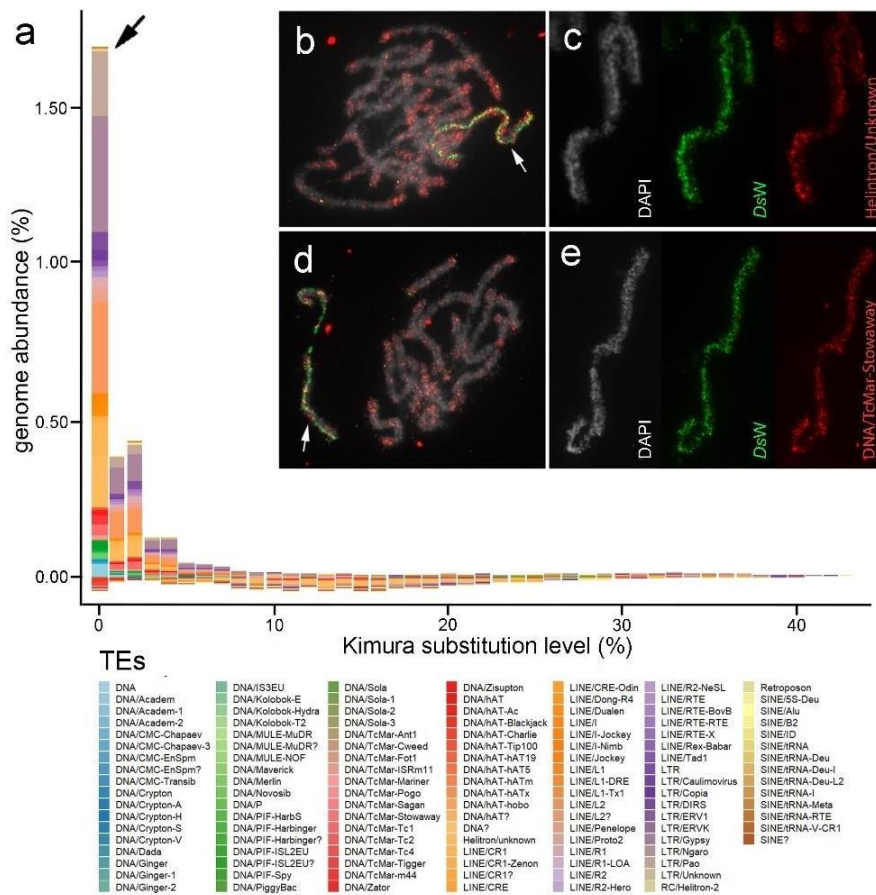


Figure 4. TEs on the genome of *Diatraea saccharalis*. **(a)** subtractive landscape (abundance versus divergence) obtained from female and male genomes. Abundance values are the difference between female minus male genomes. Positive values indicate overabundant sequences in female genome and negative values overabundant sequences on male genome. Observe overabundance of some TEs with low divergence in female genome, suggesting recent amplification. **(b,d)** FISH mapping of two TE elements (red) and W-painting probe (green) on pachytene oocytes. The white arrows show the sex chromosomes. In **(c,e)** sex chromosomes from female pachytenes were selected to show in detail the distribution of the two TEs.

Tables

Table 1. General characteristics, including, satDNA name, monomer size, A+T content, abundance, abundance ratio between female/male genomes and divergence, for the 14 satDNA families identified in the genome of *Diatraea saccharalis*.

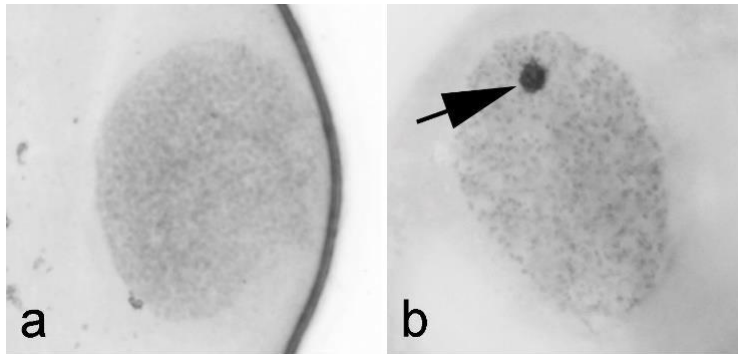
satDNA ID	Size (bp)	A+T (%)	Abundance (%)		Abundance ratio female/male (%)	Divergence (%)	
			Female	Male		Female	Male
DsaSat01	30	53.4	0,0570	0,0524	1,0878	9,44	9,03
DsaSat02	774	67.3	0,0504	0,0553	0,9114	17,85	17,6
DsaSat03	32	53.1	0,0425	0,0436	0,9748	12,41	12,39
DsaSat04	901	59.4	0,0184	0,0155	1,1871	3,96	6,05
DsaSat05	1327	66.7	0,0174	0,0196	0,8878	6,98	6,07
DsaSat06	182	70.3	0,0082	0,0087	0,9425	7,10	7,75
DsaSat07	294	70.5	0,0050	0,0045	1,1111	4,22	3,56
DsaSat08	338	61.8	0,0036	0,0023	1,5652	8,81	4,34
DsaSat09	59	47.5	0,0036	0,0035	1,0286	8,73	8,59
DsaSat10	84	72.5	0,0035	0,0030	1,1667	17,92	18,04
DsaSat11	12	50.0	0,0020	0,0017	1,1765	9,45	8,58
DsaSat12	179	62.0	0,0016	0,0025	0,6400	4,82	4,04
DsaSat13	134	67.9	0,0016	0,0013	1,2308	10,26	8,96
DsaSat14	94	39.4	0,0002	0,0002	1,0000	4,95	4,57

Table 2. Main TEs classes recovered from the genome of *Diatraea saccharalis* and their characteristics, including number of elements, genome abundances and abundance ratio between female/male genomes.

TE Class	Number of Elements		Genome abundance (%)		Abundance ratio (female/male)
	Female	Male	Female	Male	
LINE	3179	3156	11,41721	10,37860	1,10007
Helitron	2294	2293	10,79091	10,75877	1,00299
DNA	1921	1916	7,07904	6,87058	1,03034
LTR	2139	2098	6,67902	5,58523	1,19584
SINE	380	381	3,39406	3,50060	0,96957
Unknown	391	388	1,98161	2,08621	0,94986
Retroposon	1	1	0,00448	0,00492	0,91057
RC	1	1	0,00014	0,00001	14,00000

Supplementary Material

Supplementary Figure 1. Polyploid nuclei obtained from Malpighian tubule cells from **(a)** male and **(b)** female larvae. The material was stained with Lacto-acetic orcein. The arrow in **(b)** points the deeply stained W chromatin body.



Supplementary Table 1. DNA sequences oligos designed to obtain probes for the repetitive DNAs prospected for the first time in this work. Asterisks show oligos directly labelled with biotin-14-dATP to be used as probes.

Sequence names	Oligos
satDNAs	
DsaSat01-30*	TCTGATGTCAACGTCAAATGCCGTGACACTTCTGATGTCAACGTCAAATGCCGTGACACT
DsaSat02-774	F: AAATCGGTCTAAGGTTCCGCG R: CTACCCGATCCTGAACCTCC
DsaSat03-32*	ATCTGAGTGTGTCATGACACTTCATGTGACGTCGATCTGAGTGTGTCATGACACTTCATGTGACGTCG
DsaSat04-901	F: TCATTGTATACAGATAGTGATTT R: CTCGTTAATAACCGTATATC
DsaSat05-1327	F: GTTGCAACAGATCCACACC R: GCCGTCTTACTTACTGAGCG
DsaSat07-294*	TTACATAACGAACATAACTTGACGACTTTTAAAGCAGAGATTGATATTTTTCTTCAAACAGAAATGCTCAGTCAGTTCTTTTTCAAACGTTGTCTGCAC
DsaSat08-338*	TTTCAACAACAATAAAAACAAAGTATGTTTTACCTGAAGAATGGCGAGACCTATTTTGCCACGGATCAACATAGGATCTATTCCTCCAGTTAAAAAAT
DsaSat10-84*	AGAAATCACAGACATATGGTCTTTCATTAGTATGTATCCTTTATGTACGACTAAATTACTTCTTTGAGTAAATGATTTATTAC
DsaSat11-12*	CAGTTGAGTCGACAGTTGAGTCGACAGTTGAGTCGA
DsaSat13-134*	TGCTTAATTACATTATTGAAGAAGGCATGAGACTCTTTGAGAAAGGGTTTTCTGTATTATAAACATTTGTAAATGGACAGTACTCTGCAGAAGCAATATC
TEs	
Helintron/Unknown	F: GTGCCACAATTGCTGCCAA R: GCATGCGGCTACTTCCAGA
DNA/TcMAR-Stowaway	F: TCAAACGCAAGAAATAGCAATTTTAA R: AAGGATGCTGACATCAGTATTAAG

Supplementary Table 2. General characteristics of TEs prospected in the genome of *Diatraea saccharalis*. It includes number of elements, genome abundance, abundance ration between female/male genomes and genome divergence.

TE Class/Superfamily		Number of Elements		Genome abundance (%)		Abundance ratio (female/male)	Genome divergence (%)	
		Female	Male	Female	Male		Female	Male
DNA	hAT-Blackjack	115	115	0,99100	1,03027	0,96189	8,14	8,11
	DNA	179	180	0,91327	0,88745	1,02909	9,52	10,06
	hAT-Charlie	223	223	0,81968	0,85343	0,96045	11,08	10,96
	TcMar-Mariner	149	148	0,62801	0,62302	1,00800	7,16	7,40
	TcMar-Tc1	106	105	0,54815	0,46269	1,18469	6,46	7,42
	PIF-Harbinger	125	125	0,42668	0,43617	0,97822	9,76	9,97
	hAT-Tip100	150	150	0,37346	0,36075	1,03522	11,47	12,08
	hAT-Ac	101	100	0,29146	0,27603	1,05593	7,66	8,33
	P	50	50	0,19702	0,18807	1,04761	7,71	7,88
	Zator	65	65	0,17696	0,18267	0,96877	13,20	13,41
	hAT	51	51	0,15992	0,14471	1,10510	9,06	9,85
	CMC-EnSpm	47	47	0,14996	0,14349	1,04510	7,85	7,89
	DNA?	15	15	0,11228	0,10939	1,02644	6,29	6,64
	TcMar-m44	26	26	0,11291	0,07673	1,47147	4,72	6,87

CMC-Chapaev-3	35	35	0,09959	0,09316	1,06900	4,42	4,76
CMC-Transib	18	18	0,09122	0,08269	1,10315	6,80	7,33
PIF-Spy	13	13	0,09505	0,07089	1,34085	3,91	4,78
PiggyBac	59	60	0,08872	0,07775	1,14107	4,82	5,77
MULE-MuDR	33	34	0,08249	0,07799	1,05771	9,34	9,55
hAT-hATm	25	24	0,07582	0,07708	0,98365	3,63	3,69
Ginger-2	22	22	0,06376	0,06351	1,00395	11,90	12,35
Sola-1	42	40	0,06245	0,05967	1,04670	10,52	11,70
TcMar-Fot1	43	43	0,06151	0,05729	1,07365	6,67	6,91
Sola-2	21	21	0,04356	0,04715	0,92389	7,00	6,74
Maverick	20	18	0,04856	0,01888	2,57230	6,03	10,57
Ginger	13	13	0,03142	0,03235	0,97122	6,67	7,29
Academ-1	20	20	0,02740	0,02987	0,91735	9,53	10,47
MULE-MuDR?	1	1	0,02693	0,02929	0,91947	4,58	4,48
Academ	15	15	0,02584	0,02865	0,90208	11,45	11,57
TcMar-Tigger	21	21	0,02611	0,02655	0,98337	13,15	13,37
CMC-Chapaev	3	3	0,02145	0,02244	0,95611	17,08	17,55
hAT-hATx	16	16	0,02200	0,02118	1,03839	8,83	9,91
Kolobok-Hydra	2	2	0,01612	0,01736	0,92877	2,15	2,01

Crypton-V	5	5	0,01513	0,01319	1,14695	5,94	6,74
hAT-hAT19	7	7	0,01341	0,01267	1,05905	7,76	8,93
Zisupton	5	5	0,01302	0,01314	0,99076	8,86	9,74
Novosib	3	3	0,01028	0,00960	1,07102	4,84	4,48
hAT-hAT5	4	4	0,00915	0,01035	0,88405	4,22	3,71
Kolobok-E	5	5	0,00934	0,00952	0,98038	4,24	4,26
PIF-ISL2EU?	1	1	0,01098	0,00581	1,89000	2,90	2,64
TcMar-Pogo	7	7	0,00851	0,00834	1,02054	10,44	9,30
Ginger-1	6	6	0,00817	0,00796	1,02678	6,66	7,12
hAT?	7	7	0,00657	0,00667	0,98506	11,30	12,57
Sola	4	4	0,00642	0,00622	1,03273	6,69	7,26
TcMar-Tc4	3	3	0,00608	0,00535	1,13659	12,32	13,02
Dada	1	1	0,00553	0,00559	0,99008	16,31	16,86
Kolobok-T2	3	3	0,00545	0,00527	1,03504	7,56	7,68
Crypton-S	1	1	0,00492	0,00542	0,90735	4,71	3,62
PIF-Harbinger?	5	5	0,00485	0,00529	0,91657	13,82	13,29
Merlin	5	5	0,00501	0,00468	1,07206	10,08	11,73
CMC-EnSpm?	2	2	0,00449	0,00534	0,84023	18,70	18,67
hAT-hobo	2	2	0,00421	0,00412	1,02220	11,10	12,02

	MULE-NOF	1	1	0,00329	0,00320	1,02874	6,25	4,93
	Crypton-H	4	4	0,00312	0,00257	1,21578	14,73	16,68
	TcMar-ISRm11	3	3	0,00296	0,00187	1,58209	5,79	10,47
	TcMar-Ant1	1	1	0,00207	0,00195	1,06188	19,54	19,95
	Crypton-A	1	1	0,00075	0,00403	0,18504	19,97	17,92
	Crypton	2	2	0,00178	0,00170	1,05201	16,58	17,62
	PIF-ISL2EU	1	1	0,00084	0,00068	1,24948	14,80	9,73
	TcMar-Stowaway	1	1	0,00051	0,00011	4,68729	0,97	0,10
	TcMar-Cweed	1	1	0,00038	0,00031	1,24470	10,08	17,24
	PIF-HarbS	1	1	0,00032	0,00022	1,49917	4,04	4,91
	TcMar-Sagan	1	1	0,00025	0,00011	2,23718	3,60	7,09
	Academ-2	1	1	0,00022	0,00012	1,88486	4,17	6,22
	IS3EU	1	1	0,00020	0,00010	1,91933	0,36	0,00
	Sola-3	1	1	0,00005	0,00023	0,21353	13,09	0,47
	TcMar-Tc2	1	1	0,00002	0,00028	0,08766	0	1,31
Total		1921	1916	7,07904	6,87058	1,03034		
LINE	L2	975	964	3,90321	3,37825	1,15539	6,86	7,52
	R1	261	256	1,16609	1,09104	1,06879	7,73	7,85
	CR1	418	415	1,04753	0,90206	1,16127	6,76	7,48

RTE-RTE	313	312	0,95611	0,94384	1,01300	8,24	8,57
Penelope	66	66	0,86303	0,79938	1,07961	6,34	6,85
RTE	230	230	0,82707	0,80980	1,02133	9,14	9,62
RTE-BovB	250	250	0,58696	0,58228	1,00803	11,85	12,37
I	166	166	0,59965	0,54297	1,10438	4,42	4,75
R1-LOA	67	67	0,48594	0,45382	1,07078	5,59	5,45
I-Jockey	158	156	0,33545	0,27500	1,21984	8,01	8,02
CR1-Zenon	57	57	0,21808	0,19606	1,11231	7,45	8,31
L1	22	24	0,13953	0,14645	0,95276	8,71	8,80
Proto2	75	74	0,10620	0,07751	1,37011	6,65	8,47
Dong-R4	42	42	0,05698	0,05738	0,99313	6,73	7,47
R2	19	19	0,03490	0,03818	0,91407	10,51	10,50
CRE	7	7	0,02128	0,02275	0,93546	10,65	11,39
RTE-X	9	9	0,01735	0,01604	1,08183	12,12	13,51
CR1?	6	6	0,01033	0,01079	0,95691	9,28	9,33
Dualen	1	1	0,01201	0,00660	1,82000	5,05	7,28
Rex-Babar	5	5	0,00837	0,00838	0,99835	7,00	6,94
Tad1	6	5	0,00465	0,00532	0,87399	13,02	12,97
L2?	3	3	0,00397	0,00389	1,01944	8,60	8,84

	CRE-Odin	3	3	0,00294	0,00266	1,10734	9,44	11,26
	I-Nimb	4	4	0,00282	0,00259	1,08842	3,12	3,66
	Jockey	5	5	0,00290	0,00241	1,20524	7,79	6,85
	L1-Tx1	7	6	0,00258	0,00191	1,35166	8,61	12,77
	R2-Hero	1	1	0,00066	0,00059	1,12228	2,66	3,05
	R2-NeSL	2	2	0,00032	0,00034	0,95546	6,35	11,83
	L1-DRE	1	1	0,00029	0,00032	0,90834	13,66	15,66
Total		3179	3156	11,41721	10,37860	1,10007		
SINE	tRNA-RTE	181	181	2,00833	2,08188	0,96467	13,35	13,39
	5S-Deu	144	145	0,98552	0,99855	0,98695	14,26	14,55
	tRNA-Deu	17	17	0,11201	0,12005	0,93300	10,76	10,79
	tRNA	15	15	0,08611	0,08401	1,02504	7,08	7,12
	tRNA-Meta	4	4	0,07458	0,07779	0,95862	9,89	9,59
	ID	3	3	0,03983	0,04216	0,94481	14,03	14,24
	tRNA-V-CR1	2	2	0,04027	0,03973	1,01355	15,85	16,29
	tRNA-Deu-L2	8	8	0,03746	0,03750	0,99897	12,64	12,99
	Alu	1	1	0,00473	0,01406	0,33634	14,68	15,07
	SINE?	1	1	0,00191	0,00218	0,87436	22,78	23,69
	tRNA-Deu-I	2	2	0,00190	0,00188	1,01272	17,59	19,46

Total	tRNA-I	1	1	0,00098	0,00056	1,74996	5,90	5,85
	B2	1	1	0,00043	0,00025	1,74817	11,41	11,55
		380	381	3,39406	3,50060	0,96956		
LTR	Gypsy	791	772	2,99455	2,33603	1,28190	3,78	4,14
	Pao	745	729	1,98513	1,67776	1,18320	3,69	3,80
	Copia	329	326	0,65740	0,62952	1,04430	6,18	6,74
	LTR	55	54	0,44095	0,40795	1,08090	8,77	9,32
	DIRS	169	168	0,33289	0,22765	1,46226	4,57	5,02
	ERV1	21	20	0,19476	0,20492	0,95044	7,13	7,03
	Caulimovirus	19	19	0,03457	0,06118	0,56508	10,84	14,34
	ERVK	5	5	0,02225	0,02231	0,99765	10,14	10,29
	Ngaro	2	2	0,01150	0,01282	0,89696	18,73	18,24
	Unknown	3	3	0,00500	0,00509	0,98256	11,39	12,20
Total		2139	2098	6,67902	5,58523	1,19584		
Helitron	Unknown	2294	2293	10,79091	10,75877	1,00299	9,41	9,78
RC	Helitron-2	1	1	0,00014	0,00001	13,34489	6,78	20,84
Retroposon		1	1	0,00448	0,00492	0,91238	7,06	7,28
Unknown		391	388	1,98161	2,08621	0,94986	8,84	8,77

5. CONCLUSÕES FINAIS

Nossos dados revelam o dinamismo do cariótipo geral e DNAs repetitivos no genoma de *Diatraea saccharalis*. Esse dinamismo é resultante de eventos de fusão cromossômica e disseminação ou amplificação local dos distintos DNAs repetitivos. A alta incidência de fusão no cariótipo de *D. saccharalis* resultou também no surgimento de um sistema neo-sexual múltiplo WZ_1Z_2 (fêmea)/ $Z_1Z_1Z_2Z_2$ (macho), no qual o cromossomo W degenerado e diferenciado do resto do genoma, essencialmente pelo acúmulo de TEs, possibilitou novos *insights* sobre a evolução dos cromossomos sexuais em Lepidoptera. Finalmente, como os DNAs repetitivos são uma parte difícil de montagem no genoma, nossos dados contribuem em futuras melhorias no sequenciamento e montagem do genoma dessa importante espécie praga.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ABBOTT, J.; NORDÉN, A. & HANSSON, BENGT. Sex chromosome evolution: Historical insights and future perspectives. **Proceedings of the Royal Society B: Biological Sciences**. 284. 20162806. 10.1098/rspb.2016.2806, 2017

BACHTROG D. Sex chromosome evolution: molecular aspects of Y-chromosome degeneration in *Drosophila*. **Genome research**, 15(10), 1393–1401. <https://doi.org/10.1101/gr.3543605>, 2005

BACHTROG, D.; HOM, E.; WONG, K.M. *et al.* Genomic degradation of a young Y chromosome in *Drosophila miranda*. **Genome Biol** 9, R30. <https://doi.org/10.1186/gb-2008-9-2-r30>, 2008

BACHTROG, D.; MANK, J. E.; PEICHEL, C. L.; KIRKPATRICK, M.; OTTO, S. P. *et al.* Sex Determination: Why So Many Ways of Doing It? **PLoS Biol.** 12(07):e1001899, DOI: doi:10.1371/journal.pbio.1001899, 2013

BARDELLA, V. B., MILANI, D., & CABRAL-DE-MELLO, D. C. Analysis of *Holhymenia histrio* genome provides insight into the satDNA evolution in an insect with holocentric chromosomes. **Chromosome Res.** 28, 369–380. doi:10.1007/s10577-020-09642-1, 2020

BARDELLA, V.B. & CABRAL-DE-MELLO, D.C. Uncovering the molecular organization of unusual highly scattered 5S rDNA: the case of *Chariesterus armatus* (Heteroptera). **Gene**, v. 646, p. 153-158, 2018

BIÉMONT, C. & VIEIRA, C. Genetics: Junk DNA as an evolutionary force. **Nature**, 443:521-524, 2006

BISCOTTI, M.A.; OLMO, E. & HESLOP-HARRISON, J.S. Repetitive DNA in eukaryotic genomes. **Chromosome Research**, v. 23(3), p. 415-420, 2015

BOEKE, J.D.; GARFINKEL, D.J.; STYLES, C.A. & FINK, G.R. Ty elements transpose through an RNA intermediate. **Cell**, v. 40, p. 491–500, 1985

BOURQUE, G.; BURNS, K.H.; GEHRING, M.; GORBUNOVA, V. SELUANOV A, *et al.* Ten things you should know about transposable elements. **Genome Biol.** 19:199. doi: 10.1186/s13059-018-1577-z, 2018

-
- BUENO, D., PALACIOS-GIMENEZ, O.M., MARTÍ, D.A. *et al.* The 5S rDNA in two Abracris grasshoppers (Ommatolampidinae: Acrididae): molecular and chromosomal organization. **Mol Genet Genomics** 291, 1607–1613, <https://doi.org/10.1007/s00438-016-1204-1>, 2016
- CABRAL-DE-MELLO, D.C. & MAREC, F. Universal fluorescence in situ hybridization (FISH) protocol for mapping repetitive DNAs in insects and other arthropods. **Mol Genet Genomics** 296, 513–526. <https://doi.org/10.1007/s00438-021-01765-2>, 2021
- CABRAL-DE-MELLO, D.C.; OLIVEIRA, S. G.; de MOURA, R. & MARTINS, C. Chromosomal organization of the 18S and 5S rRNAs and histone H3 genes in *Scarabaeinae coleopterans*: insights into the evolutionary dynamics of multigene families and heterochromatin. **BMC Genet** 12(88). doi: 10.1186/1471-2156-12-88, 2011
- CABRAL-DE-MELLO, D.C.; ZRZAVÁ, M.; KIBICKOVÁ, S. *et al.* The Role of Satellite DNAs in Genome Architecture and Sex Chromosome Evolution in Crambidae Moths. **Frontiers in Genetics**. 12:661417. doi: 10.3389/fgene.2021.661417, 2021
- CARBAJAL PALADINO, L. Not so stable after all: Neo-sex chromosomes and speciation in Lepidoptera. In: Abstracts, 21st International Chromosome Conference, 10-13 July 2016, Foz do Iguacu, Brazil. **Cytogenet Genome Res**, 148: 83–155 (poster no. II.8), 2016
- CHARLESWORTH, B.; SNIEGOWSKI, P. & STEPHAN, W. The evolutionary dynamics of repetitive DNA in eukaryotes. **Nature**, v. 371, p. 215-220, 1994
- CHARLESWORTH, D. Evolution of recombination rates between sex chromosomes. **Phil. Trans. R. Soc. B** 372:20160456, 2017
- CHARLESWORTH, D.; CHARLESWORTH, B. & MARAIS, G. Steps in the evolution of heteromorphic sex chromosomes. **Heredity** 95, 118-128. <https://doi.org/10.1038/sj.hdy.6800697>, 2005
- COOPER, R. Experimental demonstration of holokinetic chromosomes, and of differential "radiosensitivity" during oogenesis in the grass mite, *Siteroptes graminum* (Reuter). *J Exp. Zool*, 182(1), 69-94, 1972
- CRAIG, N. L.; CHANDLER, M.; GELLERT, M. *et al.* Mobile DNA II, 3rd ed. Washington, DC. **American Society for Microbiology**, 2015
-

-
- DALIKOVÁ, M.; ZRZAVÁ, M.; HLADOVÁ, I.; *et al.* New insights into the evolution of the W chromosome in Lepidoptera. **Journal of Heredity**, v. 108(7), p. 709-719, 2017a
- DALIKOVÁ, M.; ZRZAVÁ, M.; KUBICKOVÁ, S. & MAREC, F. W-enriched satellite sequence in the Indian meal moth, *Plodia interpunctella* (Lepidoptera, Pyralidae). **Chromosome Res** 25, 241-252. <https://doi.org/10.1007/s10577-017-9558-8>, 2017b
- DERNBURG, A.F. Here, there, and everywhere: kinetochore function on holocentric chromosomes. **Journal of Cell Biology**, v. 153(6), p. F33-38, 2001
- DOPMAN, E. B.; PÉREZ, L.; BOGDANOWICZ, S. M. & HARRISON, R. G. Consequences of reproductive barriers for genealogical discordance in the European Corn Borer. **Proc National Academy of Sciences** 102(41), 14706-11. doi:10.1073/pnas.0502054102, 2005
- DRINNENBERG, I.A.; deYOUNG, D.; HENIKOFF, S. & MALIK, H.S. Recurrent loss of CenH3 is associated with independent transitions to holocentricity in insects. **Elife**. 23;3: e03676. doi: 10.7554/eLife.03676. PMID: 25247700; PMCID: PMC4359364, 2014
- EIRÍN-LOPEZ, J. M.; REBORDINOS, L.; ROONEY, A. P. & ROZAS, J. The birth-and-death evolution of multigene families revisited. **Genome dynamics** 7, 170-196. <https://doi.org/10.1159/000337119>, 2012
- ESCUDERO, M.; MÁRQUEZ-CORRO, J. & HIPPEL, A. The Phylogenetic Origins and Evolutionary History of Holocentric Chromosomes. **Systematic Botany**. 41. 580-585. doi: 10.1600/036364416X692442., 2016
- FARIA, R & NAVARRO, A. Chromosomal speciation revisited: Rearranging theory with pieces of evidence. **Trends Ecol Evol** 25(11):660–669, 2010
- FELICIELLO, I.; AKRAP, I.; BRAJKOVIC, J.; ZLATAR, I. & UGARKOVIC, D. Satellite DNA as a driver of population divergence in the red flour beetle *Tribolium castaneum*. **Genome Biol. Evol.** 7, 228-239, 2014
- FERRETTI, A. B. S. M.; MILANI, D.; PALACIOS-GIMENEZ, O. M. *et al.* High dynamism for neo-sex chromosomes: satellite DNA reveal complex evolution in a grasshopper. **Heredity** 125, 124-137. doi: 10.1038/s41437-020-0327-7, 2020
- FESCHOTTE, C. & PRITHAM, E. J. DNA transposons and the evolution of eukaryotic genomes. **Ann Rev Genet**, 41:331-368, 2007
-

-
- FREIRE, R., ARIAS, A., ÍNSUA, A.M. *et al.* Evolutionary Dynamics of the 5S rDNA Gene Family in the Mussel *Mytilus*: Mixed Effects of Birth-and-Death and Concerted Evolution. **J Mol Evol** 70, 413–426, <https://doi.org/10.1007/s00239-010-9341-3>, 2010
- FUKOVÁ, I.; TRAUT, W.; VITKOVÁ, M. *et al.* Probing the W chromosome of the codling moth, *Cydia pomonella*, with sequences from microdissected sex chromatin. **Chromosoma** 116, 135-145. doi: 10.1007/s00412-006-0086-0, 2007
- FURMAN, B.; METZGER, D.; DAROLTI, I.; WRIGHT, A.E. *et al.* Sex Chromosome Evolution: So Many Exceptions to the Rules. **Genome Biology and Evolution**, Volume 12, Issue 6, Pages 750–763, <https://doi.org/10.1093/gbe/evaa081>, 2020
- GARRIDO-RAMOS, M. A. Satellite DNA in Plants: More than Just Rubbish. **Cytogenet. Genome Res.** 146, 153-170, 2015
- GARRIDO-RAMOS, M. A. Satellite DNA: An Evolving Topic. **Genes**, 8(9), 230. <https://doi.org/10.3390/genes8090230>, 2017
- GOUBERT, C.; MODOLO, L.; VIEIRA, C. *et al.* De novo assembly and annotation of the Asian tiger mosquito (*Aedes albopictus*) repeatome with dnaPipeTE from raw genomic reads and comparative analysis with the yellow fever mosquito (*Aedes aegypti*). **Genome Biology and Evolution** 7(4), 1192-1205. doi: 10.1093/gbe/evv050, 2015
- GRABHERR, M. G.; HAAS, B. J.; YASSOUR, M. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature Biotechnology** 29(7), 644-52. doi: 10.1038/nbt.1883, 2011
- HAIZEL, T.; LIM, Y. K.; LEITCH, A. R. & MOORE, G. Molecular análisis of holocentric centromeres of *Luzula* species. **Cytogenet Genome Res** 109(1-3), 134-143. <https://doi.org/10.1159/000082392>, 2005
- HECKMANN, S.; MACAS, J. KUMKE, K.; FUCHS, J. & SCHUBERT, V. *et al.* The holocentric species *Luzula elegans* shows interplay between centromere and large-scale genome organization. **The Plant journal: for cell and molecular biology**. 73. 10.1111/tpj.12054, 2012
- HILL, J.; RASTAS, P.; HORNETT, E.A.; NEETHIRAJ, R., CLARK, N.; MOREHOUSE, N. *et al.* Unprecedented reorganization of holocentric chromosomes
-

-
- provides insights into the enigma of lepidopteran chromosome evolution. **Sci Adv.** 12;5(6):eaau3648. doi: 10.1126/sciadv.aau3648, 2019
- HOBZA, R.; KUBAT, Z.; CEGAN, R. *et al.* Impact of repetitive DNA on sex chromosome in plants. **Chromosome Res.** 23: 561, 2015
- HOUBEN, A. & SCHUBERT, I. DNA and proteins of plant centromeres. **Current Opinion in Plant Biology**, v. 6(6), p. 554-560, 2003
- HUA-VAN, A.; LE ROUZIE, A.; MAISONHAUTE, C. & CAPY, P. Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. **Cytogenet Genome Res.** 110:426-440. doi: 10.1159/000084975, 2005
- IJDO, J. M.; WELLS, R. A.; BALDINI, A. & REEDERS, S. T. Improved telomere detection using a telomere repeat probe (TTAGGG)_n generated by PCR. **Nucleic Acids Research** 19(17), 4780. doi: 10.1093/nar/19.17.4780, 1991
- KAGEYMA, D. & TRAUT, W. Opposite sex-specific effects of *Wolbachia* and interference with the sex determination of its host *Ostrinia scapularis*. **Proc R Soc Lond B**, 271: 251–258, 2004
- KEJNOVSKY, E.; HOBZA, R.; KUBAT, Z.; CERMAK, T. & VYSKOT, B. The role of repetitive DNA in structure and evolution of sex chromosomes in plants. **Heredity.** 102:533-541, 2009
- KEJNOVSKY, E.; MICHALOVOVA, M. & STEFLOVA, P. *et al.* Expansion of Microsatellites on Evolutionary Young Y Chromosome. **PloS One** 8(1): e45519. <https://doi.org/10.1371/journal.pone.0045519>, 2013
- KIDWELL, M. Chapter 3 – Transposable Elements. In book: The Evolution of the Genome (pp. 165-221) **Elsevier Inc**, 2005
- LIU, J.; HE, Y.; AMASINO, R. & CHEN, X. siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. **Genes Dev.** 18(23), 2873-8. doi:10.1101/gad.1217304, 2004
- LIU, X.; YANG, W.; MU, B.; *et al.* Engineering of 'Purple Embryo Maize' with a multigene expression system derived from a bidirectional promoter and self-cleaving 2A peptides. **Plant Biotechnol J.** 16(6):1107-1109. doi:10.1111/pbi.12883, 2018
-

-
- LÓPEZ-FLORES, I. & GARRIDO-RAMOS, M. A. The repetitive DNA content of eukaryotic genomes. **Genome Dyn.** v. 7, p.1-28, 2012
- LUKHTANOV, V. A. Sex chromatin and sex chromosome systems in nonditrysian Lepidoptera (Insecta). **J Zool Syst Evol Res.** 38(2):73–79, 2000
- LUKHTANOV, V. A.; DINČÁ, V.; FRIBERG, M. *et al.* Versatility of multivalent orientation, inverted meiosis, and rescued fitness in holocentric chromosomal hybrids. **Proc Natl Acad Sci USA** 115(41):e9610-9619. doi:10.1073/pnas.1802610115, 2018
- LYNCH, M. The Origins of Genome Architecture. 1st ed. Sunderland: Sinauer Associates, 494p., 2007
- MACAS, J.; MÉSZÁROS, T. & NOUZOVÁ, M. PlantSat: a specialized database for plant satellite repeats. **Bioinformatics** 18, 28-35. <https://doi.org/10.1093/bioinformatics/18.1.28>, 2002
- MALIK, H. S. & HENIKOFF, S. Major evolutionary transitions in centromere complexity. **Cell** 138(6), 1067-82. <https://doi.org/10.1016/j.cell.2009.08.036>, 2009
- MANDRIOLI, M. & MANICARDI, G.C. Holocentric chromosomes. **PLOS Genetics** 16(7): e1008918. <https://doi.org/10.1371/journal.pgen.1008918>, 2020
- MANK, J.; HOSKEN, D. & WEDELL, N. Some inconvenient truths about sex chromosome dosage compensation and the potential role of sexual conflict. **Evolution; international journal of organic evolution.** doi: 65. 2133-44. 10.1111/j.1558-5646.2011.01316. x, 2011
- MAREC, F.; SAHARA, K. & TRAUT, W. Rise and fall of the W chromosome in Lepidoptera. In: Goldsmith MR, Marec F, eds. **Molecular Biology and Genetics of the Lepidoptera.** Boca Raton, FL: CRC Press, 49-63, 2010
- MAREC, F.; TOTHOVÁ, A.; SAHARA, K. *et al.* Meiotic pairing of sex chromosome fragments and its relation to atypical transmission of a sex-linked marker in *Ephesia kuehniella* (Insecta: Lepidoptera). **Heredity** 87, 659-671. <https://doi.org/10.1046/j.1365-2540.2001.00958.x>, 2001
- MARQUES, A.; RIBEIRO, T.; NEUMANN, P.; MACAS, J.; NOVÁK, P.; SCHUBERT, V. *et al.*, Holocentromeres in *Rhynchospora* are associated with genome-wide
-

centromere-specific repeat arrays interspersed among euchromatin. **PNAS**. 112 (44) 13633-13638; DOI: 10.1073/pnas.1512255112, 2015

MARTÍ, E.; MILANI, D.; BARDELLA, V.B., ALBUQUERQUE, L; SONG, J.; PALACIOS-GIMENEZ, O.M. & CABRAL-DE-MELLO, D.C. Cytogenomic analysis unveils mixed molecular evolution and recurrent chromosome rearrangements shaping the multigene families on *Schistocerca* grasshopper genomes. **Evolution**, in press., 2021

MARTINS, C. Chromosomes and repetitive DNAs: a contribution to the knowledge of fish genome. In: Pisano E, Ozouf-Costaz C, Foresti F, Kapoor BG (Eds.). **Fish Cytogenetics**. Enfield, New Hampshire: Science Publisher, pp. 421-453, 2007

MELTERS, D.P.; PALIULIS L.V.; KORF, I.F. & CHAN, W.L. Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. **Chromosome Research**, v. 20, p. 579-593, 2012

MERLO, M.A., CROSS, I., MANCHADO, M. *et al.* The 5S rDNA High Dynamism in *Diplodus sargus* is a Transposon-Mediated Mechanism. Comparison with Other Multigene Families and Sparidae Species. **J Mol Evol** 76, 83–97, <https://doi.org/10.1007/s00239-013-9541-8>, 2013

MESSIER W.; Li, S. H. & STEWART, C. B. The birth of microsatellites. **Nature**. 381:483, 1996

MONTI, V.; LOMBARDO, G.; LOXDALE, H.D. *et al.* Continuous occurrence of intra-individual chromosome rearrangements in the peach potato aphid, *Myzus persicae* (Sulzer) (Hemiptera: Aphididae). **Genetica** 140, 93–103. <https://doi.org/10.1007/s10709-012-9661-x>, 2012

NANDA, I.; SCHLEGELMILCH, K.; HAAF, T.; SCHARTL, M. & SCHMID, M. Synteny conservation of the Z chromosome in 14 avian species (11 families) supports a role for Z dosage in avian sex determination. **Cytogenet Genome Res**. 122:150-156. doi: 10.1159/000163092, 2008

NEI, M. & ROONEY, A. P. Concerted and birth-and-death evolution of multigene families. **Ann Rev Genetics**, v.39 p.121-152, dec. 2005

NGUYEN, P. & CARABAJAL PALADINO, L. On the Neo-Sex Chromosome of Lepidoptera. In: Pontarotti P. (eds) **Evolutionary Biology**. Springer, Cham. https://doi.org/10.1007/978-3-319-41324-2_11, 2016

NGUYEN, P.; SAHARA, K.; YOSHIDO, A. & MAREC, F. Evolutionary dynamics of rDNA clusters on chromosomes of moths and butterflies (Lepidoptera). **Genetica**, 138: 343-354, 2010

NGUYEN, P.; SYKOROVÁ, M.; ŠÍCHOVÁ, J.; *et al.* Neo-sex chromosomes and adaptive potential in tortricid pests. **Proceedings of the National Academy of Sciences**, v. 110(17), p. 6931-6936, 2013

NOVÁK, P.; NEUMANN, P. & MACAS, J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. **BMC Bioinformatics**, v. 11, p. 378, 2010

NOVÁK, P.; NEUMANN, P.; PECH, J.; STEINHAIŠL, J. & MACAS, J. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next generation sequence reads. **Bioinformatics**, v. 29, p. 792-793, 2013

NOVÁK, P.; ROBLIEDILLO, L. A.; KOBLIZKOVÁ, A. *et al.* TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. **Nucleic Acids Res** 45(12), e111. doi: 10.1093/nar/gkx257, 2017

OHNO S. Ancient linkage groups and frozen accidents. **Nature** 244:259–262, 1973

OSANAI-FUTAHASHI, M.; SUETSUGU, Y.; MITA, K. & FUJIWARA H. Genome-wide screening and characterization of transposable elements and their distribution analysis in the silkworm, *Bombyx mori*. **Insect Biochem Mol Biol.** (12):1046-57. doi: 10.1016/j.ibmb.2008.05.012. PMID: 19280695, 2008

PALACIOS-GIMENEZ, O. M., MILANI, D., SONG, H., MARTÍ, D. A., LÓPEZ-LEÓN, M. D., RUIZ-RUANO, F. J., *et al.* Eight million years of satellite DNA evolution in grasshoppers of the genus *Schistocerca* illuminate the ins and outs of the library hypothesis. **Genome Biol. Evol.** 12, 88–102. doi: 10.1093/gbe/evaa018, 2020

-
- PALACIOS-GIMENEZ, O. M.; BARDELLA, V. B.; LEMOS, B. & CABRAL-DE-MELLO, D. C. Satellite DNAs are conserved and differentially transcribed among *Gryllus* cricket species. **DNA Research**. DOI: 10.1093/dnares/dsx044, 2017
- PEZER, Z.; BRAJKOVIC, J.; FELICIELLO, I. & UGARKOVIC, D. Satellite DNAMediated Effects on Genome Regulation. **Genome Dyn.** 7, 153-169, 2012
- PLOHL, M.; LUCHETTI, A.; MESTROVIC, N. & MANTOVANI, B. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. **Gene**, v. 409, p. 72-82, 2008
- PLOHL, M.; MESTROVIC, N. & MRAVINAC, B. Centromere identity from the DNA point of view. **Chromosoma**. 123, 313-325, 2014
- POMBO, A. & DILLON, N. Three-dimensional genome architecture: players and mechanisms. **Nat Rev Mol Cell Biol** 16, 245–257. <https://doi.org/10.1038/nrm3965>, 2015
- PRINGLE, E. G.; BAXTER, S. W.; WEBSTER, C. L. *et al.* Synteny and Chromosome Evolution in the Lepidoptera: Evidence From Mapping in *Heliconius melpomene*. **Genetics** 177(1), 417-426. <https://doi.org/10.1534/genetics.107.073122>, 2007
- ROBINSON, R. Lepidoptera Genetics. Oxford: **Pergamon Press**, 1ed. 1971
- RUIZ-RUANO, E. J.; LÓPEZ-LEÓN, M.D.; CABRERO, J. & CAMACHO, J. P. M. High-throughput analysis of the satellitome illuminates satellite DNA evolution. **Sci. Rep.**, 6, 28333, 2016
- SAHARA, K.; YOSHIDO, A.; MAREC, F. *et al.* Conserved synteny of genes between chromosome 15 of *Bombyx mori* and a chromosome of *Manduca sexta* shown by five-color BAC-FISH. **Genome** 50(11), 1061-65. doi:10.1139/g07-082, 2007
- SAHARA, K.; YOSHIDO, A. & TRAUT, W. Sex chromosome evolution in moths and butterflies. **Chromosome Res.** 20:83-94, 2012
- SERRATO-CAPUCHINA, A. & MATUTE, D.R. The Role of Transposable Elements in Speciation. **Genes**. 9(5):254. <https://doi.org/10.3390/genes9050254>, 2018
- SHAPIRO, J. & STERNBERG, R. Why repetitive DNA is essential to genome function. **Biol Rev Camb Philos Soc** 80: 227-250. doi:10.1017/S1464793104006657, 2005
-

-
- ŠICHOVÁ, J., VOLENÍKOVÁ, A., DINCA, V., NGUYEN, P., VILA, R., SAHARA, K., *et al.* Dynamic karyotype evolution and unique sex determination systems in Leptidea wood white butterflies. *BMC Evol. Biol.* 15:89. doi: 10.1186/s12862-015-0375-4, 2015
- ŚLIWINSKA, E.B.; MARTYKA, R. & TRVIANOWSKI, P. Evolutionary interaction between W/Y chromosome and transposable elements. *Genetica* 144, 267–278. <https://doi.org/10.1007/s10709-016-9895-0>, 2016
- SMIT, A. F.; HUBLEY, R. & GREEN, P. RepeatMasker Open-4.0. Available at <http://www.repeatmasker.org>, 2017
- SPERLING, F. Sex-linked genes and species-differences in Lepidoptera. *Can Entomol* 126(3), 807-818. doi:10.4039/Ent126807-3, 1994
- TRAUT, W. & MAREC, F. Sex chromatin in lepidoptera. *Q Rev Biol* 71(2), 239-256. doi:10.1086/419371, 1996
- TRAUT, W.; SAHARA, K. & MAREC, F. Sex chromosomes and sex determination in Lepidoptera. *Sex Dev.* 1:332-346, 2007
- TRAUT, W.; SAHARA, K.; OTTO, T. D. & MAREC, F. Molecular differentiation of sex chromosomes probed by comparative genomic hybridization. *Chromosoma* 108, 173-80. doi: 10.1007/s00412005036, 1999
- UGARKOVIC, D. Functional elements residing within satellite DNAs. *EMBO Rep.* 6(11), 1035-1039. doi:10.1038/sj.embor.7400558, 2005
- UGARKOVIC, D. & PLOHL, M. Variation in satellite DNA profiles-causes and effects. *EMBO J*, 21:5955-5959, 2002
- VIRKKI, N. Gametogenesis in the sugarcane borer moth, *Diatraea saccharalis* (F.) (Crambidae). *J. Agricult.* University Puerto Rico 47, 102–137. doi: 10.46429/jaupr.v47i2.12944, 1963
- VÍTKOVÁ, M.; FUKOVÁ, I.; KUBICKOVÁ, S. & MAREC, F. Molecular divergence of the W chromosomes in pyralid moths (Lepidoptera). *Chromosome Res* 15:917–930, 2007
-

-
- WHITE, M. A.; KITANO, J. & PEICHEL, C. L. Purifying Selection Maintains Dosage-Sensitive Genes during Degeneration of the Threespine Slickeback Y Chromosome. **Mol Biol Evol.** 32(8): 1981-95, 2015
- WICKER, T.; SABOT, F.; HUA-VAN, A. *et al.* A unified classification system for eukaryotic transposable elements. **Nature Reviews genetics** 8, 973-982. <https://doi.org/10.1038/nrg2165>, 2007
- WRENSCH, D. L.; KETHLEY, J. B. & NORTON, R. A. Cytogenetics of Holokinetic Chromosomes and Inverted Meiosis: Keys of the Evolutionary Success of Mites, with Generalizations on Eukaryotes. **Springer**. https://doi.org/10.1007/978-1-4615-2389-5_11, 1994
- WRIGHT, A.; DEAN, R.; ZIMMER, F. *et al.* How to make a sex chromosome. **Nat Commun** 7, 12087. <https://doi.org/10.1038/ncomms12087>, 2016
- YASUKOCHI, Y. OHNO, M. SHIBATA, F.; JOURAKU, A.; NAKANO, R.; ISHIKAWA, Y. & SAHARA, K. A FISH-based chromosome map for the European corn borer yields insights into ancient chromosomal fusions in the silkworm. **Heredity**, 116:75-83, 2016
- YASUKOCHI, Y.; ASHAKUMARY, L. A.; BABA, K.; YOSHIDO, A. & SAHARA, K. A Second-Generation Integrated Map of the Silkworm Reveals Synteny and Conserved Gene Order Between Lepidopteran Insects. **Genetics** 173(3), 1319-28. <https://doi.org/10.1534/genetics.106.055541>, 2006
- YOSHIDO, A.; MAREC, F. & SAHARA, K. Resolution of sex chromosome constitution by genomic in situ hybridization and fluorescence in situ hybridization with (TTAGG) n telomeric probe in some species of Lepidoptera. **Chromosoma** 114, 193–202. <https://doi.org/10.1007/s00412-005-0013-9>, 2005
- YOSHIDO, A.; SAHARA, K.; MAREC, F. & MATSUBA, Y. Step-by-step evolution of neo-sex chromosome in geographical populations of wild silkmooths, *Samia cynthia* spp. **Heredity**. 106:614-624, 2011
- ZEDEK, F. & BURES, P. Absence of positive selection on CenH3 in *Luzula* suggests that holokinetic chromosomes may suppress centromere drive, **Annals of Botany**, Volume 118, Issue 7, Pages 1347–1352, <https://doi.org/10.1093/aob/mcw186>, 2016
-

ZHANGWU, Z. *et al.* Chromosome studies of the millet stem borer, *Chilo infuscatellus* Snellen (Lepidoptera: Pyralidae). **Journal of Shanxi Agricultural University**. 14:394-396. DOI: 10.13842/j.cnki.issn1671-8151.1994.04.017, 1994
