Chromosoma (2022) 131:253–267
https://doi.org/10.1007/s00412-022-00781-4

RESEARCH

A step forward in the genome characterization of the sugarcane borer, *Diatraea saccharalis*: karyotype analysis, sex chromosome system and repetitive DNAs through a cytogenomic approach

Ana E. Gasparotto[1] · Diogo Milani[1] · Emiliano Martí[1] · Ana Beatriz S. M. Ferretti[1] · Vanessa B. Bardella[1] · Frederico Hickmann[2] · Magda Zrzavá[3,4] · František Marec[3] · Diogo C. Cabral-de-Mello[1,5]

Received: 22 June 2022 / Revised: 10 August 2022 / Accepted: 16 September 2022 / Published online: 11 October 2022

Abstract

Moths of the family Crambidae include a number of pests that cause economic losses to agricultural crops. Despite their economic importance, little is known about their genome architecture and chromosome evolution. Here, we characterized the chromosomes and repetitive DNA of the sugarcane borer *Diatraea saccharalis* using a combination of low-pass genome sequencing, bioinformatics, and cytogenetic methods, focusing on the sex chromosomes. Diploid chromosome numbers differed between the sexes, i.e., $2n = 33$ in females and $2n = 34$ in males. This difference was caused by the occurrence of a WZ_1Z_2 trivalent in female meiosis, indicating a multiple sex-chromosome system $WZ_1Z_2/Z_1Z_1Z_2Z_2$. A strong interstitial telomeric signal was observed on the W chromosome, indicating a fusion of the ancestral W chromosome with an autosome. Among repetitive DNAs, transposable elements (TEs) accounted for 39.18% (males) to 41.35% (females), while satDNAs accounted for only 0.214% (males) and 0.215% (females) of the genome. FISH mapping revealed different chromosomal organization of satDNAs, such as single localized clusters, spread repeats, and non-clustered repeats. Two TEs mapped by FISH were scattered. Although we found a slight enrichment of some satDNAs in the female genome, they were not differentially enriched on the W chromosome. However, we found enriched FISH signals for TEs on the W chromosome, suggesting their involvement in W chromosome degeneration and differentiation. These data shed light on karyotype and repetitive DNA dynamics due to multiple chromosome fusions in *D. saccharalis*, contribute to the understanding of genome structure in Lepidoptera and are important for future genomic studies.

Keywords Chromosome fusion · FISH · Holocentric chromosome · Multiple sex chromosomes · satDNA · W chromatin

Introduction

Repetitive DNA elements are sequences that repeat hundreds or thousands of times and make up a large proportion of the total nuclear DNA in eukaryotic genomes. These sequences may occur in long tandem arrays of similar motifs, such as satellite DNAs (satDNAs) and some multigene families, while others may be scattered throughout the genome as transposable elements (TEs) with medium to high repetition rates. All types of repetitive sequences can be enriched and localized in specific chromosomal regions, mainly in the heterochromatic regions, including centromeres and telomeres, and they are important organizational and functional components of genomes (Charlesworth et al. 1994; López-Flores and Garrido-Ramos 2012; Biscotti et al. 2015). Moreover, the accumulation, expansion, and lability of repetitive DNAs

✉ Diogo C. Cabral-de-Mello
cabral.mello@unesp.br

1 Departamento de Biologia Geral e Aplicada, Instituto de Biociências/IB, UNESP - Universidade Estadual Paulista, Rio Claro, SP 13506-900, Brazil

2 Department of Entomology and Acarology, Luiz de Queiroz College of Agriculture, (USP/ESALQ), University of São Paulo, Piracicaba, SP, Brazil

3 Biology Centre of the Czech Academy of Sciences, Institute of Entomology, České Budějovice, Czech Republic

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

5 Department of Experimental Biology, Genetics Area, University of Jaén, Paraje las Lagunillas s/n, 23071 Jaén, Spain

contribute to the formation and differentiation of heteromorphic sex chromosomes, as described in various species from most eukaryote groups (e.g., Charlesworth et al. 1994; Kejnovský et al. 2013; Chalopin et al. 2015; Hobza et al. 2015; Dalíková et al. 2017a; Palacios-Gimenez et al. 2017; Ferretti et al. 2020).

The Lepidoptera (moths and butterflies) are an extremely diverse insect order with more than 160,000 valid species (Van Nieukerken et al. 2011). From a karyotypic point of view, the chromosomes of Lepidoptera are numerous (with a modal number of $2n = 60–62$), small and uniform, without morphological landmarks, such as centromeres, i.e., holocentric chromosomes (Robinson 1971; Lukhtanov 2000; Traut et al. 2007; Sahara et al. 2012). This order represents the largest group with female heterogamety, with most species having a WZ/ZZ sex chromosome system. This differs from most other insect groups that have male heterogamety, with the exception of the Trichoptera (caddisflies), a sister order of Lepidoptera, which also have heterogametic females (Traut et al. 2007; Marec et al. 2010; Šíchová et al. 2015). Chromosome rearrangements, such as fusions, fissions, and inversions, have led to the dynamic karyotype evolution with the occurrence of diverse chromosome numbers and derived sex chromosome systems in Lepidoptera (Marec et al. 2001; Yasukochi et al. 2006; Sahara et al. 2007, 2012; Lukhtanov 2015; Šíchová et al. 2015; Hill et al. 2019). This is particularly evident in some groups, such as *Agrodiaetus* sp. (Lycaenidae) with $n = 10$ to $n = 134$ (Kandul et al. 2007), *Godyris* sp. (Nymphalidae) with $n = 13$ to $n = 120$ (Brown et al. 2004), and *Lysandra* sp. (Lycaenidae) with $n = 24$ to $n = 93$ (Talavera et al. 2013). In the case of *Leptidea* sp. (Pieridae), in addition to a large variability in diploid chromosome number ($2n = 51$ to $2n = 208$), different but species-specific systems of multiple sex chromosomes were found, including W_1-$_3Z_1$-$_4/Z_1$-$_4Z_1$-$_4$, W_1-$_3Z_1$-$_3/Z_1$-$_3Z_1$-$_3$, W_1-$_4Z_1$-$_4/Z_1$-$_4Z_1$-$_4$, and W_1-$_3Z_1$-$_6/Z_1$-$_6Z_1$-$_6$ (Šíchová et al. 2015, 2016; Yoshido et al. 2020).

Repetitive DNAs may be involved in the organization of the genome and its plasticity, as they may be frequent sites for chromosomal rearrangements. Despite the large number of chromosomal studies conducted in Lepidoptera and the extensive genome sequencing in this group, understanding the chromosomal organization of some repetitive DNA fractions remains elusive, especially for satDNAs and TEs. To date, most cytogenetic studies in Lepidoptera have mapped mainly multigene families, including rDNAs and H3 histone genes (Nguyen et al. 2010; Provazníková et al. 2021), and only 12 satDNAs have been identified (Cabral-de-Mello et al. 2021). These studies have enabled the generation of markers for comparative analyses useful for understanding genomic organization and pattern changes between species. Specifically for sex chromosomes, although the W chromosome is known to be largely heterochromatic and rich in DNA repeats, the specific content of this chromosome is poorly understood. Nevertheless, some TEs, satDNAs, microsatellites, and, in a few cases, rDNA clusters have been identified on the W chromosome (Abe et al. 2005; Traut et al. 2013; Věchtová et al. 2016; Dalíková et al. 2017b; Zrzavá et al. 2018; Cabral-de-Mello et al. 2021).

Grass moths of the family Crambidae include a number of pest species that cause severe economic losses to agricultural crops (Munroe and Solis 1999). Despite the economic importance of crambid pests, very little is known about their genome architecture, as only a few cytogenetic studies have been carried out. In the few species studied so far, the haploid chromosome number ranged from $n = 10$ to $n = 41$, the WZ/ZZ sex chromosome system has been determined for four species, and a chromosome map based on fluorescence in situ hybridization (FISH) has been constructed for *Ostrinia nubilalis* (Virkki 1963; Robinson 1971; Kageyma and Traut 2004; Nguyen et al. 2010; Yasukochi et al. 2016; Cabral-de-Mello et al. 2021). Recently, a combination of cytogenetics and genomic approaches in four Crambidae species revealed differences in gross genome architecture (including chromosome numbers) and high variability in repetitive DNAs, which may have contributed to the plasticity of sex chromosomes in this group (Cabral-de-Mello et al. 2021).

Diatraea saccharalis, commonly known as the sugarcane borer, is one of the most important crambid pests responsible for significant crop damage and economic losses in maize and sugarcane (Flynn et al. 1984; Reagan and Mulcahy 2019). This moth is widespread throughout the Americas, including the Caribbean, South and Central America, and the southeastern United States (Myers 1935; Palacio Cortés et al. 2010). Because of this economic importance, *D. saccharalis* has attracted the attention of the scientific community for population, structural, and functional genetic studies (e.g., Joyce et al. 2014; Lopes et al. 2014; Merlin and Cônsoli 2019; Borges dos Santos et al. 2020; Noriega et al. 2020), but lacks a complete understanding of genome architecture at the chromosomal level, except for chromosome number, which is $n = 17$ (Virkki 1963). In this study, we have characterized the repeat content of the sugarcane borer using a combination of cytogenetic and bioinformatic methods, allowing a detailed description of the karyotype, sex chromosomes, and chromosomal distribution of repetitive DNAs. The data contribute to the understanding of genome organization and sex chromosome evolution in Crambidae and may assist in future structural and functional genomic analyzes in *D. saccharalis*, which could also be useful for pest management programs.

Materials and methods

Insects and chromosome preparations

Specimens of *Diatraea saccharalis* were obtained from the laboratory culture of the Luiz de Queiroz College of Agriculture, University of São Paulo – ESALQ/USP (Piracicaba, São Paulo, Brazil). Larvae were fed an artificial diet and maintained in a BOD (Bio-Oxygen Demand) incubator (Solab, Piracicaba, Brazil) under controlled conditions (temperature 25 ± 4 °C, 60% humidity, and a 12 h light/12 h dark regime) (see King et al. 1975).

Spread chromosome preparations were made as described in Mediouni et al. (2004) and Šíchová et al. (2013). Meiotic chromosomes at the pachytene stage were obtained from the testes of third instar larvae and ovaries of the last instar larvae and in the initial stage of pupation. Mitotic chromosomes were obtained from the wing imaginal discs of fifth instar larvae of both sexes. Briefly, testes and wing imaginal discs were dissected in physiological solution, hypotonized in 75 mM KCl for 15 min and fixed in Carnoy's solution (ethanol/chloroform/acetic acid, 6:3:1) for 15 min. The ovaries were transferred to Carnoy's solution without hypotonic treatment to preserve the heterochromatin pattern of the putative W chromosome. 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. Slides of sufficient quality were dehydrated in an ethanol series (70%, 80%, and 100%, 30 s each) and stored at –20 °C until use. All bodies of dissected larvae were stored in 100% ethanol in a freezer at –20 °C until extraction of genomic DNA (gDNA).

Preparation of polyploid nuclei and microdissection of sex chromatin

Malpighian tubules from the 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's fixative on a slide for about 1 min and then stained with 1.25% lactic acetic orcein for 5 min. The slides were inspected under a light microscope for the presence or absence of female-specific sex chromatin, formed by multiple copies of the W chromosome (Marec and Traut 1994; Traut and Marec 1996).

For microdissection of sex chromatin, we used Malpighian tubules dissected from female larvae of the last instar, swollen for 10 min in hypotonic solution (75 mM KCl), and fixed in Carnoy's fixative for 15 min. The tubules were dissociated with tungsten needles in a drop of 60% acetic acid, spread on coverslips using 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 to a Nikon Axiphot inverted microscope.

Low-pass genome sequencing, identification and analysis of satDNAs and TEs

Genomic DNA (gDNA) was extracted from a female larva and a male larva of the third instar using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI) according to the manufacturer's protocol. Genomic libraries were prepared using the NEBNext Ultra DNA Library Prep Kit (Biolabs, Cambridge, MA). The genomes of both sexes of *D. saccharalis* were sequenced by paired-end sequencing (2 × 150 bp) using the Illumina HiSeq 4000 system, performed by Novogene (HK) Co., Ltd. (Hong Kong, China). Sequencing yielded 1.49 Gb for the female genome and 943 Mb for the male genome, corresponding to coverage of ~ 4.15 × for female and ~ 2.63 × for male genomes, as the genome of the species is 359 Mb in size (Borges dos Santos et al. 2020). The genome reads were deposited in the Sequence Read Achieve (SRA) under accession numbers SAMN30196781 and SAMN30196782. Reads were processed, filtered, and interlaced using the tools available on the Galaxy/RepeatExplorer platform (Novák et al. 2013). Identification of satDNAs was performed separately using 500,000 reads for each sex as input. After clustering with RepeatExplorer, the satDNAs identified with the TAREAN tool (Novák et al. 2017) were selected for subsequent analysis. In addition, the tandem arrangement was checked using dotplot and Tandem Repeats Finder, TRF (Benson 1999). RepeatMasker (Smit et al. 2017) and the Python script "rm_homology.py" (https://github.com/fjruizruano/satminer/blob/master/rm_homology.py) were used to perform all-against-all comparison of monomers of the recovered satDNAs to analyze similarities and define all satDNA families.

We also used RepeatMasker to calculate the abundance for each satDNA, comparing the female and male genomes. We used the calcDivergenceFromAlign.pl script from the RepeatMasker utilities tool (Smit et al. 2017) to estimate the average Kimura 2-parameter (K2P) distances. The K2P model (Kimura 1980) takes into account transition and transversion substitution rates and assumes that the four nucleotide frequencies are the same and that substitution rates do not vary between sites. To do this, we randomly selected approximately 2.65 million previously preprocessed read pairs per library obtained using 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 a genome fraction calculated from the number of nucleotides mapped relative to the library size. We also compared the divergence of

satDNAs between sexes and created repeat landscapes representing the relative abundance of repeat elements on one axis and the K2P distance from consensus on the other axis. The satDNAs families were named according to descending abundance, taking into account the mean value of abundance between the sexes. The satDNAs presumed to be associated with the W chromosome of females were examined by looking for bias in the female/male abundance ratio for each satDNA family.

We identified and quantified TEs in both sexes using dnaPipeTE (Goubert et al. 2015). dnaPipeTE relies on Trinity (Grabherr et al. 2011) to assemble the most abundant sequences (i.e., repetitive sequences), followed by homology-based characterization (Goubert et al. 2015). For automated homology-based annotation of contigs, we used a custom TEs library as described elsewhere (Martí et al. 2021). TEs superfamilies were then aligned to Illumina libraries of both sexes using RepeatMasker (Smit et al. 2017), and K2P divergence and normalization for both sexes were estimated as described above. To detect TEs superfamilies potentially associated with the W chromosome, we estimated the female/male abundance ratio for each TE group.

We checked the similarity of the identified satDNAs with previously described sequences by a comparative analysis with nucleotide sequences deposited in the Repbase/GIRI or NCBI databases. Finally, to verify whether the genome of *D. saccharalis* harbors the two satDNA families DposSat01 (GenBank accession number MW369068) and DposSat02 (GenBank accession number MW369069) present in the genome of its sister species *D. postlineella* (Cabral-de-Mello et al. 2021), we mapped them to the genome of *D. saccharalis* using RepeatMasker (Smit et al. 2017).

DNA probes and chromosomal mapping by fluorescence in situ hybridization

We used non-species-specific probes to localize conserved sequences, including telomeric repeats using the $(TTAGG)_n$ probe and major ribosomal DNA using the 18S rDNA probe. We also prepared specific probes, such as the W-chromosome painting probe, male and female gDNA probes, satDNAs probes, and TE-derived probes. For TEs, we chose the most representative superfamily of TEs in the species genome and the most enriched TEs in the female library compared to the male library. For satDNAs, we chose the five most abundant families in the *D. saccharalis* genome, as well as sequences enriched in the female genome. These probes were isolated and mapped to chromosomes of *D. saccharalis* by fluorescence in situ hybridization (FISH).

18S rDNA and satDNA probes were amplified by PCR from female gDNA and labeled by nick-translation with biotin-14-dATP (Invitrogen, San Diego, CA). The 18S rDNA probe was obtained using primers described in Cabral-de-Mello et al. (2011). For amplification of satDNAs and TEs, we designed primers manually or using Geneious software with the consensus sequence of each satDNA (Supplementary Table 1). Each PCR was performed using a mixture of 10 × PCR Rxn Buffer, 0.2 mM $MgCl_2$, 0.16 mM dNTPs, 2 mM of each primer, 1 U of *Taq* Platinum DNA Polymerase (Invitrogen), and 50–100 ng/µL of template DNA. PCR conditions 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), with a final extension at 72 °C for 5 min. PCR products were separated on a 1% electrophoretic agarose gel. For seven satDNAs, we used synthetic oligonucleotides directly labeled with biotin-14-dATP at the 5' end designed in conserved and exclusive regions of the satDNAs (Supplementary Table 1).

Genomic DNA extracted from male and female larvae was 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. Labeled gDNA probes were used for comparative genomic hybridization, which was performed according to the protocol described in Traut et al. (1999), with modifications described in Dalíková et al. (2017a). W-chromosome painting probes were prepared as described in Fuková et al. (2007). Samples of manually microdissected W chromatin were subjected to PCR amplification using the GenomePlex Single Cell Whole Genome Amplification Kit, WGA4 (Sigma-Aldrich, St. Louis, MO). The amplified product was labeled with either digoxigenin-11-dUTP (Roche) or biotin-14-dATP (Invitrogen) by PCR using the GenomePlex WGA Reamplification Kit, WGA3 (Sigma-Aldrich). In addition, we used a W-chromosome painting probe from *D. postlineella*, previously prepared by Cabral-de-Mello et al. (2021). We also combined W-painting with a $(TTAGG)_n$ telomeric probe to check for the presence or absence of interstitial telomeric sequences (ITS). The insect telomeric probe was synthesized by non-template PCR according to Ijdo et al. (1991) using the self-complementary primers $(TTAGG)_5$ and $(CCTAA)_5$ and labeled with digoxigenin-11-dUTP by nick translation.

Chromosome preparations from ovaries were used for 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), whereas probes labeled with biotin-14-dATP were detected with Alexa-Fluor 488-conjugated streptavidin (Invitrogen). In all FISH experiments, chromosomes were counterstained with DAPI (4',6-diamidino-2-phenylindole; Sigma-Aldrich) and slides were mounted in VECTASHIELD (Vector, Burlingame, CA). Images were captured in grayscale using a cooled DP70 digital camera coupled to the Olympus BX51 fluorescence microscope equipped with appropriate filter sets.

Images were pseudocolored, merged (chromosomes and probe signals), and optimized for brightness and contrast using Adobe Photoshop CC6. To characterize the patterns of repetitive DNA distribution, at least 10 pachytene were analyzed for each probe combination.

Results

Karyotype characteristics and multiple sex chromosomes of Diatraea saccharalis

The diploid chromosome number was determined by analysis and repeated counting of pachytene complements from testes and ovaries and mitotic metaphases from testes, ovaries, and wing imaginal discs. To facilitate identification of individual chromosomes, we used FISH with telomeric probes. Mitotic karyotypes consisted of chromosomes of different sizes, such as large, medium, and small. We found differences in chromosome number between sexes, i.e., $2n = 34$ in males (Fig. 1a) and $2n = 33$ in females (Fig. 1b). In some metaphases of both sexes, we detected two small chromosomes with a constriction that colocalized with the hybridization signals of the 18S rDNA probe (Fig. 1c), suggesting that the constriction represents the nucleolus organizer region. No condensed segments were observed in pachytene bivalents, indicating low abundance of heterochromatin and non-enrichment of the repetitive DNA fraction in certain regions of the genome. The exception was a chromosome highlighted with DAPI, corresponding to the W chromosome, which was clearly seen in pachytene oocytes (Fig. 1d). Besides the occurrence of canonical telomeric repeats in the terminal regions of the chromosomes, an interstitial signal was clearly observed in the putative W chromosome, indicating the presence of ITS (Fig. 1b, d–h). In addition, the difference in diploid number between males and females indicated the presence of a multiple sex chromosome system in the species, consisting of one W chromosome and two Z chromosomes (Z_1 and Z_2), i.e., the WZ_1Z_2 system. The multiple sex chromosomes and the ITS site in the W chromosome were confirmed by more detailed analysis (see below).

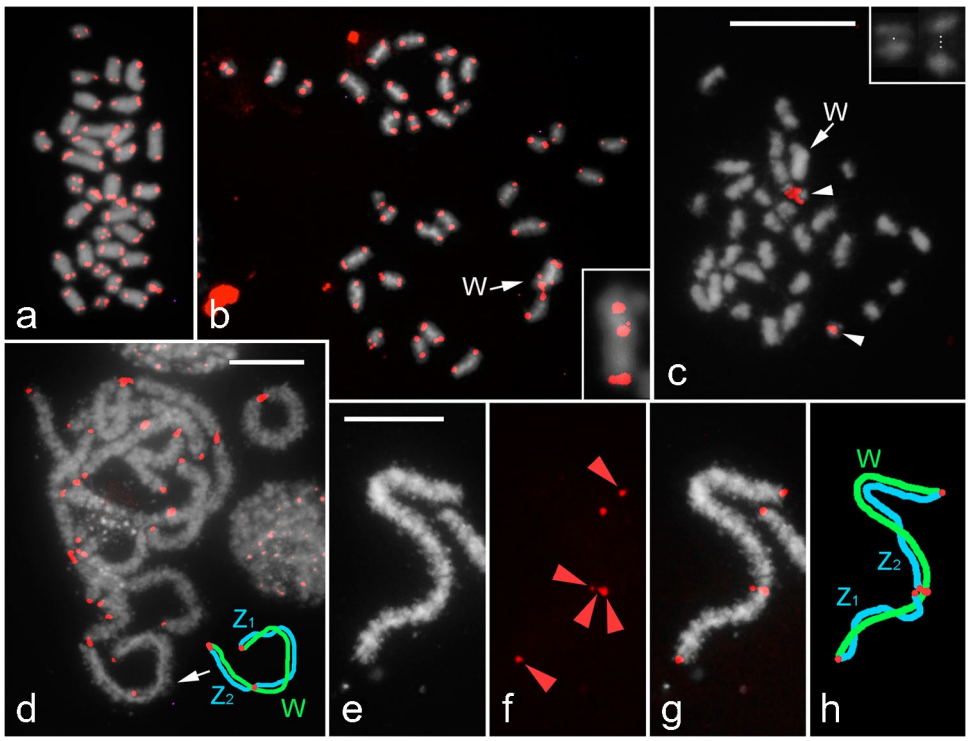

Fig. 1 Karyotype analysis of *Diatraea saccharalis*. Chromosomes were stained with FISH probes (red), either with the $(TTAGG)_n$ telomeric probe (**a**, **b, d**, **f**, **g**) or with the 18S rDNA probe (**c**) and counterstained with DAPI (gray). Mitotic metaphase chromosomes from wing discs of male (**a**) and female (**b**) larvae labeled with the telomeric probe, showing $2n = 34$ and $2n = 33$, respectively. In (**b**), the W chromosome is indicated and another W is magnified in the inset. Note interstitial hybridization signals indicating the presence of ITS on this chromosome. (**c**) Mitotic metaphase from ovary probed for 18S rDNA. Note two small autosomes identified by the probe and the large W chromosome. The inset shows two rDNA-bearing autosomes, each containing an apparent constriction (white dots). (**d**) Female pachytene with FISH-labeled telomeric sequences. Note interstitial signals of the telomeric probe in the sex chromosome trivalent, schematically shown on the right. In the WZ_1Z_2 trivalent, the W chromosome is deeply stained with DAPI. (**e**–**h**) Sex chromosome trivalent showing the distribution of hybridization signals of the telomeric probe: (**e**) DAPI, (**f**) $(TTAGG)_n$ signals indicated by red arrowheads, (**g**) merged image, (**h**) schematic representation. Bar = 10 µm

Preparations of polyploid nuclei from the Malpighian tubules of female larvae revealed a single conspicuous sex chromatin body of regular spherical shape, confirming the presence of the W chromosome. As expected, no sex chromatin was found in the polyploid nuclei of Malpighian tubules from males (Supplementary Fig. 1).

Comparative genomic hybridization used to examine and characterize the gross molecular differentiation of the W and Z chromosomes showed strong hybridization of the female gDNA-derived probe along the entire length of the W chromosome, whereas the hybridization signals of the male gDNA-derived probe were very weak (Fig. 2a–e). This

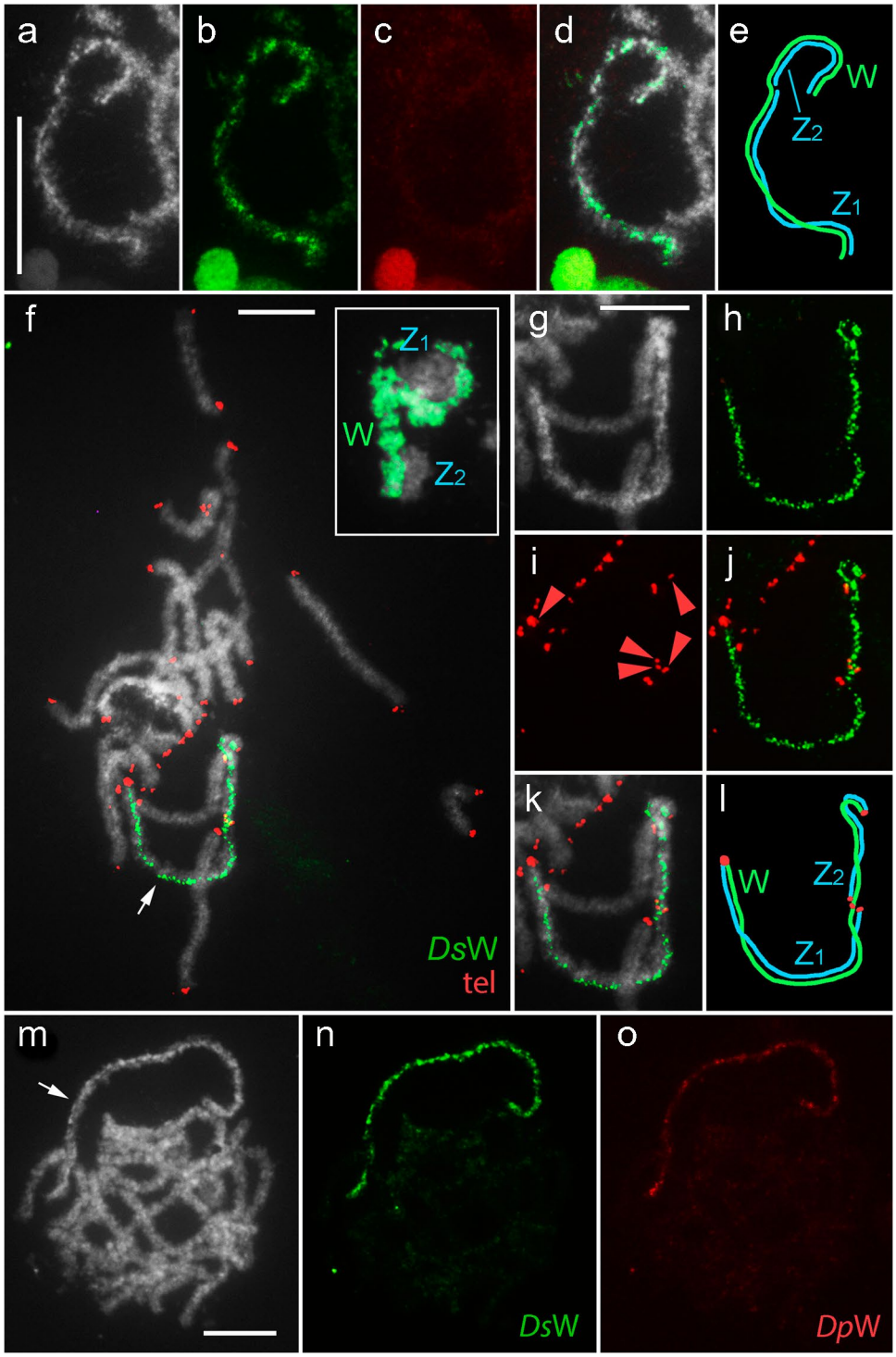

Fig. 2 Analysis of sex chromosomes of *Diatraea saccharalis*. (**a**–**e**) Sex chromosome trivalent from pachytene oocyte stained by comparative genomic hybridization. (**a**) DAPI staining; (**b**) female-derived genomic probe (green); (**c**) male-derived genomic probe (red); (**d**) merged image; (**e**) schematic representation of the WZ_1Z_2 trivalent. (**f**–**l**) Pachytene oocyte stained by two-color FISH with the W-painting probe (green) and the $(TTAGG)_n$ telomeric probe (red). (**f**) Complete pachytene; the inset shows a selected WZ_1Z_2 trivalent from a nurse cell of female ovary. (**g**) WZ_1Z_2 trivalent shown in (**f**) stained with DAPI, (**h**) W-painting probe, (**i**) telomeric probe (hybridization signals located on sex chromosomes are indicated by red arrowheads), (**j**) merged image of W-painting and telomeric signals, (**k**) merged image of both probes and DAPI counterstaining, (**l**) schematic representation of the WZ_1Z_2 trivalent. (**m**–**o**) Cross hybridization of W-painting probe from *D. postlineella* in female pachytene of *D. saccharalis*. (**m**) Chromosomes counterstained with DAPI; (**n**) W-painting probe from *D. saccharalis*; (**o**) W-painting probe from *D. postlineella*. Note weak signals of the W-painting probe from *D. postlineella* along the entire length of the *D. saccharalis* W chromosome. Bar = 10 µm

suggests that the W chromosome of *D. saccharalis* consists mainly of female-specific DNA sequences.

The W-painting probe strongly highlighted the entire W chromosome, which allowed easy identification of this sex chromosome and confirmed the molecular differentiation of the W chromosome from other chromosomes as well as the Z chromosome (Fig. 2f). In addition, W-painting combined with the telomeric probe revealed a strong interstitial signal of the telomeric probe in the W chromosome, confirming the presence of ITS on this chromosome. Near the ITS site of the W chromosome, we observed two additional telomeric signals corresponding to the ends of the two Z chromosomes, confirming the presence of the WZ_1Z_2 trivalent (Fig. 2g–l). In nurse cells, the W chromosome was clearly associated with two Z chromosomes that differed slightly in size (Fig. 2f, inset).

Cross-hybridization with a W-painting probe derived from the W chromatin of the sister species *Diatraea postlineella* showed hybridization signals along the entire length of the W chromosome of *D. saccharalis*. The intensity of the signals was similar along the entire W chromosome, but much weaker compared to the signals of the W-painting probe from *D. saccharalis* (Fig. 2m–o).

Genomic and chromosomal characterization of satDNAs and TEs

By clustering analysis in RepeatExplorer using the TAREAN report, we identified a total of 14 putative satDNAs on *D. saccharalis*, including male and female genomic libraries (GenBank accession numbers OP172884–OP172897). We did not find any relevant similarity with the previously described repetitive elements, not even in RepBase/GIRI or NCBI. Two satDNA repeats previously identified in the *D. postlineella* genome were also found in the *D. saccharalis* genome. The identified satDNA sequences represent only a small fraction of the male and female genomes, 0.214% and 0.215%, respectively. The monomers of these satDNAs were highly variable in size, ranging from 12 to 1327 bp, mostly with A + T rich content ranging from 39.4 to 72.5% (Table 1). Considering the mean between sexes, the abundance of different satDNA families in the *D. saccharalis* genome ranged from 0.0002 to 0.0570% and the K2P distance from 3.96 to 17.85%. The satellite DNAs landscape (abundance vs. divergence), considering both sexes, showed a higher peak of abundance at low divergence (around zero) and a smaller peak of abundance between 6 and 10% of divergence. Repeats with higher divergence tended to decrease in amounts from 10% of K2P (Fig. 3a).

Eight of 14 satDNAs families (DsaSat01-30, DsaSat04-901, DsaSat07-294, DsaSat08-338, DsaSat10-84, DsaSat11-12, and DsaSat13-134) were slightly enriched in the female genome (Fig. 3b), indicating possible enrichment on the W chromosome. However, the enrichment of these repeats on the W chromosome was ruled out by FISH, as with DsaSat01-30 (Fig. 3c) and DsaSat04-901 (Fig. 3e), one autosomal cluster was observed, terminally and interstitially located, respectively. The other female-enriched satDNAs (DsaSat07-294, DsaSat08-338, DsaSat10-84, DsaSat11-12, and DsaSat13-134) showed no visible FISH signals, and FISH signals were also absent for DsaSat03-32 (results not shown). Finally, DsaSat02-774 and DsaSat05-1327 showed scattered signals in all chromosomes, without clear evidence of enrichment in specific chromosomes or chromosomal regions (Fig. 3d, f).

TEs were much more abundant than satDNAs in the genome of *D. saccharalis*. They made up about 41.35% of the female genome and 39.18% of the male genome

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

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

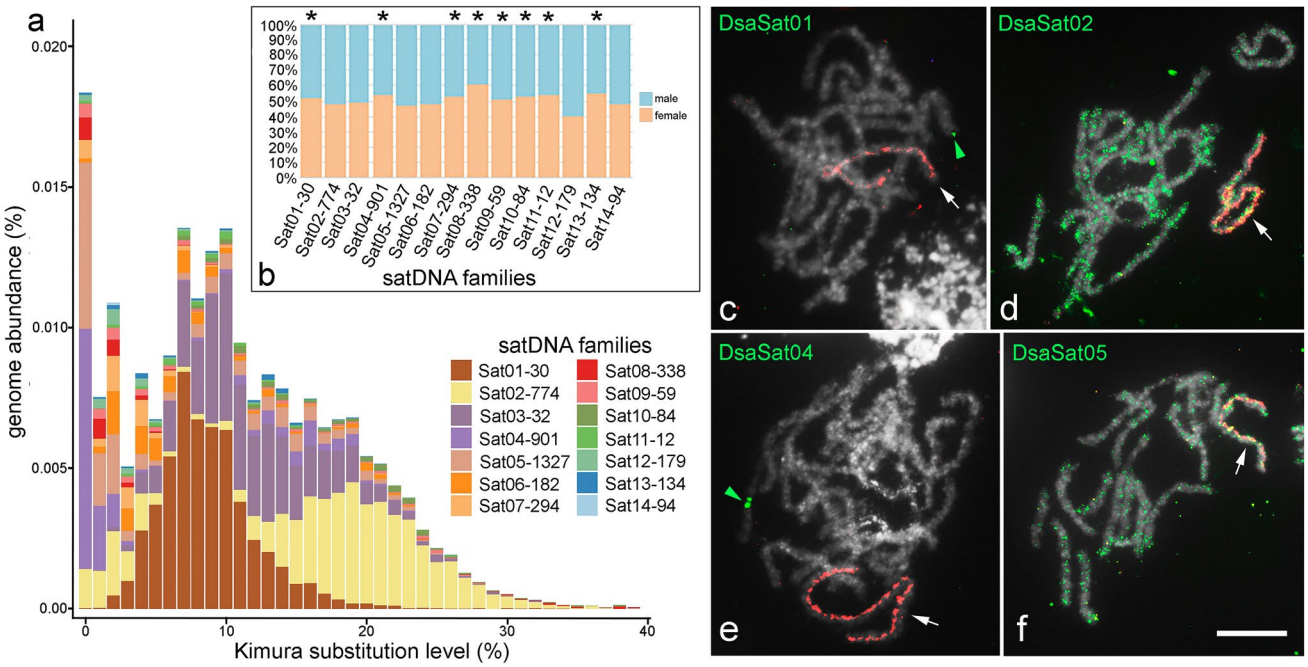

Fig. 3 SatDNAs in the genome of *Diatraea saccharalis*, genomic characteristics (**a**, **b**) and chromosomal mapping (**c**–**f**). (**a**) satDNAs landscape (mean of abundance *versus* K2P divergence of male and female genomes). Note two peaks of abundance, one with K2P divergence of 0 and the other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

and included at least 123 elements, taking into account unknown (not annotated) elements. *LINEs* and *Helitrons* were the most representative TEs (about 10% each), followed by LTR, which accounted for about 6% of the *D. saccharalis* genome. However, retroposons were extremely underrepresented (mean of 0.0047% of abundance). The number of elements for the different types of TEs correlated with their abundance (Table 2, Supplementary Table 2). Almost no differences in abundance were observed between male and female genomes and the number of elements for the different classes of TEs (Table 2). However, four elements showed enrichment in the female genome, with an abundance ratio > 2, namely *Helitron-2*, *TcMarStowaway*, *Maverik*, and *TcMar-Sagan* (Supplementary Table 2). The comparative TE landscape between female and male genomes clearly showed a recent amplification of some TEs in the female genome, with the occurrence of some elements with higher abundance and low divergence (Fig. 4a). FISH mapping of the most abundant TEs, i.e., *Helitron/Unknown* (GenBank accession numbers OP172899) with a mean abundance in the genome of 10.774% and DNA/*TcMar-Stowaway* (GenBank accession numbers OP172898), DNA TE elements enriched in the female genome showed similar patterns on pachytene chromosomes from oocytes. Both TEs were distributed as small dots along virtually the entire length of all chromosomes, but hybridization signals on the W chromosome were more intense compared to other chromosomes, especially compared to the Z chromosome (Fig. 4b–e).

Table 2 Main TEs classes recovered from the genome of *Diatraea saccharalis* and their characteristics, including number of elements, genome abundance, 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	2295	2294	10.79105	10.75878	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

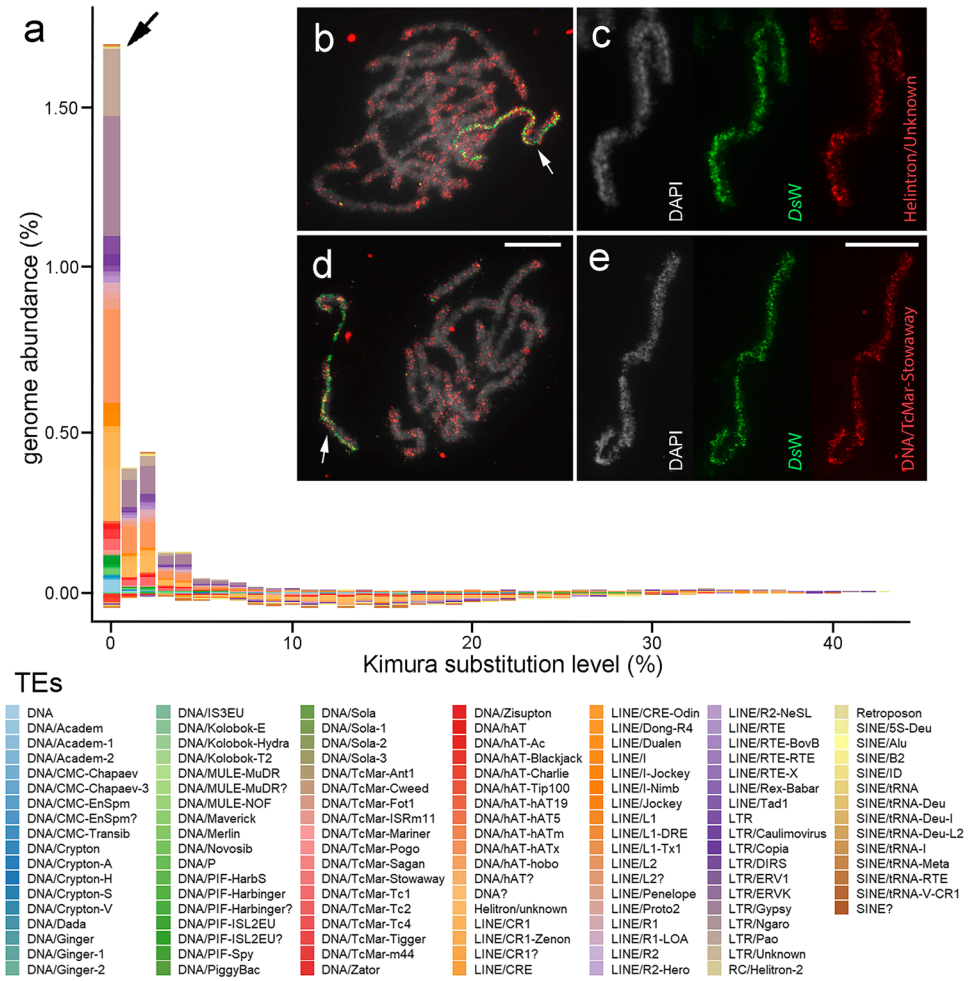

Fig. 4 TEs in the genome of *Diatraea saccharalis*. (**a**) Subtractive landscape (mean of abundance versus K2P divergence) obtained from female and male genomes. Abundance values are the difference between female and male genomes. Positive values indicate overabundant sequences in the female genome and negative values overabundant sequences in the male genome. Note overabundance of some TEs with low divergence in the female genome, suggesting recent amplification. (**b**, **d**) FISH mapping of two TEs elements (red) and W-painting probe (green) on pachytene oocytes. White arrows point to the sex chromosome trivalent. (**c**, **e**) Sex chromosome trivalents from pachytene oocytes showing in detail the distribution of the two TEs. Bar = 10 µm

Discussion

To date, few data have been published on genomes and karyotypes of Crambidae in general and the genus *Diatraea* in particular, although it is economically very important. To our knowledge, besides *D. saccharalis* (Virkki 1963), the only other member of the genus with a known karyotype is *D. postlineella* (Cabral-de-Mello et al. 2021). While the ancestral chromosome number in Lepidoptera is $2n = 62$ (Lukhtanov and Puplesene 1996) and most species have $2n = 60–62$ chromosomes (Robinson 1971), both *Diatraea* species show a large reduction in chromosome number, $2n = 42$ in *D. postlineella* (Cabral-de-Mello et al. 2021) and $2n = 34$ (male) and $2n = 33$ (female) in *D. saccharalis* (this study). This feature suggests chromosome fusions that have shaped the karyotypes of these species. However, with the exception of the W chromosome, we did not find any interstitial telomeric sites, suggesting their elimination. Species with chromosome numbers that deviate from the ancestral number are not uncommon in Lepidoptera, but surprisingly large deviations are not that common either (Robinson 1971), although lepidopteran chromosomes should be able to fuse and break up easily due to their holokinetic nature (Murakami and Imai 1974), without adverse consequences from the presence of di- or acentric chromosomes (reviewed in Wrensch et al. 1994; Zedek and Bureš 2018). Nevertheless, there are lepidopteran taxa whose karyotypes are exceptionally chromosome-poor or -rich, such as the Pieridae and Lycaenidae (Robinson 1971; Kandul et al. 2007). One possible trigger for chromosome rearrangements could be the activity of mobile elements (Zhang et al. 2011; Chénais et al. 2012). Interestingly, in *Leptidea juvernica* (Pieridae), a species from the genus known for sex chromosome rearrangements and karyotype variability, high TE genome content (38%), and activity were found (Talla et al. 2017). In *D. saccharalis*, the proportion of TEs in the genome is even higher (39.18% in males and 41.35% in females), making it and its relatives an excellent model for future studies on the potential role of TEs in chromosome rearrangements.

The six species of the Pyraloidea (Pyralidae and Crambidae) studied to date show great diversity in the number and location of major rDNA clusters. While one terminal cluster of rDNA is considered an ancestral feature in Lepidoptera (Provazníková et al. 2021), the Pyraloidea have either one

interstitial cluster (both *Diatraea* species), one terminal plus one interstitial cluster, or multiple terminal clusters (Nguyen et al. 2010; Provazníková et al. 2021; Cabral-de-Mello et al. 2021; this study). Because the single interstitial cluster is located on a small chromosome in both *Diatraea* species, we hypothesize that the chromosome carrying this rDNA cluster was involved in an inversion that led to its placement at the interstitial position, as has been proposed for other few Lepidoptera species (Nguyen et al. 2010). Fusion involving the rDNA-bearing chromosomes in the two *Diatraea* species is less likely, as these smallest chromosomes in the karyotype were probably not involved in fusions. In addition, ectopic recombination, a mechanism involved in the restructuring of rDNA chromosome 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), may also have occurred.

Regarding the sex chromosomes in *D. saccharalis*, the combination of ITS in the W chromosome together with one chromosome less in females indicates the presence of a neo-W chromosome, i.e., the $W_1Z_1Z_2$ (female)/$Z_1Z_1Z_2Z_2$ (male) sex chromosome system. This system most likely arose from the fusion of the W chromosome with an autosome, making the autosome homolog the Z_2 chromosome. Although most Lepidoptera species have a WZ/ZZ or Z0/ZZ sex chromosome system (Traut et al. 2007), neo-sex chromosomes have been repeatedly observed in Lepidoptera (reviewed in Marec et al. 2010; Yoshido et al. 2011; Sahara et al. 2012), and the growing number of Lepidoptera species with neo-sex chromosomes suggests that this is not an exceptional trait (e.g., Nguyen et al. 2013; Carabajal Paladino et al. 2019; Yoshido et al. 2020; Hejníčková et al. 2021). The parsimony chain of events in the genus *Diatraea* is that the reduction in diploid number to $2n = 42$ occurred in the common ancestor of *D. postlineella* and additional fusions occurred independently in the karyotype of *D. saccharalis*, which could also be shared with other species in the genus, but this requires further analysis.

Fusion of an autosome with the W chromosome, as observed in *D. saccharalis*, could play an important role in adaptive potential, contributing to ecological specialization and the formation of reproductive barriers, and eventually leading to speciation (Yoshido et al. 2011, 2020; Nguyen et al. 2013; Šíchová et al. 2015; Carabajal Paladino et al. 2019). One such example has been described in the family Tortricidae, where the origin of a neo-sex system included an autosome rich in genes that allow detoxification of insecticides and plant metabolites, linking these genes to sex inheritance and increasing the adaptive potential of tortricid moths (Nguyen et al. 2013). In *D. saccharalis*, it is particularly important to investigate the composition of the Z_2 chromosome, a homolog of the autosome involved in the formation of the neo-W to test for the possible presence of genes on this chromosome that may be related to the species' performance as a sugarcane and maize pest.

Although some information has been collected on the content of the W chromosome in Lepidoptera, providing insights into its evolution (reviewed in Traut et al. 2007; Nguyen and Carabajal Paladino 2016), the specific molecular composition of the W chromosome remains largely unresolved. There are several reasons for this. First, the major components of the W chromosome are clearly repetitive, as the W chromosome is typically rich in heterochromatin, as shown in pioneering studies that have examined the W chromosome in only a few species (Abe et al. 2005; Fuková et al. 2007; Traut et al. 2013). Because early genome sequencing platforms based on short reads did not allow assembly of repeat-rich regions, genome sequencing projects focused on male genomes to avoid the repeat-rich W chromosome. In addition, the high cost of sequencing mostly limited early genome sequences to model species. Second, unlike the Z chromosome, which is highly conserved between distant lepidopteran clades (Dalíková et al. 2017a), the W chromosome evolves rapidly and the randomness of degenerative processes leads to highly variable W chromosome compositions even in related species (e.g., Vítková et al. 2007; Zrzavá et al. 2018). Any study of the W chromosome in a new species is therefore uncharted territory with unpredictable results.

Our comparative genomic hybridization analysis revealed that the W chromosome in *D. saccharalis* is well differentiated from the Z chromosome and consists mainly of sequences enriched in the female genome, in contrast to species in which the W chromosome consists of sequences abundant in both the male and female genomes (e.g., Vítková et al. 2007; Hejníčková et al. 2021). Enrichment of the W chromosome with repeats derived from the female genome has been reported for other Crambidae (Cabral-de-Mello et al. 2021) and many other Lepidoptera (Yoshido et al. 2006; Dalíková et al. 2017a; Hejníčková et al. 2021), including species with neo-sex chromosome systems (Yoshido et al. 2011; Šíchová et al. 2016; Carabajal Paladino et al. 2019). This mainly reflects an accumulation of repetitive sequences, TEs in the case of *D. saccharalis* (this study) and probably satDNAs and/or TEs on the W chromosome in other species (Sahara et al. 2003; Mediouni et al. 2004; Abe et al. 2005; Dalíková et al. 2017b). In addition, cross-hybridization experiments have shown that the W chromosome of *D. saccharalis* shares at least some repetitive sequences with its congeneric species, *D. postlineella*. The complete painting of the W chromosome of *D. saccharalis* with the W probe of *D. postlineella* also suggests that repetitive sequences from the ancestral W chromosome shared by the two *Diatraea* species have populated the new portion of the W chromosome of *D. saccharalis* corresponding to the ancestral autosome involved in the fusion. Some degree of

W sequence conservation has been observed in other Crambidae species (Cabral-de-Mello et al. 2021) and in Pyralidae species (Vítková et al. 2007), in contrast to two congeneric geometrid species of *Abraxas*, which exhibited high molecular divergence between W chromosomes (Zrzavá et al. 2018), suggesting that the rate of molecular differentiation of W chromosomes varies between Lepidoptera groups.

Our combined analysis of low-pass genome sequencing, computational tools, and chromosome mapping provided a deeper understanding of the repeat content composition and chromosome structure in *D. saccharalis*. Because we were interested in the sequences responsible for sex chromosome differentiation, we performed a comparative analysis of the male and female genomes to identify the sequences whose abundance differed between the two sexes, thus revealing the composition of the W chromosome in this particular species. Our analysis of the sex chromosomes revealed that the W chromosome of *D. saccharalis* is heterochromatic, as in most Lepidoptera (Traut et al. 2007), but is not highly enriched in satDNAs compared to other chromosomes. Although we identified eight satDNAs that are slightly overrepresented in the female genome compared to the male genome, these repeats did not show any obvious enrichment on the W chromosome. Therefore, the observed variation in genome abundance might represent only interindividual variability and might not be related to sex. On the other hand, two TEs identified as female-enriched in silico hybridized strongly on the W chromosome of *D. saccharalis* compared to other chromosomes and decorated its entire length. This suggests that TEs are involved in the degeneration of the W chromosome and its differentiation from the Z chromosome in this species. An abundance of TEs on W chromosomes was also found in other lepidopteran species, *Bombyx mori* (Abe et al. 2005) and *Ephestia kuehniella* (Traut et al. 2013).

The presence of satDNAs in low abundance in the genome of *D. saccharalis* (0.214% in males and 0.215% in females) is apparently a common feature in Crambidae (Cabral-de-Mello et al. 2021) and is consistent with the low heterochromatin content observed in *D. saccharalis* and most Lepidoptera species, in which heterochromatin is present mainly on the W chromosome (Traut et al. 2007). This could also be due to the small genome size of this group (Gregory 2020) and the predominance of interspersed elements, i.e., TEs (Talla et al. 2017). This last assumption is supported by the large number of TEs identified in the *D. saccharalis* genome, representing approximately 40% of the species' genome. The number of interspersed repeats identified here is similar to those described in the first draft genome of *D. saccharalis* (Borges dos Santos et al. 2020).

The species studied so far in which more than one satDNA was mapped showed exclusively scattered or singly localized signals (Cabral-de-Mello et al. 2021). The situation in *D. saccharalis* is different because the chromosome organization of the repeats identified here (satDNAs and TEs) was variable. The location of satDNAs suggests that they had either (i) scattered signals decorating all chromosomes, as in most Lepidoptera (Mahendran et al. 2006; Věchtová et al. 2016; Cabral-de-Mello et al. 2021), (ii) were arranged in a single cluster, as in *D. postlineela*, *Plodia interpunctella*, and *Mamestra brassicae* (Mandrioli et al. 2003; Dalíková et al. 2017b; Cabral-de-Mello et al. 2021), or (iii) had no obvious signals on the chromosomes, i.e., non-clustered patterns or quite small clusters, below the resolution of FISH. This is the first intragenomic variation in the chromosomal distribution of satDNAs discovered in Lepidoptera. The reasons could be that this variability reflects the usual situation and was uncovered here due to the number of satDNAs described and mapped, which is the highest published to date in any Lepidoptera species. Alternatively, the variability of repeats could actually indicate unusual whole-genome dynamics, also suggested by chromosomal rearrangements. To confirm either of these possibilities, additional satDNAs from genomes of species with ancestral chromosome numbers need to be examined.

In summary, we present new comprehensive information on the karyotype, sex chromosomes, and repeat content of *D. saccharalis*, a pest that causes major economic losses in sugarcane. Overall, our data reveal a dynamic karyotype and repetitive DNAs in the genome of *D. saccharalis*. The occurrence of fusions in the karyotype of *D. saccharalis* also led to 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 from the rest of the genome essentially by the accumulation of TEs, providing new insights into the evolution of sex chromosomes in Lepidoptera. To do this, we used a chromosomics approach, a combination of bioinformatic genome analysis followed by in situ mapping of sequences of interest using molecular cytogenetics methods (Deakin et al. 2019). This approach, while laborious, provides a comprehensive view of the genome or a selected portion of it, such as the sex chromosomes, as in silico genomic analysis provides a wealth of potential markers that are then confirmed and localized directly on chromosomes. Because this workflow requires knowledge of both cytogenetics and bioinformatics, this type of study is still rare despite the enormous number of sequenced genomes. Since repetitive DNA is a difficult part of the genome to assemble, our data will contribute to future improvements in the assembly of genomic sequences of this important pest species, which could also be relevant for pest control programs.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00412-022-00781-4.

Author contribution Ana E Gasparotto: performed the research, analyzed the data and wrote the original draft.
Diogo Milani: performed the research and analyzed the data.
Emiliano Martí: performed the research and analyzed the data.
Ana BSM Ferretti: performed the research.
Vanessa B Bardella: performed the research.
Frederico Hickmann: performed the research.
Magda Zrzavá: analyzed the data and wrote the original draft.
František Marec: conceived, designed the study, analyzed the data and wrote the original draft.
Diogo C Cabral-de-Mello: conceived, designed the study, analyzed the data and wrote the original draft.
All authors revised and edited the text.

Funding This study was supported in part by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (process number 2019/19069–7), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES), Conselho Nacional de Desenvolvimento Científiico e Tecnológico (CNPq). DM, VBB, ABSMF, and FH were supported by FAPESP (processes numbers 2021/05908–7, 2018/21772–5, 2020/06188–5, and 2018/20351–6) and AEGS and EM by CAPES. DCC-d-M is a recipient of a research productivity fellowship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico-CNPq (process number 308290/2020–8). MZ and FM acknowledge support from grant 20-13784S of the Czech Science Foundation.

Data availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication All authors approved the final version of the manuscript.

Competing interests The authors declare no competing interests.

References

Abe H, Mita K, Yasukochi Y, Oshiki T, Shimada T (2005) Retrotransposable elements on the W chromosome of the silkworm, *Bombyx mori*. Cytogenet Genome Res 110:144–151. https://doi.org/10.1159/000084946

Bardella VB, Fernandes JAM, Cabral-de-Mello DC (2016) Chromosomal evolutionary dynamics of four multigene families in Coreidae and Pentatomidae (Heteroptera) true bugs. Mol Genet Genomics 291:1919–1925. https://doi.org/10.1007/s00438-016-1229-5

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 2:573–580. https://doi.org/10.1093/nar/27.2.573

Biscotti MA, Olmo E, Heslop-Harrison JS (2015) Repetitive DNA in eukaryotic genomes. Chromosome Res 23:415–420. https://doi.org/10.1007/s10577-015-9499-z

Borges dos Santos L, Viana JPG, Francischini FJB, Fogliata SV, Joyce AL, Pereira de Souza A, Murúa MG, Clough SJ (2020) A first draft genome of the sugarcane borer, *Diatraea saccharalis*. F1000Research 9:1269. https://doi.org/10.12688/f1000research.26614.1

Brown KS, Von Schoultz B, Suomalainen E (2004) Chromosome evolution in Neotropical Danainae and Ithomiinae (Lepidoptera). Hereditas 141:216–236. https://doi.org/10.1111/j.1601-5223.2004.01868.x

Cabral-de-Mello DC, Marec F (2021) 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

Cabral-de-Mello DC, Oliveira SG, de Moura R, 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. https://doi.org/10.1186/1471-2156-12-88

Cabral-de-Mello DC, Zrzavá M, Kubíčková S, Rendón P, Marec F (2021) The role of satellite DNAs in genomic architecture and sex chromosome evolution in Crambidae moths. Front Genet 12:661417. https://doi.org/10.3389/fgene.2021.661417

Carabajal Paladino L, Provazníková I, Berger M, Bass C, Aratchige NS, López SN, Marec F, Nguyen P (2019) Sex chromosome turnover in moths of the diverse superfamily Gelechioidea. Genome Biol Evol 11:1307–1319. https://doi.org/10.1093/gbe/evz075

Chalopin D, Volff JN, Galiana D, Anderson JL, Schartl M (2015) Transposable elements and early evolution of sex chromosomes in fish. Chromosome Res 23:545–560. https://doi.org/10.1007/s10577-015-9490-8

Charlesworth B, Sniegowski P, Stephan W (1994) The evolutionary dynamics of repetitive DNAs in eukaryotes. Nature 371:215–220. https://doi.org/10.1038/371215a0

Chénais B, Caruso A, Hiard S, Casse N (2012) The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. Gene 509:7–15. https://doi.org/10.1016/j.gene.2012.07.042

Dalíková M, Zrzavá M, Hladová I, Nguyen P, Šonský I, Flegrová M, Kubíčková S, Voleníková A, Kawahara AY, Peters RS, Marec F (2017a) New insights into the evolution of the W chromosome in Lepidoptera. J Hered 108:709–719. https://doi.org/10.1093/jhered/esx063

Dalíková M, Zrzavá M, Kubíčková S, Marec F (2017b) 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

Deakin JE, Potter S, O'Neill R, Ruiz-Herrera A, Cioffi MB et al (2019) Chromosomics: bridging the gap between genomes and chromosomes. Genes 10:627. https://doi.org/10.3390/genes10080627

Ferretti ABSM, Milani D, Palacios-Gimenez OM, Ruiz-Ruano FJ, Cabral-de-Mello DC (2020) High dynamism for neo-sex chromosomes: satellite DNA reveal complex evolution in a grasshopper. Heredity 125:124–137. https://doi.org/10.1038/s41437-020-0327-7

Flynn JL, Reagan TE, Ogunwolu EO (1984) Establishment and damage of the sugarcane borer (Lepidoptera: Pyralidae) in corn as influenced by plant development. J Econ Entomol 77:691–697. https://doi.org/10.1093/jee/77.3.691

Fuková I, Traut W, Vítková M, Kubíčková S, Marec F (2007) Probing the W chromosome of the coding moth, *Cydia pomonella*, with sequences from microdissected sex chromatin. Chromosoma 116:135–145. https://doi.org/10.1007/s00412-006-0086-0

Goubert C, Modolo L, Vieira C, ValienteMoro C, Mavingui P, Boulesteix M (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 Biol Evol 7:1192–1205. https://doi.org/10.1093/gbe/evv050

Grabherr MG, Haas BJ, Yassour M et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652. https://doi.org/10.1038/nbt.1883

Gregory TR (2020) Animal genome size database. Available online at: http://www.genomesize.com

Gunderina L, Golygina V, Broshkov A (2015) Chromosomal organization of the ribosomal RNA genes in the genus *Chironomus* (Diptera, Chironomidae). Comp Cytogenet 9:201–220. https://doi.org/10.3897/CompCytogen.v9i2.9055

Hejníčková M, Dalíková M, Potocký P, Tammaru T, Trehubenko M, Kubíčková S, Marec F, Zrzavá M (2021) Degenerated, undifferentiated, rearranged, lost: high variability of sex chromosomes in Geometridae (Lepidoptera) indentified by sex chromatin. Cells 10:2230. https://doi.org/10.3390/cells10092230

Hill J, Rastas P, Hornett EA et al (2019) Unprecedented reorganization of holocentric chromosomes provides insights into the enigma of lepidopteran chromosome evolution. Sci Adv 5:eaau3648. https://doi.org/10.1126/sciadv.aau3648

Hobza R, Kubat Z, Cegan R, Jesionek W, Vyskot B, Kejnovsky E (2015) Impact of repetitive DNA on sex chromosomes in plants. Chromosome Res 23:561–570. https://doi.org/10.1007/s10577-015-9496-2

Ijdo JM, Wells RA, Baldini A, Reeders ST (1991) Improved telomere detection using a telomere repeat probe $(TTAGGG)_n$ generated by PCR. Nucleic Acids Res 19:4780. https://doi.org/10.1093/nar/19.17.4780

Joyce AL, White WH, Nuessly GS, Alma Solis M, Scheffer SJ, Lewis ML, Medina RF (2014) Geographic population structure of the sugarcane borer, *Diatraea saccharalis* (F.) (Lepidoptera: Crambidae), in the Southern United States. PLoS One 9:e110036. https://doi.org/10.1371/journal.pone.0110036

Kageyma D, Traut W (2004) Opposite sex-specific effects of *Wolbachia* and interference with the sex determination of its host *Ostrinia scapulalis*. Proc Biol Sci 271:251–258. https://doi.org/10.1098/rspb.2003.2604

Kandul NP, Lukhtanov VA, Pierce NE (2007) Karyotypic diversity and speciation in *Agrodiaetus* butterflies. Evolution 61:546–559. https://doi.org/10.1111/j.1558-5646.2007.00046.x

Kejnovský E, Michalovova M, Steflova P, Kejnovska I, Manzano S, Hobza R, Kubat Z, Kovarik J, Jamilena M, Boris V (2013) Expansion of microsatellites on evolutionary young Y chromosome. PLoS ONE 8:e45519. https://doi.org/10.1371/journal.pone.0045519

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120. https://doi.org/10.1007/BF01731581

King EG, Brewer FD, Martin DF (1975) Development of *Diatraea saccharalis* [Lep.: Pyralidae] at constant tempratures. Entomophaga 20:301–306. https://doi.org/10.1007/BF02371955

Lopes DA, Cantagalli LB, Stuchi ALPB, Mangolin CA, Ruvolo-Takasusuki MCC (2014) Population genetics of the sugarcane borer *Diatraea saccharalis* (Fabr.) (Lepidoptera: Crambidae). Acta Sci Agron 36:189–194. https://doi.org/10.4025/actasciagron.v36i2.16211

López-Flores I, Garrido-Ramos MA (2012) The repetitive DNA content of eukaryotic genomes. Genome Dyn 7:1–28. https://doi.org/10.1159/000337118

Lukhtanov VA (2000) Sex chromatin and sex chromosome systems in nonditrysian Lepidoptera (Insecta). J Zool Syst Evol Res 38:73–79. https://doi.org/10.1046/j.1439-0469.2000.382130.x

Lukhtanov VA (2015) The blue butterfly *Polyommatus* (*Plebicula*) *atlanticus* (Lepidoptera, Lycaenidae) holds the record of the highest number of chromosomes in the non-polyploid eukaryotic organisms. Comp Cytogenet 9:683–690. https://doi.org/10.3897/CompCytogen.v9i4.5760

Lukhtanov VA, Puplesene YuV (1996) Karyotypical peculiarities and main features pf karyotype evolution in lepidopterans of the nepticuloid, tischerioid, gelechioid and tineoid complexes (Lepidoptera: Nepticuloidea, Tischerioidea, Gelechioidea s.l., Psychoidea-Gracillarioidea). Entomologitsheskoe Obozrenie 75:310–323

Mahendran B, Acharya C, Dash R, Ghosh SK, Kundu SC (2006) Repetitive DNA in tropical tasar silkworm *Antheraea mylitta*. Gene 370:51–57. https://doi.org/10.1016/j.gene.2005.11.010

Mandrioli M, Manicardi GC, 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. https://doi.org/10.1023/A:1022058032217

Marec F, Traut W (1994) Sex chromosome pairing and sex chromatin bodies in W-Z translocation strains of *Ephestia kuehniella* (Lepidoptera). Genome 37:426–435. https://doi.org/10.1139/g94-060

Marec F, Sahara K, Traut W (2010) Rise and fall of the W chromosome in Lepidoptera. In: Goldsmith MR, Marec F (eds) Molecular biology and genetics of the Lepidoptera. CRC, Boca Raton, pp 49–63

Marec F, Tothová A, Sahara K, Traut W (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. https://doi.org/10.1046/j.1365-2540.2001.00958.x

Martí E, Milani D, Bardella VB, Albuquerque L, Song H, Palacios-Gimenez OM, Cabral-de-Mello DC (2021) Cytogenomic analysis unveils mixed molecular evolution and recurrent chromosome rearrangements shaping the multigene families on *Schistocerca* grasshopper genomes. Evolution 75:2027–2041. https://doi.org/10.1111/evo.14287

Mediouni J, Fuková I, Frydrychová R, Dhouibi MH, Marec F (2004) Karyotype, sex chromatin and sex chromosome differentiation in the carob moth, *Ectomyelois ceratoniae* (Lepidoptera: Pyralidae). Caryologia 57:184–194. https://doi.org/10.1080/00087114.2004.10589391

Merlin BL, Cônsoli FL (2019) Regulation of the larval transcriptome of *Diatraea saccharalis* (Lepidoptera: Crambidae) by maternal and other factors of the parasitoid *Cotesia flavipes* (Hymenoptera: Braconidae). Front Physiol 10:1106. https://doi.org/10.3389/fphys.2019.01106

Munroe E, Solis MA (1999) The Pyraloidea. In: Kristensen N (ed) Lepidoptera, Moths and Butterflies, Vol. 1. Handbook of Zoology Vol. IV, Atrthropoda: Insecta. Walter de Gruyter & Co., Berlin, pp 233–256. https://doi.org/10.1515/9783110804744.233

Murakami A, Imai HT (1974) Cytological evidence for holocentric chromosomes of silkworms, *Bombyx mori* and *B. mandarina* (Bombycidae, Lepidoptera). Chromosoma 47:167–178. https://doi.org/10.1007/BF00331804

Myers J (1935) The ecological distribution of some South American grass and sugar-cane borers (*Diatraea* spp., Lep., Pyralidae). Bull Entomol Res 26:335–342. https://doi.org/10.1017/s0007485300036646

Nguyen P, Carabajal Paladino L (2016) On the neo-sex chromosomes of Lepidoptera. In: Pontarotti P (ed) Evolutionary Biology: Convergent Evolution, Evolution of Complex Traits, Concepts and Methods. Springer, Heidelberg, pp 171–185. https://doi.org/10.1007/978-3-319-41324-2_11

Nguyen P, Sahara K, Yoshido A, Marec F (2010) Evolutionary dynamics of rDNA clusters on chromosomes of moths and butterflies (Lepidoptera). Genetica 138:343–354. https://doi.org/10.1007/s10709-009-9424-5

Nguyen P, Sýkorová M, Šíchová J, Kůta V, Dalíková M, Čapková Frydrychová R, Neven LG, Sahara K, Marec F (2013) Neo-sex chromosomes and adaptive potential in tortricid pests. Proc Natl Acad Sci U S A 110:6931–6936. https://doi.org/10.1073/pnas.1220372110

Noriega DD, Arraes FBM, Antonino JD, Macedo LLP, Fonseca FCA, Togawa RC, Crynberg P, Silva MCM, Negrisoli AS Jr, Morgante CV, Grossi-de-Sa MF (2020) Comparative gut transcriptome analysis of *Diatraea saccharalis* in response to the dietary source. PLoS ONE 15:0235575. https://doi.org/10.1371/journal.pone.0235575

Novák P, Neumann P, Pech J, Steinhaisl J, Macas J (2013) RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next generation sequence reads. Bioinformatics 29:792–793. https://doi.org/10.1093/bioinformatics/btt054

Novák P, Robledillo LA, Koblížková A, Vrbová I, Neumann P, Macas J (2017) TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. Nucleic Acids Res 45:e111. https://doi.org/10.1093/nar/gkx257

Palacio Cortés AM, Zarbin PHG, Takiya DM, Bento MS, Guidolin AS, Consoli FL (2010) Geographic variation of sex pheromone and mitochondrial DNA in *Diatraea saccharalis* (Fab., 1974) (Lepidoptera: Crambridae). J Insect Physiol 56:1624–1630. https://doi.org/10.1016/j.jinsphys.2010.06.005

Palacios-Gimenez OM, Bardella VB, Lemos B, Cabral-de-Mello DC (2017) Satellite DNAs are conserved and differentially transcribed among *Gryllus* cricket species. DNA Res 25:137–147. https://doi.org/10.1093/dnares/dsx044

Panzera Y, Pita S, Ferreiro MJ, Ferrandis I, Lages C, Pérez R, Silva AE, Guerra M, Panzera F (2012) High dynamics of rDNA cluster location in kissing bug holocentric chromosomes (Triatominae, Heteroptera). Cytogenet Genome Res 138:56–67. https://doi.org/10.1159/000341888

Provazníková I, Hejníčková M, Visser S, Dalíková M, Carabajal Paladino LZ, Zrzavá M, Voleníková A, Marec F, Nguyen P (2021) Large-scale comparative analysis of cytogenetic markers across Lepidoptera. Sci Rep 11:12214. https://doi.org/10.1038/s41598-021-91665-7

Reagan TE, Mulcahy MM (2019) Interaction of cultural, biological, and varietal controls for management of stalk borers in Louisiana sugarcane. Insects 10:305. https://doi.org/10.3390/insects10090305

Robinson R (1971) Lepidoptera genetics. Pergamon Press, Oxford

Sahara K, Yoshido A, Kawamura N, Ohnuma A, Abe H, Mita K, Oshiki T, Shimada T, Asano S, Bando H, Yasukochi Y (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. https://doi.org/10.1007/s00412-003-0245-5

Sahara K, Yoshido A, Marec F, Fuková I, Zhang HB, Wu CC, Goldsmith MR, Yasukochi Y (2007) Conserved synteny of genes between chromosome 15 of *Bombyx mori* and a chromosome of *Manduca sexta* shown by five-color BAC-FISH. Genome 50:1061–1065. https://doi.org/10.1139/g07-082

Sahara K, Yoshido A, Traut W (2012) Sex chromosome evolution in moths and butterflies. Chromosome Res 20:83–94. https://doi.org/10.1007/s10577-011-9262-z

Šíchová J, Nguyen P, Dalíková M, Marec F (2013) Chromosomal evolution in tortricid moths: conserved karyotypes with diverged features. PLoS ONE 8:e64520. https://doi.org/10.1371/journal.pone.0064520

Šíchová J, Ohno M, Dincă V, Watanabe M, Sahara K, Marec F (2016) Fissions, fusions and translocations shaped the karyotype and multiple sex chromosome constitution of the northeast-Asian wood-white butterfly, *Leptidea amurensis*. Biol J Linn Soc 118:457–471. https://doi.org/10.1111/bij.12756

Šíchová J, Voleníková A, Dincă V, Nguyen P, Vila R, Sahara K, Marec F (2015) Dynamic karyotype evolution and unique sex determination systems in *Leptidea* wood white butterflies. BMC Evol Biol 15:89. https://doi.org/10.1186/s12862-015-0375-4

Smit AFA, Hubley R, Green P (2017) RepeatMasker Open-4.0. Available at http://www.repeatmasker.org

Talavera G, Lukhtanov VA, Rieppel L, Pierce NE, Vila R (2013) In the shadow of phylogenetic uncertainty: the recent diversification of *Lysandra* butterflies through chromosomal change. Mol Phylogenet Evol 69:469–478. https://doi.org/10.1016/j.ympev.2013.08.004

Talla V, Suh A, Kalsoom F, Dincă V, Vila R, Friberg M, Wiklund C, Backström N (2017) Rapid increase in genome size as a consequence of transposable element hyperactivity in wood-white (*Leptidea*) butterflies. Genome Biol Evol 9:2491–2505. https://doi.org/10.1093/gbe/evx163

Traut W, Marec F (1996) Sex chromatin in Lepidoptera. Q Rev Biol 71:239–256. https://doi.org/10.1086/419371

Traut W, Sahara K, Marec F (2007) Sex chromosomes and sex determination in Lepidoptera. Sex Dev 1:332–346. https://doi.org/10.1159/000111765

Traut W, Sahara K, Otto TD, Marec F (1999) Molecular differentiation of sex chromosomes probed by comparative genomic hybridization. Chromosoma 108:173–180. https://doi.org/10.1007/s004120050366

Traut W, Vogel H, Glöckner G, Hartmann E, Heckel DG (2013) High-throughput sequencing of a single chromosome: a moth W chromosome. Chromosome Res 21:491–505. https://doi.org/10.1007/s10577-013-9376-6

Van Nieukerken EJ, Kaila L, Kitching IJ et al (2011) Order Lepidoptera Linnaeus, 1758. In: Zhang ZQ (ed) Animal Biodiversity: An Outline of Higher-Level Classification and Survey Taxonomic Richness. Magnolia Press, Auckland. https://doi.org/10.11646/ZOOTAXA.3148.1.41

Věchtová P, Dalíková M, Sýkorová M, Žurovcová M, Füssy Z, Zrzavá M (2016) CpSAT-1, a transcribed satellite sequence from the coding moth, *Cydia pomonella*. Genetica 144:385–395. https://doi.org/10.1007/s10709-016-9907-0

Virkki N (1963) Gametogenesis in the sugarcane borer moth, *Diatraea saccharalis* (f.) (Crambidae). J Agric Univ Puerto Rico 47:102–137. https://doi.org/10.46429/jaupr.v47i2.12944

Vítková M, Fuková I, Kubíčková S, Marec F (2007) Molecular divergence of the W chromosomes in pyralid moths (Lepidoptera). Chromosome Res 15:917–930. https://doi.org/10.1007/s10577-007-1173-7

Wrensch DL, Kethley JB, Norton RA (1994) Cytogenetics of holokinetic chromosomes and inverted meiosis: keys to the evolutionary success of mites, with generalizations on eukaryotes. In: Houck MA (ed) Mites: Ecological and Evolutionary Analyses of Life-History Patterns. Springer, Dordrecht, pp 282–343. https://doi.org/10.1007/978-1-4615-2389-5_11

Yasukochi Y, Ashakumary LA, Baba K, Yoshido A, Sahara K (2006) A second-generation integrated map of the silkworm reveals synteny and conserved gene order between lepidopteran insects. Genetics 173:1319–1328. https://doi.org/10.1534/genetics.106.055541

Yasukochi Y, Ohno M, Shibata F, Jouraku A, Nakano R, Ishikawa Y, Sahara K (2016) A FISH-based chromosome map fot the European corn borer yields insights into ancient chromosomal fusions in the silkworm. Heredity 116:75–83. https://doi.org/10.1038/hdy.2015.72

Yoshido A, Sahara K, Marec F, Matsuba Y (2011) Step-by-step evolution of neo-sex chromosome in geographical populations of wild silkmoths, *Samia cynthia* spp. Heredity 106:614–624. https://doi.org/10.1038/hdy.2010.94

Yoshido A, Šíchová J, Pospíšilová K, Nguyen P, Voleníková A, Šafář J, Provazník J, Vila R, Marec F (2020) Evolution of multiple

sex-chromosomes associated with dynamic genome reshuffling in *Leptidea* wood-white butterflies. Heredity 125:138–154. https://doi.org/10.1038/s41437-020-0325-9

Yoshido A, Yamada Y, Sahara K (2006) The W chromosomes detection in several lepidopteran species by Genomic *in situ* hybridization (GISH). J Insect Biotec Sericol 75:147–151. https://doi.org/10.11416/jibs.75.147

Zedek F, Bureš P (2018) Holocentric chromosomes: from tolerance to fragmentation to colonization of the land. Ann Bot 121:9–16. https://doi.org/10.1093/aob/mcx118

Zhang J, Yu C, Krishnaswamy L, Peterson T (2011) Transposable elements as catalysts for chromosome rearrangements. Methods Mol Biol 701:315–326. https://doi.org/10.1007/978-1-61737-957-4_18

Zrzavá M, Hladová I, Dalíková M, Šíchová J, Õunap E, Kubíčková S, Marec F (2018) Sex chromosomes of the iconic moth *Abraxas grossulariata* (Lepidoptera, Geometridae) and its congener *A. sylvata*. Genes 9:279. https://doi.org/10.3390/genes9060279