Out of patterns, the euchromatic B chromosome of the grasshopper *Abracris flavolineata* is not enriched in high-copy repeats

Diogo Milani[1], Francisco J Ruiz-Ruano[2,3], Juan Pedro M Camacho[4], Diogo C Cabral-de-Mello[1,*]

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

[2]Uppsala University, Evolutionary Biology Centre, Department of Organismal Biology – Systematic Biology, Uppsala, Sweden

[3]University of East Anglia, Norwich Research Park, School of Biological Sciences, Norwich, United Kingdom

[4]UGR - Univ de Granada, Facultad de Ciencias, Departamento de Genética, Granada, Spain

**Corresponding author:* UNESP - Univ Estadual Paulista, Instituto de Biociências/IB, Departamento de Biologia Geral e Aplicada, 13506-900 Rio Claro, SP, Brazil Phone/Fax: 55 19 35264152. E-mail: cabral.mello@unesp.br

Abstract

In addition to the normal set of standard (A) chromosomes, some eukaryote species harbor supernumerary (B) chromosomes. In most cases, B chromosomes show differential condensation with respect to A chromosomes and display dark C-bands of heterochromatin, and some of them are highly enriched in repetitive DNA. Here we perform a comprehensive NGS (Next Generation Sequencing) analysis of the repeatome in the grasshopper *Abracris flavolineata* aimed at uncovering the molecular composition and origin of its B chromosome. Our results have revealed that this B chromosome shows a DNA repeat content highly similar to the DNA repeat content observed for euchromatic (non-C-banded) regions of A chromosomes. Moreover, this B chromosome shows little enrichment for high-copy repeats, with only a few elements showing overabundance in B-carrying individuals compared to the 0B individuals. Consequently, the few satellite DNAs (satDNAs) mapping on the B chromosome were mostly restricted to its centromeric and telomeric regions, and they displayed much smaller bands than those observed on the A chromosomes. Our data support the intraspecific origin of the B chromosome from the longest autosome by misdivision, isochromosome formation and additional restructuring, with accumulation of specific repeats in one or both B chromosome arms, yielding a submetacentric B. Finally, the absence of B-specific satDNAs, which are frequent in other species, along with its euchromatic nature, suggest that this B chromosome arose recently and might still be starting a heterochromatinization process. On this basis, it could be a good model to investigate the initial steps of B chromosome evolution.

Key words: FISH, genome, RepeatExplorer, repetitive DNAs, supernumerary chromosome

Introduction

B chromosomes are dispensable genomic elements reported in many plant, animal, and fungal species (Jones and Rees 1982; Camacho 2005; Houben et al. 2014; Jones 2017). B chromosomes were discovered more than a century ago (Wilson 1907) and, for many years, only repetitive DNA had been found on them (for review, see Camacho 2005). However, it is now known that B chromosomes also contain protein-coding genes (Martis et al. 2012; Valente et al. 2014; Navarro-Dominguez et al. 2017). A common characteristic of most B chromosomes is the accumulation of repetitive DNA, which accounts for its evolution and differentiation from A chromosomes (Camacho 2005; Houben et al. 2014). These accumulated repeats include microsatellites and satellite DNAs (satDNAs), multiple classes of transposable elements (TEs), and multigene families (see for example Nur et al. 1988; Ziegler et al. 2003; Coleman et al. 2009; Poletto et al. 2010; Peng and Cheng 2011; Bueno et al. 2013; Klemme et al. 2013; Milani et al. 2014; Silva et al. 2014; Coan and Martins 2018; Hanlon et al. 2018; Malinpensa et al. 2018; Marques et al. 2018; Ruiz-Ruano et al. 2018; Felicettia et al. 2021; Stornioli et al. 2021).

In grasshoppers, cytological and molecular analysis in multiple species revealed that most B chromosomes show C-banded heterochromatin and plenty of DNA repeats (for instance, see Ruiz-Ruano et al. 2016a, 2018; Milani et al. 2017a, 2018). For example, most B chromosome variants found in *Eyprepocnemis plorans* are mostly made of rDNA and a satDNA family (Cabrero et al. 1999; 2014; López-León et al. 2008) and they are enriched in R2 retrotransposons (Montiel et al. 2014). The most complete quantification of the repeatome in a grasshopper was recently performed in *Locusta migratoria* and showed that the B chromosome contains 94.9% of repetitive DNA, with a single satDNA comprising 55% of the B chromosome (Ruiz-Ruano et al.

2018). In addition, this B chromosome showed a 17 kb region, including 29 different TEs, which was apparent as a FISH band on the B chromosome. Similarly, heterochromatic B chromosomes rich in a variety of repetitive DNAs have been reported in other grasshopper species, such as, *Eumigus monticola*, *Rhammatocerus brasiliensis*, *Xyleus* (*discoideus*) *angulatus*, *Schistocerca rubiginosa Podisma sapporensis* and *Dichroplus pratensis* (Bidau et al. 2004; Loreto et al. 2008; Oliveira et al. 2011; Ruiz-Ruano et al. 2016a; Jetybayev et al. 2018; Milani et al. 2018).

An exception to this general pattern is the South American grasshopper *Abracris flavolineata* (2n= 22+X0♂/XX♀) where a submetacentric B chromosome failed to show heterochromatin defined by the C-banding technique, i.e. C-positive blocks (Cella and Ferreira 1991; Bueno et al. 2013). This B chromosome is mitotically stable thus showing the same number in all cells from the same individual and occurring in one or two copies in a natural population sampled at Rio Claro, São Paulo, Brazil (Milani et al. 2017b). Current evidence supports the origin of this B chromosome from the longest chromosome pair (L1), based on the U2 snDNA being only visualized by FISH on these two chromosomes (Bueno et al. 2013). Additionally, we detected other repeats on this B chromosome, a satDNA family (Milani et al. 2017a), some microsatellite repeats (Milani et al. 2014) and two TEs (Palacios-Gimenez et al. 2014), which were shared with many A chromosomes. Intrigued by the absence of C-positive heterochromatin on the B chromosome of *A. flavolineata*, we decided to perform high-throughput complementary bioinformatic and cytogenetic analyses to characterize its repetitive DNA content. Repeatome analysis including 1,744 TEs, 53 satDNAs and 9 multigene families revealed that, consistent with its C-heterochromatin scarcity, this B chromosome is not enriched in high-copy repetitive DNAs, which makes it unusual among B chromosomes in general. Exceptionally, we found a few repetitive DNAs

present on the euchromatic (C-negative) regions of the A chromosomes, which also decorate the interstitial regions of the B chromosome. In contrast, other repetitive DNAs that were enriched in heterochromatic regions (C-bands) of the A complement were mostly restricted to centromeric and distal regions of the B chromosome. In addition, satellitome analysis revealed that the B chromosome shared one satDNA family in exclusivity with the L1 autosome thus supporting B ancestry from this A chromosome. We finally suggest that this B chromosome could be a young element currently being in an initial step of heterochromatinization.

Materials and methods

Biological materials, genomic DNA extraction and chromosome preparations

For molecular and bioinformatic analysis, we used the same seven male individuals of *Abracris flavolineata* (three 0B, two 1B and two 2B) previously studied by Bueno et al. (2013), Milani et al. (2017b), and Ahmad et al. (2020). The hind legs of these animals, previously stored in 100% ethanol at -80 °C, were used for genomic DNA (gDNA) extraction following the phenol/chloroform-based protocol (Sambrook and Russel 2001), which was used for genomic sequencing and PCR assays (see next topics).

For chromosomal mapping we collected five gravid females, which were maintained alive in cages at the laboratory until oviposition, allowing embryos to be obtained. Mitotic embryo chromosome spreads were prepared according to the protocol proposed by Webb et al. (1978). Chromosome spreads were performed by maceration and spreading of portions of embryos on a slide within a drop of 50% acetic acid, under a hot plate at 45 ^{o}C.

Genome sequencing and identification of repetitive DNA sequences being overabundant in B-carrying individuals

Genomic DNA sequencing was performed by the Illumina HiSeq 4000 platform using the Macrogen Inc. service (Seoul, Republic of Korea). Sequencing yielded 27–41 Gb DNA (per sample) of 151 bp paired reads. The genomes from seven individuals are deposited in the Sequence Read Archive (SRA) under the accession numbers SRX7784770-SRX7784772. Repetitive sequences making up the repeatome of *A. flavolineata* were recovered and characterized using different approaches, including a thorough search for the satDNA families making up the satellitome, multigene families, and TEs (see details below).

To find and characterize the maximum number of different satDNA families, we applied the satMiner protocol (Ruiz-Ruano et al. 2016b). For this purpose, we randomly selected 2 x 5,000,000 reads from each individual using SeqTK (https://github.com/lh3/seqtk) and pooled those belonging to the same type of genome (0B, 1B or 2B) by concatenating them. We then performed sequence preprocessing for each group of reads using the "rexp_prepare_normaltag.py" script (https://github.com/fjruizruano/ngs-protocols), which uses Trimmomatic (Bolger et al. 2014) to remove adapters and low-quality nucleotides (Q<20), and finally selected only completely paired reads after trimming, i.e., those read pairs with 151 bp in both members. The script then interleaves forward and reverse reads and converts them to fasta format. We obtained 100,000 read pairs for each of the three libraries (0B, 1B and 2B) and concatenated them into a single file. We then applied the satMiner protocol (Ruiz-Ruano et al. 2016b) consisting of several rounds of clustering with RepeatExplorer (RE) software (Novák et al. 2013) alternated with the DeconSeq filtering tool (Schmieder and Edwards 2011) to remove those satDNA sequences

identified in previous RE rounds and added 100,000 of these cleaned read pairs from each pool sample (0B, 1B and 2B) prior to each new RE round (again summing up 300,000 read pairs).

RE clusters putatively containing satDNAs were selected by visual graph inspection to identify those showing spherical or ring shapes, which are characteristic of this type of DNA sequence. Then, we performed manual curation of the selected contigs by Geneious v4.8 software (Drummond et al. 2009), checked their tandem structure by dotplot graphic inspection and recovered the consensus sequence for repeat units of each satDNA family or subfamily. To search for homology between different satDNA families we first compared their consensus sequences using multiple sequence alignments with Muscle (Edgar 2004) implemented in Geneious v4.8 software (Drummond et al. 2009), and second, we ran a homology test based on RepeatMasker (Smit et al. 2017) with "rm_homology.py" (https://github.com/fjruizruano/ngs-protocols). The results of these analyses were used to classify the satDNA collection into superfamilies, families or subfamilies according to the identity criterion proposed in Ruiz-Ruano et al. (2016b).

For TE identification, we randomly selected 100,000 read pairs from each pool of genomes (0B, 1B and 2B), for a total of 600,000 reads, which were used as input for a single RE round followed by a reclustering-specific tool available in the Galaxy platform (https://repeatexplorer-elixir.cerit-sc.cz/galaxy/). This tool was used for merging clusters showing homology into larger contigs, which are prone to improve TE assembly. Then, we analyzed all the cluster contigs for sequence extraction with Geneious v4.8 software (Drummond et al. 2009). Since this method allowed the recovery of fewer than 200 different TE families, we also used the dnaPipeTE pipeline (Goubert et al. 2015), which uses Trinity (Grabherr et al. 2011) as an assembler,

followed by recurrent TE annotation and quantification in the raw reads compared with a custom database previously built by Ruiz-Ruano et al. (2018) from B-carrying genomes of *Locusta migratoria*. This analysis was performed using only forward reads and default parameters recommended for dnaPipeTE. Next, by means of a custom script (https://github.com/fjruizruano/ngs-protocols/blob/master/dnapipete_createdb.py) we used dnaPipeTE assembly and annotation to generate a fasta file with annotated contigs in the RepeatMasker format (Smit et al. 2017) for further analysis.

Finally, the multigene families (H3 histone gene, 18S, 28S, 5.8S and 5S rDNAs, U1, U2 and U6 snDNAs) and full mitochondrial DNA (mtDNA) were recovered using MITObim (Hahn et al. 2013) with the seed sequences used for *Locusta migratoria* in Ruiz-Ruano et al. (2018).

All the repeats obtained by these different methods were later concatenated, and redundancy was removed by CD-HIT-EST clustering (Li and Godzik 2006) using an 80% sequence identity level, implying that those repeats showing at least 80% identity were considered the same family.

Estimation of repetitive DNA sequence abundances and divergences in the* A. flavolineata *genome

Sequence abundance and divergence of each repetitive DNA family were determined in each of the seven genomes analyzed by means of RepeatMasker (Smit et al. 2017) using the Cross_match search engine on 5,000,000 read pairs from each library. SatDNA families were named in decreasing order of abundance in 0B genomes, following Ruiz-Ruano et al. (2016b). Sequence divergence was estimated by the Kimura 2-parameter (K2P) model using the calcDivergenceFromAlign.pl script within RepeatMasker software (Smit et al. 2017). Abundance for a given repetitive DNA

family was calculated as a genome proportion, represented by the sum of all mapped nucleotides belonging to it (including all subfamilies) with respect to the total number of nucleotides in the selected reads from each Illumina library. Abundance and divergence for each family were separately estimated for each individual and later averaged for 0B (three individuals), 1B (two individuals) and 2B (two individuals) genomes. We then calculated two sequence abundance quotients, 1B/0B and 2B/0B, to search for repeats being overabundant in the B-carrying genomes so that those repeats showing both quotients clearly higher than 1 and that 2B/0B was higher than 1B/0B were considered overabundant in B-carrying individuals and thus enriched in the B chromosome. However, those repeats showing quotients lower than 1 are considered less abundant (or absent) in the B chromosome rather than in the average A chromosome. All satDNA families and some TEs showing overabundance in B-carrying genomes were selected for subsequent chromosomal mapping (see below).

DNA amplification and chromosomal mapping of repetitive DNAs

We designed primers for PCR amplification either manually or else using Primer3 software (Untergasser et al. 2012) (Supplementary Table 1), and PCR conditions followed the same protocol described in Milani et al. (2018). For satDNA sequences, the monomeric bands were isolated and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corp., The Epigenetics Company, CA, USA) according to the manufacturer's recommendations. The same method was applied for TE isolation, taking care of isolating fragments showing the size expected from computational annealing of primers. These products were used for reamplification using the same PCR conditions. All amplified sequences were sequenced by the Sanger

method using Macrogen Inc. (Seoul, Republic of Korea) service to confirm the actual amplification of the target sequence.

We performed fluorescence *in situ* hybridization (FISH) on mitotic chromosome spreads from embryos using one or two probes simultaneously, according to Cabral-de-Mello and Marec (2021). Probes were labeled by digoxigenin-11-dUTP (Roche, Mannheim, Germany) or biotin-14-dATP (Invitrogen) and detected by antidigoxigenin-rhodamine (Roche) and streptavidin, Alexa Fluor 488-conjugated (Invitrogen), respectively. The chromosomes were counterstained using 4',6-diamidine-20-phenylindole dihydrochloride (DAPI) and slides were mounted in VECTASHIELD (Vector, Burlingame, CA, USA). The preparations were observed and images were captured using a BX61 Olympus microscope equipped with a fluorescence lamp and appropriate filters and a DP70 cooled digital camera. All images were processed and optimized using Adobe Photoshop CS6. According to the results observed, we classified the satDNA families into three types: i) visible FISH bands covering the whole chromosome width (B-pattern), ii) occurrence of dot-like scattered signals across the chromosome (D-pattern,), and iii) no FISH signal at all (NS-pattern).

Statistical methods

We compared repeat abundance between the 0B, 1B and 2B genomic libraries by means of nonparametric Friedman ANOVA and the Wilcoxon matched pairs test.

Results

Comparative genomic abundance reveals little enrichment for high-copy repeats in the B chromosome

The overall mean repetitive DNA abundance in *A. flavolineata* genomes from the Rio Claro, São Paulo, Brazil population was 52.94% in 0B individuals, 52.59% in 1B individuals and 52.00% in 2B individuals; this figure thus decreased with an increasing number of B chromosomes (Friedman ANOVA: χ^2= 8.08, N= 1806, df= 2, P<0.018). This result suggests that this B chromosome shows lower repetitive DNA content than the A chromosomes, on average, so that when a given repetitive element is scarce in the B chromosome, its genomic proportion will decrease as the number of Bs grows. This "dilution effect" was significant for TEs (χ^2= 10.12, N= 1744, df= 2, P<0.0064), marginally significant for satDNA (χ^2= 4.57, N= 53, df= 2, P>0.10) and not significant for multigene families (χ^2= 1.56, N= 9, df= 2, P>0.45) (Figure 1). However, a few repeats showed the reverse pattern, i.e., their abundance increased with increasing numbers of B chromosomes. This pattern suggested the presence of these repeats in the B chromosome. For quantitative application of this criterion, we calculated the 1B/0B and 2B/0B quotients and selected those elements showing 1B/0B>1 and 2B/0B>1B/0B, as the two conditions, as a whole, allowed selection for repeats showing increasing abundance with B number.

We found 53 satDNA families in *A. flavolineata*, all of which were present in the three genome libraries analyzed (Supplementary Table 2), thus revealing the absence of B-specific satDNAs. The dilution effect was also apparent for satDNA, as its genomic content decreased in the presence of B chromosomes (4.52% in 0B, 4.03% in 1B and 3.99% in 2B) (see Supplementary Table 2 and Figure 1c) (Wilcoxon matched pairs test: 0B vs. 1B: z= 3.27, P= 0.001; 0B vs. 2B: z= 2.19, P= 0.028). However, we found no significant difference in satDNA content between the 1B and 2B libraries (z= 0.27, P= 0.79), perhaps due to some degree of B chromosome heterogeneity. Consistent with the general dilution effect for satDNA, abundance comparisons between libraries

revealed that only two satDNA families (AflSat52-23 and AflSat53-17) were overabundant in the B-carrying genomes (see Table S2 and Figure 1c).

The analysis of coding tandem repeats (including rRNA, U snRNA and H3 histone multigene families) revealed that only the U2 snRNA family showed overabundance in the B-carrying genomes (Figure 1d). The presence of U2 snDNA on the *A. flavolineata* B chromosome was previously shown by FISH analyses (Bueno et al. 2013; Menezes-de-Carvalho et al. 2015; Milani et al. 2017b). The remaining gene families and mtDNA failed to show differences in relative abundance between B-carrying and B-lacking genomes, but some of them displayed the dilution effect (Figure 1d).

In the case of TEs, we found 212 elements (out of the 1,744 analyzed) meeting the 1B/0B>1 and 2B/0B>1B/0B criteria. These elements belonged to 28 families (Supplementary Table 3), the most abundant being LTR/gypsy elements (Figure 2). To test whether these results actually reflect overabundance in the B chromosome, we performed FISH for one element belonging to three distinct superfamilies, LTR/Gypsy (Gypsy_17), DNA/Tc1 (Tc1_74) and LINE/Jockey (Jockey_72). This analysis revealed their concentration on certain B chromosome regions with the appearance of chromosome bands (Figure 2). As these three families were among the seven most abundant, additional FISH work would reveal whether the observed pattern critically depends on abundance, a highly feasible possibility (see also Supplementary Figure 1).

High-throughput analysis of the satellitome reveals that satDNA is scarce on the B chromosome

One of the 53 satDNA families found (named here as AflSat02-391) had previously been described as AflaSAT-1 (Milani et al. 2017a). The repeat unit length

(RUL) of the 53 families ranged from 7 to 832 bp (mean= 224, SD= 167.6), and the total A+T content ranged from 30.43% to 76.50% (mean= 57.1%, SD= 8%). Homology tests between all satDNA families revealed the occurrence of only two superfamilies (SFs), with AflSat15-299, AflSat16-298 and AflSat26-296 comprising SF1, and AflSat20-233 and AflSat28-247 constituting SF2. As expected, the families belonging to each SF showed highly similar sequence properties (RUL and A+T content) (Supplementary Table 2).

A subtractive landscape (2B/0B) revealed a clear dilution effect for satDNA abundance, as the 2B genome showed a high deficit for most satDNA families, especially for the most abundant ones (Figure 3). To analyze whether these genomic results are reflected at the cytogenetic level, we performed the physical mapping by FISH on A and B chromosomes of *A. flavolineata* for all 53 satDNA families identified by bioinformatic analysis. Similar to other grasshopper species (for instance, see Ruiz-Ruano et al. 2016a, 2018), we observed three different patterns, with 44 families showing bands on chromosomes (B-pattern), three families showing many small dots scattered on chromosomes (D-pattern) and the six remainder showing no FISH signals (NS-pattern) (Table 1 and Supplementary Figure 2).

A summary of chromosome locations for the 53 satDNA families (Table 1) indicated that 47% of the 205 FISH bands found on A chromosomes were located on pericentromeric regions involving the centromere and the short chromosomal arm. The location of these satDNAs thus coincided with the heterochromatin location in this species, as revealed by C-banding (Bueno et al. 2013). However, the other half of the satDNA bands were found on euchromatic regions at proximal (5%), interstitial (30%) or distal (18%) locations of the long A chromosome arms (Table 1 and Supplementary Figure 2a-x). Notwithstanding, it is clear that the pericentric heterochromatic regions

were enriched in satDNA as they contained the five most abundant families representing 81% of all satDNA content in the 0B genome (Supplementary Table 2) (i.e., 3.67% out of the total 4.52%) (Figure 4, Supplementary Figure 2a,b, Table 1). Remarkably, of these five satDNAs, only the satDNA showing the highest abundance (AflSat01-179) was present on all A chromosomes (Figure 4a, Table 1), thus most likely playing a centromeric function (Melters et al. 2013). However, the least abundant satDNA families tended to show FISH bands on a single chromosome pair (Figure 4b, Supplementary Figure 2), as 15 of the 20 families with this condition showed abundance under the median value of all 53 families, and only 5 showed abundance above the median (Table 1). X was the A chromosome showing more satDNA FISH bands in exclusivity (three interstitially and three distally located), followed by S10 (4), M6 (3), L2 and M8 (2) and L3 and M7 (1). The X chromosome harbored the highest number of satDNA families (25) and it was the A chromosome showing the highest number of interstitial and distal satDNA bands (Table 1).

We noticed a clear-cut difference in chromosome location between the two superfamilies existing in the genome of *A. flavolineata*, as the three families belonging to SF1 always showed proximal locations on one (AflSat16-298 and AflSat26-296) or two (AflSat15-299) A chromosome pairs (Table 1, Supplementary Figure 2g,i,n), whereas the two SF2 family members showed either proximal (AflSat20-233) or interstitial (AflSat28-247) locations (Table 1, Supplementary Figure 2m,w).

Finally, there were nine other satDNA families where the location on A chromosomes was not in the form of FISH bands, three of which showed the dotted pattern (D) (Table 1, Supplementary Figure 3), and the six remaining showed no FISH signal at all (NS) (Table 1, Supplementary Figure 2y, 4).

Regarding the B chromosome, we observed that eight of the 44 satDNA families showing the B-pattern on A chromosomes (AflSat01-179, AflSat02-391, AflSat03-17, AflSat07-36, AflSat025-40, AflSat40-218, AflSat46-153, and AflSat52-23), were also present on the B chromosome, whereas the three families showing the D-pattern also showed multiple small dots on the B chromosome (Table 1 and Figure 4, Supplementary Figure 3). Among the 13 satDNA bands observed on the B chromosome, eight were pericentromeric, one was interstitial and four were distal. Most of the eight satDNA families showing FISH bands on the B chromosome showed multichromosomal locations on A chromosomes, except two showing locations on only one (AflSat46-153 on L1) or two (AflSat40-218 on S10 and X) A chromosomes (Table 1 and Figure 4b). Among all A chromosomes, L1 and X were the A chromosomes sharing the highest number of satDNA families with the B chromosome (seven each; see Table 1). Bearing in mind that L1 also shares the U2 snDNA in exclusivity with the B chromosome (Milani et al. 2017b), we consider that, with the available data (repetitive DNA only), L1 is the best candidate to be the ancestor of this B chromosome.

The satDNA families with dotted patterns occupied virtually the entire extension of the B chromosome, but AflSat08-184 and AflSat42-75 were less abundant on pericentromeric and terminal regions (Supplementary Figure 3a,c) whereas AflSat13-177 was less evident on the proximal region of the short arm (Supplementary Figure 3b). They also showed FISH signals on the euchromatic (non-C-banded) regions of the long arm of all A chromosomes, but they were absent in their C-banded regions located on the pericentromeric region and the short arm (Supplementary Figure 3).

A comparative analysis of abundance for the eight satDNA families displaying the B FISH pattern on the B chromosome revealed why the global abundance of

satDNA in the 0B, 1B and 2B genomes showed a dilution effect. For this purpose, we separately represented the most and the least abundant families (Figure 4), thus revealing that three families (AflSat01-179, AflSat07-36 and AflSat25-40) showed a clear decrease in abundance with an increasing number of B chromosomes, whereas only two (AflSat40-218 and AflSat52-23) showed the reverse pattern, due to B-enrichment, but these two satDNA families were among the least abundant in the genome.

Discussion

Genome low-pass sequencing combined with computational and chromosomal analysis provides a comprehensive understanding of the organization and evolution of DNA repeats on B chromosomes (Kumke et al. 2016; Ruiz-Ruano et al. 2018; Milani et al. 2018; Ebrahimzadegan et al. 2019; Serrano-Freitas et al. 2019). Through this approach, we found that the B chromosome of the grasshopper *A. flavolineata* is poorly enriched in repetitive DNA. Only three of the 53 satDNA families found in this species (AflSat40-218, AflSat52-23 and AflSat53-17), which are among the less abundant in the 0B genome, were overabundant in B-carrying genomes. Likewise, only 28 TE families, containing 212 elements, representing only 12% of the 1,744 TEs found, showed overabundance in B-carrying genomes. This scenario contrasts with the general idea that B chromosomes are enriched in repetitive DNA (Camacho 2005; Houben et al. 2014; Marques et al. 2018). Consistently, repeat-enriched B chromosomes have been reported in fish (Ziegler et al. 2003; Coan and Martins 2018; Stornioli et al. 2021), reptiles (Kichigin et al. 2019), plants (Martis et al. 2012; Kumke et al. 2016; Ebrahimzadegan et al. 2019), and insects (Hanlon et al. 2018; Ruiz-Ruano et al. 2018). Among the repeats found in B chromosomes, satDNA is the most frequent component

(McAllister 1995; Klemme et al. 2013; Hanlon et al. 2018; Ruiz-Ruano et al. 2018; Ebrahimzadegan et al. 2019; Langdon et al. 2000; Stornioli et al. 2021).

This accumulation of repetitive DNAs on B chromosomes is commonly assumed to be due to their genetic isolation from A chromosomes, with which they do not recombine (Camacho 2005; Houben et al. 2014). In this way, the nonenrichment in repetitive DNA, the absence of C-heterochromatin blocks and the absence of B-specific satDNA families would be consistent with the hypothesis that this B is a young element, resembling the composition of the A chromosome from which it derived (most likely L1, see below). The high similarity between B and A chromosomes is also supported also by B chromosome microdissection of *A. flavolineata* followed by chromosome painting, as all C-negative A chromosome regions and the B chromosome were similarly labeled (Menezes-de-Carvalho et al. 2015).

Based on FISH mapping of the U2 snDNA, the B chromosome in *A. flavolineata* was suggested to have derived from the L1 autosome, as only these two chromosomes harbor this sequence (Bueno et al. 2013). Here, chromosomal mapping of the full satellitome of this species has provided additional clues about B chromosome ancestry and evolution. We observed that the L1 autosome and the X chromosome both share the highest number of satDNA families with the B chromosome, i.e., seven families. However, the absence of U2 on the X chromosome and the fact that the L1 autosome is the only A chromosome sharing AflaSat46-153 with the B chromosome, reinforce the conclusion that L1 is the most likely B ancestor. Although our data support the possible derivation of the B chromosome from the L1 autosome, with possible subsequent restructuring of the B chromosome, additional research is necessary to obtain accurate information on the possible synteny of the repeats shared by these chromosomes, as it would help to unveil the precise origin of the B chromosome.

Additionally, some repeats present on the L1 autosome were not found on the B chromosome, indicating some additional degree of B differentiation attributed to the intense dynamism of repetitive DNAs. Among grasshoppers, the origin of B chromosomes from large A chromosomes, as the current results suggest in *A. flavolineata* appears to be uncommon, as the few cases where B chromosome ancestry was claimed involved medium (M) or small (S) A chromosomes, such as S11 in *E. plorans* (Teruel et al. 2014), S8 in *E. monticola* (Ruiz-Ruano et al. 2016a), M8 and S9 in *L. migratoria* (Ruiz-Ruano et al. 2018), S9 in *S. rubiginosa*, S11 in *R. brasiliensis* and S10 in *X. d. angulatus* (Milani et al. 2018). These medium- or small-sized A chromosomes are enriched in repetitive DNAs because their pericentromeric C-banded regions are the same size as the pericentromeric C-banded regions in long A chromosomes, but their non-C-banded regions are much smaller. Therefore, M and S chromosomes are more prone to be involved in chromosome rearrangements, which might be an initial step for B chromosome origin (Hewitt 1974; Perfectti and Werren 2001; Camacho 2005; Raskina et al. 2008; Houben et al. 2014; Ruiz-Ruano et al. 2016a; Milani et al. 2018). In *A. flavolineata*, the low amount of repeats on the B chromosome would be consistent with the low proportion of the C-banded region in L1 (see Figure 5) and the loss of most of the C-banded chromatin in the B. In contrast, B derivation from medium or small A chromosomes with a lower proportion of non-C-banded chromatin should most likely render heterochromatic Bs, likewise in cases with B chromosome ancestry related to highly heterochromatic chromosomes, such as sex chromosomes (Sharbel et al. 1998; Pansonato-Alves et al. 2014; Ventura et al. 2015; Serrano-Freitas et al. 2019).

Remarkably, satDNAs displaying FISH bands on the *A. flavolineata* B chromosome frequently showed a symmetrical pattern for the FISH bands located on

pericentromeric and distal regions, such as the U2 snDNA in both B chromosome arms (Figure 5, Table 1), suggesting the isochromosome nature of this B chromosome and the involvement of centromeric misdivision in its origin. The small size of the FISH bands observed on the B chromosome for most satDNA families (e.g. AflSat01-179, AflSat02-391 and AflSat03-17), would be consistent with the loss of the L1 short arm (which contains the largest amount of C-heterochromatin and satDNA families) during the B-forming misdivision. Isochromosomes arising from misdivision have been reported in grasshoppers such as *Eyprepocnemis plorans* (López-León et al. 1993), *Omocestus burri* (Del Cerro et al. 1994) and *Metaleptea brevicornis adspersa* (Grieco and Bidau 2000), plants such as *Zea mays* (Carlson and Phillips 1986), *Crepis capillaris* (Leach et al. 2005) and *S. cereale* (Marques et al. 2012), the fish *Astyanax scabripinnis* (Mestriner et al. 2000) and *Drosophila melanogaster* (Hanlon et al. 2018). Against the isochromosome hypothesis in *A. flavolineata* would be the B chromosome being not perfectly metacentric, as this would require additional events of inversion or differential duplications or deletions between B arms. More intense amplification of DNA repeats on one of the B chromosome arms has been noticed for TEs such as Gypsy_17, Tc1_74 and *Afmar2* (Palacios-Gimenez et al. 2014), and two satDNAs analyzed here (AflaSat07-36 and AflaSat40-218). This kind of event could have contributed to the emergence of the submetacentric B chromosome, which is currently prevalent in *A. flavolineata*. Notwithstanding, the evidence for L1 derivation of the B chromosome is still preliminary, as we all are still in the initial steps to disentangle the conundrum of B chromosome origin.

Altogether our results indicate that the B chromosome in *A. flavolineata* is unusually little enriched in repetitive DNAs, presumably because this B chromosome arose from the longest A chromosome, with a low proportion of C-heterochromatin, the

most part of which was lost during the misdivision that yielded the B chromosome from the L1 autosome. The B chromosome is enriched in only a few repetitive elements, to a low extent, and the absence of B-specific satDNAs suggests that this B chromosome is young. This fact might be helpful in testing the L1-derivation hypothesis of the B chromosome, as a putatively young element could still conserve high similarity in gene content with its ancestor chromosome.

Acknowledgements

We acknowledge the three anonymous reviewers for helpful comments that contributed significantly to improve the manuscript. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES), by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (process numbers 2014/11763-8 and 2015/16661-1), and Conselho Nacional de Desenvolvimento Científiico e Tecnológico (CNPq). 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). FJRR hold posdoctoral fellowships from Junta de Andalucía fellowship (Spain), Sven och Lilly Lawskis fond (Sweden) and a Marie Skłodowska-Curie Individual Fellowship (grant agreement 875732, European Commission).

Conflict of interest

The authors declare no conflict.

Data archiving

Genomes have been deposited at the Sequence Read Archive (SRA) under accession numbers SRX7784770-SRX7784772.

References

Ahmad SF, Jehangir M, Cardoso AL, Wolf IR, Margarido VP, Cabral-de-Mello DC, et al. (2020) B chromosomes of multiple species have intense evolutionary dynamics and accumulated genes related to important biological processes. BMC Genomics 21:656

Bidau CJ, Rosato M, Marti DA (2004) FISH detection of ribosomal cistrons and assortment-distortion for X and B chromosomes in *Dichroplus pratensis* (Acrididae). Cytogenet Genome Res 106:295–301

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120

Bueno D, Palacios-Gimenez OM, Cabral-de-Mello DC (2013) Chromosomal mapping of repetitive DNAs in the grasshopper *Abracris flavolineata* reveal possible ancestry of the B chromosome and H3 histone spreading. PLoS One 8:e66532

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

Cabrero J, López-León MD, Bakkali M, Camacho JPM (1999) Common origin of B chromosome variants in the grasshopper *Eyprepocnemis plorans*. Heredity 83: 435–439

Cabrero J, López-León MD, Ruíz-Estévez M, Gómez R, Petitpierre E, Rufas JS, et al. (2014) B1 was the ancestor B chromosome variant in the western Mediterranean area in the grasshopper *Eyprepocnemis plorans*. Cytogenet Genome Res 142:54–58

522 Camacho JPM (2005) B chromosomes. In: Gregory TR (ed) The evolution of the
523 genome. Elsevier San Diego 223–286
524 Carlson WR, Phillips RL (1986) The B chromosome of maize. Critical Reviews in Plant
525 Sciences 3:201–226
526 Cella MD, Ferreira A (1991) The cytogenetics of *Abracris flavolineata* (Orthoptera,
527 Caelifera, Ommatolampinae, Abracrini). Rev Brasileira Genética 14:315–329
528 Coan RL, Martins C (2018) Landscape of transposable elements focusing on the B
529 chromosome of the cichlid fish *Astatotilapia latifasciata*. Genes 9:269
530 Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, et
531 al. (2009) The genome of *Nectria haematococca*: contribution of supernumerary
532 chromosomes to gene expansion. PLoS Genet 5:e1000618
533 Del Cerro AD, Fernández A, Santos JL (1994) Spreading synaptonemal complexes of B
534 isochromosomes in the grasshopper *Omocestus burri*. Genome 37:1035–1040
535 Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M (2009) Geneious v.4.8.5,
536 biomatters ltd. Aukland, New Zealand
537 Ebrahimzadegan R, Houben A, Mirzaghaderi G (2019) Repetitive DNA landscape in
538 essential A and supernumerary B chromosomes of *Festuca pratensis* Huds. Sci
539 Rep 9:1–11
540 Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high
541 throughput. Nucleic Acids Res 32:1792–1797
542 Felicetti D, Haerter CA, Baumgärtner L, Paiz LM, Takagui FH, Margarido VP, et al.
543 (2021) A New Variant B Chromosome in Auchenipteridae: The Role of
544 (GATA)n and (TTAGGG)n Sequences in Understanding the Evolution of
545 Supernumeraries in *Trachelyopterus*. Cytogenet Genome Res 161:70-81

546 Goubert C, Modolo L, Vieira C, ValienteMoro C, Mavingui P, Boulesteix M (2015) De
547 novo assembly and annotation of the Asian tiger mosquito (*Aedes albopictus*)
548 repeatome with dnaPipeTE from raw genomic reads and comparative analysis
549 with the yellow fever mosquito (*Aedes aegypti*). Genome Biol Evol 7:1192–
550 1205
551 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. (2011)
552 Trinity: reconstructing a full-length transcriptome without a genome from RNA-
553 Seq data. Nat Biotechnol 29:644
554 Grieco ML, Bidau CJ (2000) The dicentric nature of the metacentric B chromosome of
555 *Metaleptea brevicornis adspersa* (Acridinae, acrididae). Heredity 84:639-646
556 Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial genomes
557 directly from genomic next-generation sequencing reads—a baiting and iterative
558 mapping approach. Nucleic Acids Res 41:e129–e129
559 Hanlon SL, Miller DE, Eche S, Hawley RS (2018) Origin, composition, and structure of
560 the supernumerary B chromosome of *Drosophila*
561 *melanogaster*. Genetics 210:1197–1212
562 Hewitt GM (1974) The integration of supernumerary chromosomes into the orthopteran
563 genome. In Cold Spring Harbor symposia on quantitative biology 38:183–194
564 Cold Spring Harbor Laboratory Press
565 Houben A, Banaei-Moghaddam AM, Klemme S, Timmis JN (2014) Evolution and
566 biology of supernumerary B chromosomes. Cell Mol Life Sci 71:467–478
567 Jetybayev IY, Bugrov AG, Dzuybenko VV, Rubtsov NB (2018) B chromosomes in
568 grasshoppers: Different origins and pathways to the modern Bs. Genes 9:509
569 Jones RN (2017) New species with B chromosomes discovered since 1980. Nucleus
570 60:263–281

Jones RN, Rees H (1982) B chromosomes. Academic press

Kichigin IG, Lisachov AP, Giovannotti M, Makunin AI, Kabilov MR, O'Brien PC, et al. (2019) First report on B chromosome content in a reptilian species: The case of *Anolis carolinensis*. Mol Genet Genomics 294:13-21

Klemme S, Banaei-Moghaddam AM, Macas J, Wicker T, Novák P, Houben A (2013) High-copy sequences reveal distinct evolution of the rye B chromosome. New Phytol 199:550–558

Kumke K, Macas J, Fuchs J, Altschmied L, Kour J, Dhar MK, et al. (2016) *Plantago lagopus* B chromosome is enriched in 5S rDNA-derived satellite DNA. Cytogenet Genome Res 148:68-73

Langdon T, Seago C, Jones RN, Ougham H, Thomas H, Forster JW, Jenkins G (2000) De novo evolution of satellite DNA on the rye B chromosome. Genetics 154:869–884

Leach CR, Houben A, Field B, Pistrick K, Demidov D, Timmis JN (2005) Molecular evidence for transcription of genes on a B chromosome in *Crepis capillaris*. Genetics 171:269–278

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659

López-León MD, Cabrero J, Dzyubenko VV, Bugrov AG, Karamysheva TV, Rubtsov, NB, Camacho JPM (2008) Differences in ribosomal DNA distribution on A and B chromosomes between eastern and western populations of the grasshopper *Eyprepocnemis plorans plorans*. Cytogenet Genome Res, 121: 260–265

López-León MD, Cabrero J, Pardo MC, Viseras E, Camacho JPM, Santos JL (1993) Generating high variability of B chromosomes in *Eyprepocnemis plorans* (grasshopper). Heredity 71:352–362

Loreto V, Cabrero J, López-León MD, Camacho JPM, Souza MJ (2008) Possible autosomal origin of macro B chromosomes in two grasshopper species. Chromosome Res 16:233–241

Malimpensa GC, Traldi JB, Toyama D, Henrique-Silva F, Vicari MR, Moreira-Filho O (2018) Chromosomal mapping of repeat DNA in *Bergiaria westermanni* (Pimelodidae, Siluriformes): Localization of 45S rDNA in B chromosomes. Cytogenet Genome Res 154:99-106

Marques A, Klemme S, Guerra M, Houben A (2012) Cytomolecular characterization of de novo formed rye B chromosome variants. Mol Cytogenet 5:34

Marques A, Klemme S, Houben A (2018) Evolution of plant B chromosome enriched sequences. Genes 9:515

Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, et al. (2012) Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. Proc Natl Acad Sci USA 109:13343–13346

McAllister BF (1995) Isolation and characterization of a retroelement from B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. Insect Mol Biol 4:253–262

Melters DP, Bradnam KR, Young HA, Telis N, May MR, Ruby JG, et al. (2013) Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. Genome Biol 14:R10

Menezes-de-Carvalho NZ, Palacios-Gimenez OM, Milani D, Cabral-de-Mello DC (2015) High similarity of U2 snDNA sequence between A and B chromosomes in the grasshopper *Abracris flavolineata*. Mol Genet Genomics 290:1787–1792

620 Mestriner CA, Galetti PM, Valentini SR, Ruiz IR, Abel LD, Moreira-Filho O, Camacho
621 JPM (2000) Structural and functional evidence that a B chromosome in the
622 characid fish *Astyanax scabripinnis* is an isochromosome. Heredity 85:1–9
623 Milani D, Bardella VB, Ferretti AB, Palacios-Gimenez OM, Melo ADS, Moura RC, et
624 al. (2018) Satellite DNAs unveil clues about the ancestry and composition of B
625 chromosomes in three grasshopper species. Genes 9:523
626 Milani D, Cabral-de-Mello DC (2014) Microsatellite organization in the grasshopper
627 *Abracris flavolineata* (Orthoptera: Acrididae) revealed by FISH mapping:
628 remarkable spreading in the A and B chromosomes. PLoS One 9:e97956
629 Milani D, Palacios-Gimenez OM, Cabral-de-Mello DC (2017b) The U2 snDNA is a
630 useful marker for B chromosome detection and frequency estimation in the
631 grasshopper *Abracris flavolineata*. Cytogenet Genome Res 151:36–40
632 Milani D, Ramos É, Loreto V, Martí DA, Cardoso AL, de Moraes KCM, et al. (2017a)
633 The satellite DNA AflaSAT-1 in the A and B chromosomes of the grasshopper
634 *Abracris flavolineata*. BMC Genet 18:81
635 Montiel EE, Cabrero J, Ruiz-Estévez M, Burke WD, Eickbush TH, Camacho JPM,
636 López-León MD (2014) Preferential occupancy of R2 retroelements on the B
637 chromosomes of the grasshopper *Eyprepocnemis plorans*. PLoS One 9:e91820
638 Navarro-Dominguez B, Ruiz-Ruano FJ, Cabrero J, Corral JM, LópezLeón MD, Sharbel
639 TF, Camacho JPM (2017) Protein-coding genes in B chromosomes of the
640 grasshopper *Eyprepocnemis plorans*. Sci Rep 7:45200
641 Novák P, Neumann P, Pech J, Steinhaisl J, Macas J (2013) RepeatExplorer: a galaxy-
642 based web server for genome-wide characterization of eukaryotic repetitive
643 elements from next-generation sequence reads. Bioinformatics 29:792–793

644 Nur U, Werren JH, Eickbush DG, Burke WD, Eickbush TH (1988) A" selfish" B
645 chromosome that enhances its transmission by eliminating the paternal
646 genome. Science 240:512–514
647 Oliveira NL, Cabral-de-Mello DC, Rocha MF, Loreto V, Martins C, Moura RC (2011)
648 Chromosomal mapping of rDNAs and H3 histone sequences in the grasshopper
649 *Rhammatocerus brasiliensis* (Acrididae, Gomphocerinae): extensive
650 chromosomal dispersion and co-localization of 5S rDNA/H3 histone clusters in
651 the A complement and B chromosome. Mol Cytogenet 4:24
652 Palacios-Gimenez OM, Bueno D, Cabral-De-Mello DC (2014) Chromosomal mapping
653 of two Mariner-like elements in the grasshopper *Abracris flavolineata*
654 (Orthoptera: Acrididae) reveals enrichment in euchromatin. Eur J Entomol
655 111:329–334
656 Pansonato-Alves JC, Serrano ÉA, Utsunomia R, Camacho JPM, da Costa Silva GJ,
657 Vicari MR, et al. (2014) Single origin of sex chromosomes and multiple origins
658 of B chromosomes in fish genus *Characidium*. PLoS One 9:e107169
659 Peng SF, Cheng YM (2011) Characterization of satellite CentC repeats from
660 heterochromatic regions on the long arm of maize B-chromosome. Chromosome
661 Res 19:183–191
662 Perfectti F, Werren JH (2001) The interspecific origin of B chromosomes: experimental
663 evidence. Evolution 55:1069–1073
664 Poletto AB, Ferreira IA, Martins C (2010) The B chromosomes of the African cichlid
665 fish *Haplochromis obliquidens* harbour 18S rRNA gene copies. BMC Genet
666 11:1

667 Raskina O, Barber JC, Nevo E, Belyayev A (2008) Repetitive DNA and chromosomal
668 rearrangements: speciation-related events in plant genomes. Cytogenet Genome
669 Res 120:351–357
670 Ruiz-Ruano FJ, Cabrero J, López-León MD, Camacho JPM (2016a) Satellite DNA
671 content illuminates the ancestry of a supernumerary (B)
672 chromosome. Chromosoma 126:487–500
673 Ruiz-Ruano FJ, Cabrero J, López-León MD, Sánchez A, Camacho JPM (2018)
674 Quantitative sequence characterization for repetitive DNA content in the
675 supernumerary chromosome of the migratory locust. Chromosoma 127:45–57
676 Ruiz-Ruano FJ, López-León MD, Cabrero J, Camacho JPM (2016b) High-throughput
677 analysis of the satellitome illuminates satellite DNA evolution. Sci Rep 6:28333
678 Sambrook J, Russell DW (2001) Molecular Cloning. Sambrook & Russel Vol. 1, 2,
679 3. Cold Springs Harbour Laboratory Press
680 Schmieder R., Edwards R (2011) Fast identification and removal of sequence
681 contamination from genomic and metagenomic datasets. PloS One 6:e17288
682 Serrano-Freitas ÉA, Silva DM, Ruiz-Ruano FJ, Utsunomia R, Araya-Jaime C, Oliveira
683 C, et al. (2019) Satellite DNA content of B chromosomes in the characid fish
684 *Characidium gomesi* supports their origin from sex chromosomes. Mol Genet
685 Genomics 295:195–207
686 Sharbel TF, Green DM, Houben A (1998) B-chromosome origin in the endemic New
687 Zealand frog *Leiopelma hochstetteri* through sex chromosome
688 devolution. Genome 41:14–22
689 Silva DMDA, Pansonato-Alves JC, Utsunomia R, Araya-Jaime C, Ruiz-Ruano FJ,
690 Daniel SN, et al. (2014) Delimiting the origin of a B chromosome by FISH

691 mapping, chromosome painting and DNA sequence analysis in *Astyanax*
692 *paranae* (Teleostei, Characiformes). PLoS One 9:e94896
693 Smit AFA, Hubley R, Green P (2017) RepeatMasker Open-4.0. http://
694 www.repeatmasker.org
695 Stornioli JHF, Goes CAG, Calegari RM, dos Santos RZ, Giglio LM, Foresti F, et al.
696 (2021) The B Chromosomes of Prochilodus lineatus (Teleostei, Characiformes)
697 Are Highly Enriched in Satellite DNAs. Cells 10:1527
698 Teruel M, Ruiz-Ruano FJ, Marchal JA, Sánchez A, Cabrero J, Camacho JPM, Perfectti
699 F (2014) Disparate molecular evolution of two types of repetitive DNAs in the
700 genome of the grasshopper *Eyprepocnemis plorans*. Heredity 112:531–542
701 Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG
702 (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115–
703 e115
704 Valente GT, Conte MA, Fantinatti BEA, Cabral-de-Mello DC, Carvalho RF, Vicari
705 MR, Kocher TD, Martins C (2014) Origin and evolution of B chromosomes in
706 the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses.
707 Mol Biol Evol 31:2061–2072
708 Ventura K, O'Brien PCM, do Nascimento Moreira C, Yonenaga-Yassuda Y, Ferguson-
709 Smith MA (2015) On the origin and evolution of the extant system of B
710 chromosomes in *Oryzomyini* radiation (Rodentia, Sigmodontinae). PloS
711 one 10:e0136663
712 Webb GC, White MJD, Contreras N, Cheney J (1978) Cytogenetics of the
713 parthenogenetic grasshopper *Warramaba* (formely *Moraba*) *virgo* and its
714 bisexual relatives. IV. Chromosome banding studies. Chromosoma 67:309–339

715 Wilson EB (1907) The supernumerary chromosomes of Hemiptera. Science
716 NY 26:870–871
717 Ziegler CG, Lamatsch DK, Steinlein C, Engel W, Schartl M, Schmid M (2003) The
718 giant B chromosome of the cyprinid fish *Alburnus alburnus* harbours a
719 retrotransposon-derived repetitive DNA sequence. Chromosome Res 11:23–35
720

721 **Tables**

722 **Table 1.** Chromosome location of the 53 satDNA families found in *Abracris*
723 *flavolineata*. SF = superfamily, L1-L3 = three long autosomes, M4-M8 = five medium
724 autosomes, S9-S11 = three short autosomes, X = X chromosome, B = B chromosome, p
725 = pericentromeric (centromere plus short chromosomal arms), pr = proximal (near to
726 centromere), i = interstitial, d = distal. B = banded pattern, D = dotted pattern and NS =
727 no signal. Repeats with dotted patterns occurred on all chromosomes, including the B
728 chromosome (see text for details). Asterisks (*) indicate the occurrence of
729 heteromorphisms for the presence/absence of bands in homologous chromosomes.
730

731 **Figures**

732 **Figure 1**. Relative abundance of several types of repetitive DNA in *A. flavolineata*
733 genomes carrying 1B and 2B, in comparison with the B-lacking genome, measured by
734 the log2 transformed 1B/0B and 2B/0B quotients. a) Five examples of TEs showing
735 overabundance in the B chromosome (solid lines), and five others showing the dilution
736 effect (dotted lines) with negative values for both quotients. b) Overabundant TEs in the
737 B chromosome, indicating TE type for the five showing the highest quotients. c) Only
738 two satDNA families were overabundant in the B chromosome. Note the dilution effect
739 for many other satDNAs. d) Only the U2 snDNA showed clear overabundance in the B-

carrying genomes, whereas the other families showed the dilution effect or quotients close to zero, suggesting their scarcity in the B chromosome.

Figure 2. Comparative genomic proportion between the 0B, 1B and 2B genomes for the most abundant TEs in *Abracris flavolineata*. The asterisks indicate the superfamilies in which one representative was selected for FISH mapping on chromosomes (a-c). Note the spread distribution on long arms and absence of signals on pericentromeric C-heterochromatic region of A chromosomes. In the B chromosome (arrowheads) observe the differential distribution of TEs, i.e., first interstitial half of long arm (a), spread signal along the entire extension of the B chromosome, except distal regions (b) and enrichment on interstitial areas of both arms, and faint signals in proximal half of long arm (c). This last repeat was also absent in the terminal regions. Bar = 10μm.

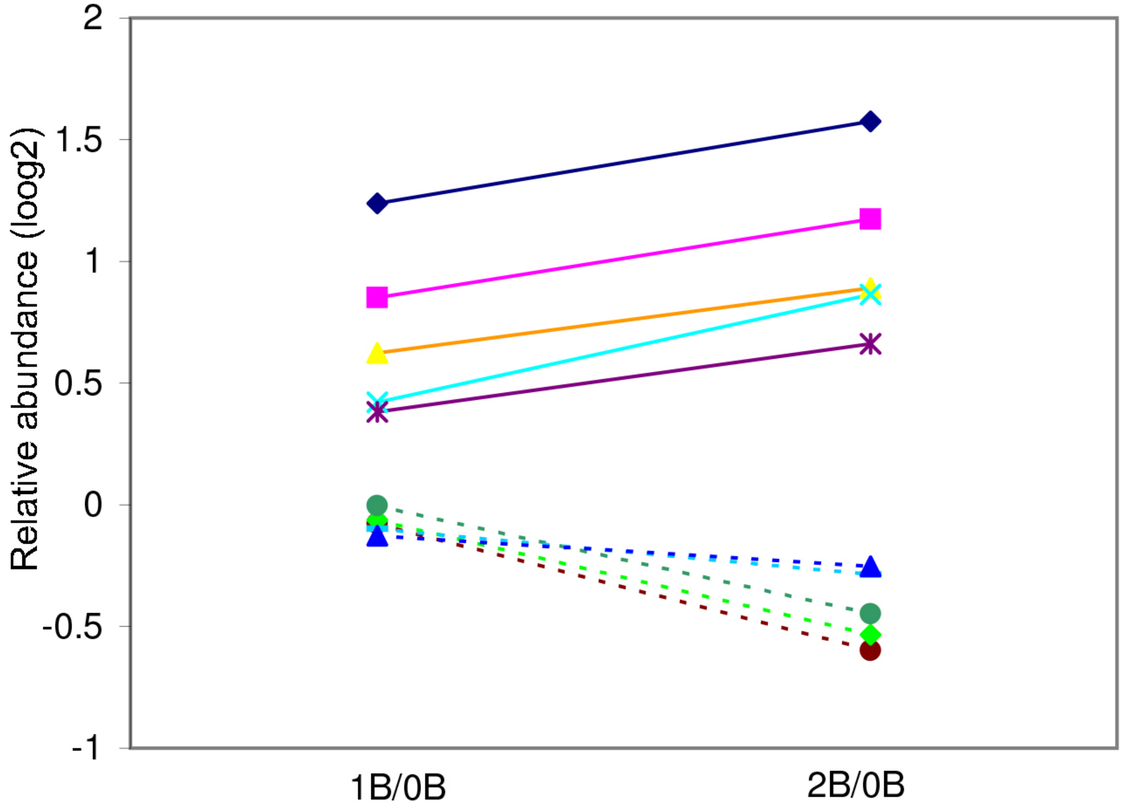
Figure 3. Subtractive repetitive landscape (genome proportion *versus* sequence divergence based on Kimura substitution level) obtained from average counts for satDNAs in two males with 2B chromosomes and three with no B chromosome of *Abracris flavolineata*. Abundance values show the difference between the 2B minus the 0B genomes. Thus, positive values indicate overabundance in the 2B genomes, and negative values indicate overabundance in the 0B genomes. Note the occurrence of mainly negative values indicating the low enrichment of satDNAs in 2B-carrying genomes.

Figure 4. Comparative genomic proportion and FISH mapping for eight satDNAs occurring in the B chromosomes with a banded pattern, showing high (a) or low (b) abundance (expressed as genome proportion). Repeat names are indicated on the left.

765 Some A chromosomes are indicated on each embryo mitotic metaphase plate, and the B
766 chromosome is indicated by arrowheads. The differential satDNA distribution on the B
767 chromosome was observed, with pericentromeric signals for AflSat01 and AflSat02,
768 pericentromeric plus distal signals for AflSat03, AflSat07, AflSat25, AflSat46, and
769 AflSat52, and pericentromeric plus interstitial (on the long arm) signals for AfSat40.
770 Additionally, note the exclusive presence of AflSat46 bands on the B chromosome and
771 the L1 pair. Bar = 10μm.
772
773 **Figure 5.** Comparative C-banding and FISH for repeat location between the L1 and B
774 chromosomes of *A. flavolineata*. The L1 autosome showed, like the remaining A
775 chromosomes, a large pericentromeric C-band including the pericentromeric region and
776 the short arm. The FISH analysis for seven satDNA families and the U2 snDNA repeat
777 showed pericentromeric and telomeric locations on the B chromosome whereas they
778 were located on the pericentromeric region and the short arm of L1 (AflSat01, AflSat02,
779 AflSat03, the pericentromeric region (AflSat46), interstitial (AflSat07), interstitial
780 region and the short arm (AflSat25 and AflSat52). satDNAs thus might suggest that B
781 originates from the proximal third of L1, including the interstitial region containing
782 several satDNAs. However, the U2 snDNA is located on an L1 region outside the
783 former proximal region, so the presence of U2 on B is not explained by a single
784 rearrangement event.
785

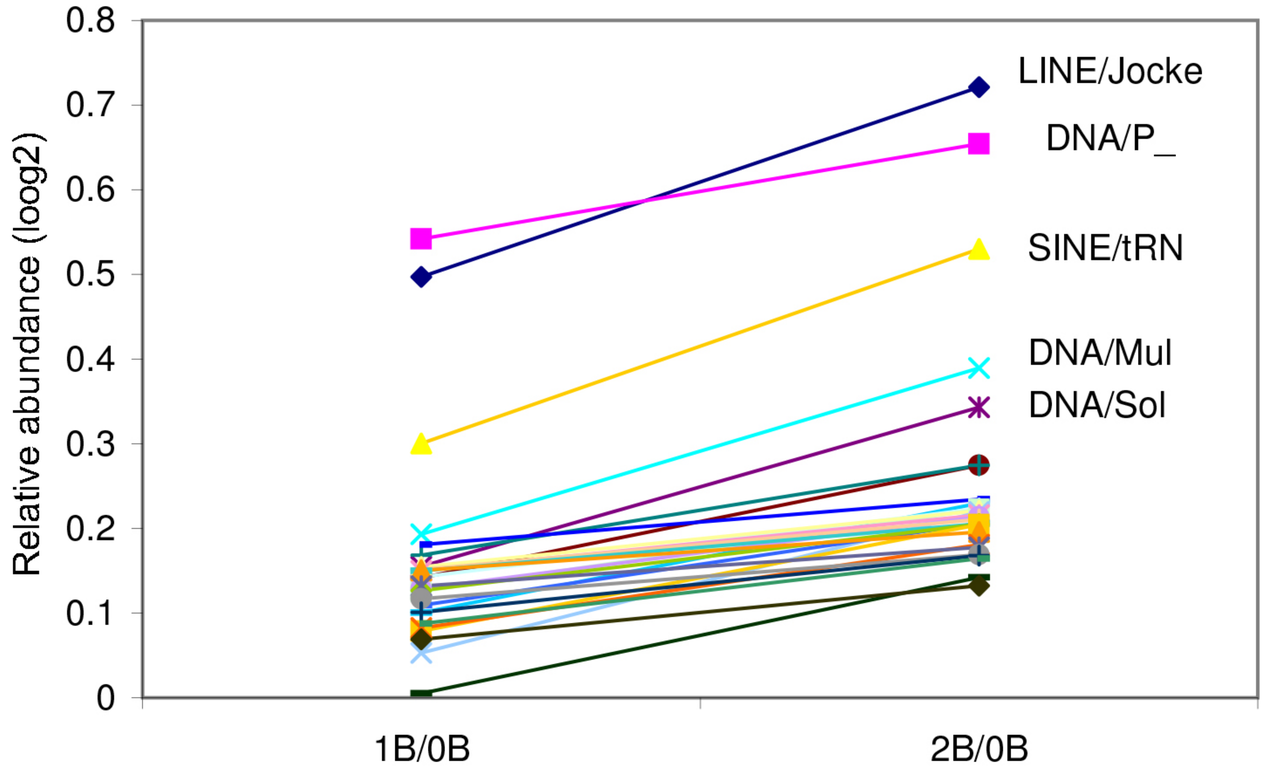
a

TEs examples



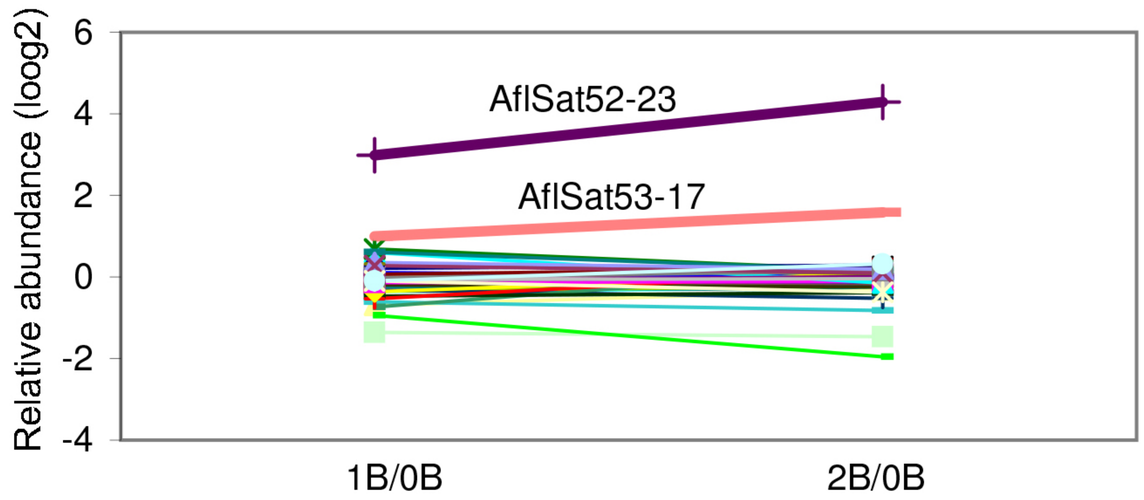
b

Overbundant TEs in the B chromosome



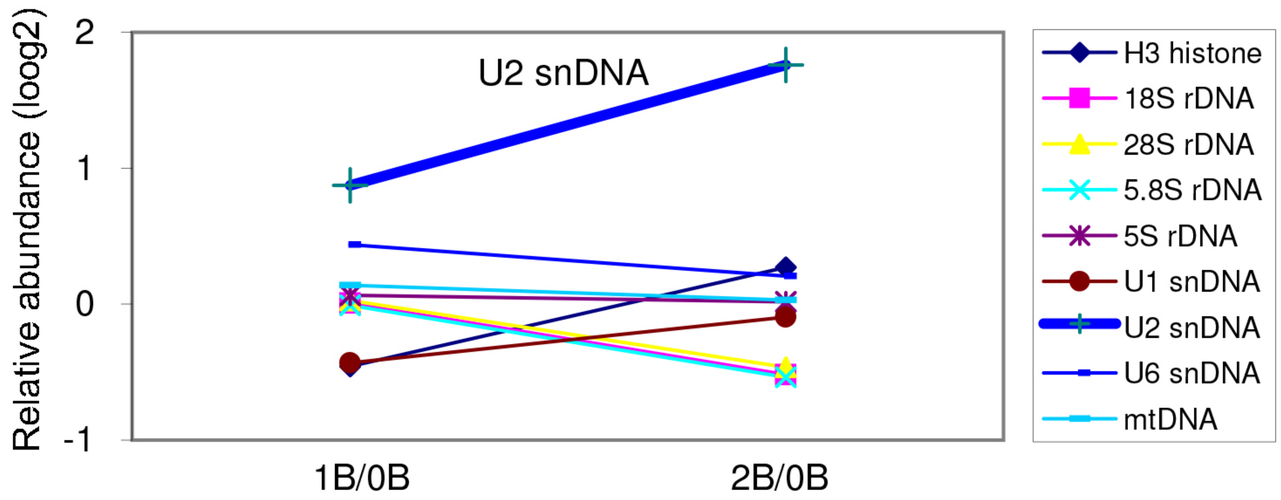
c

Satellite DNAs

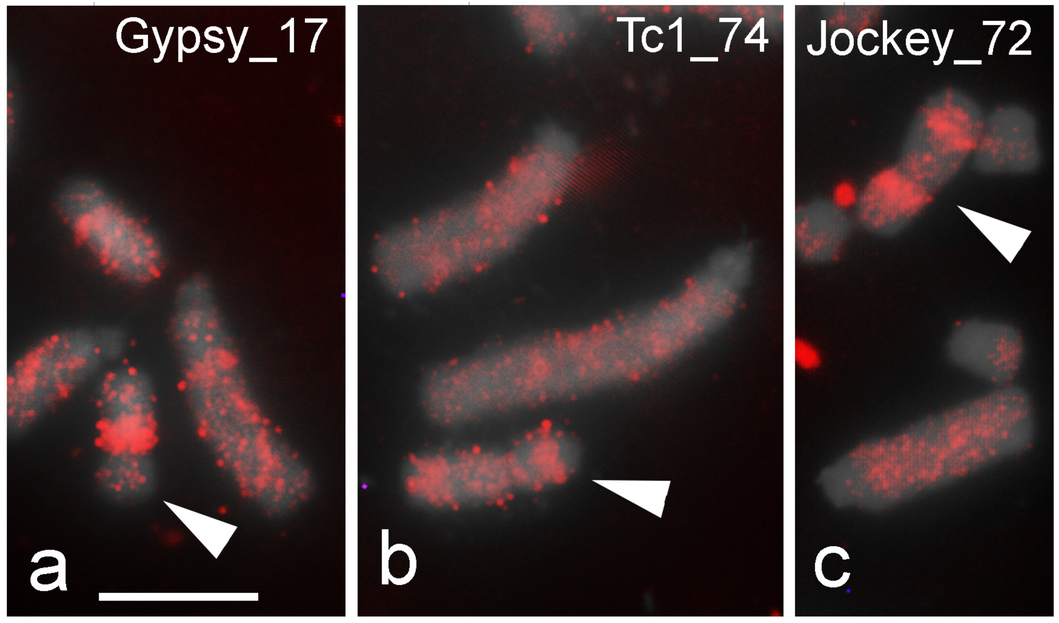


d

Gene families and mtDNA



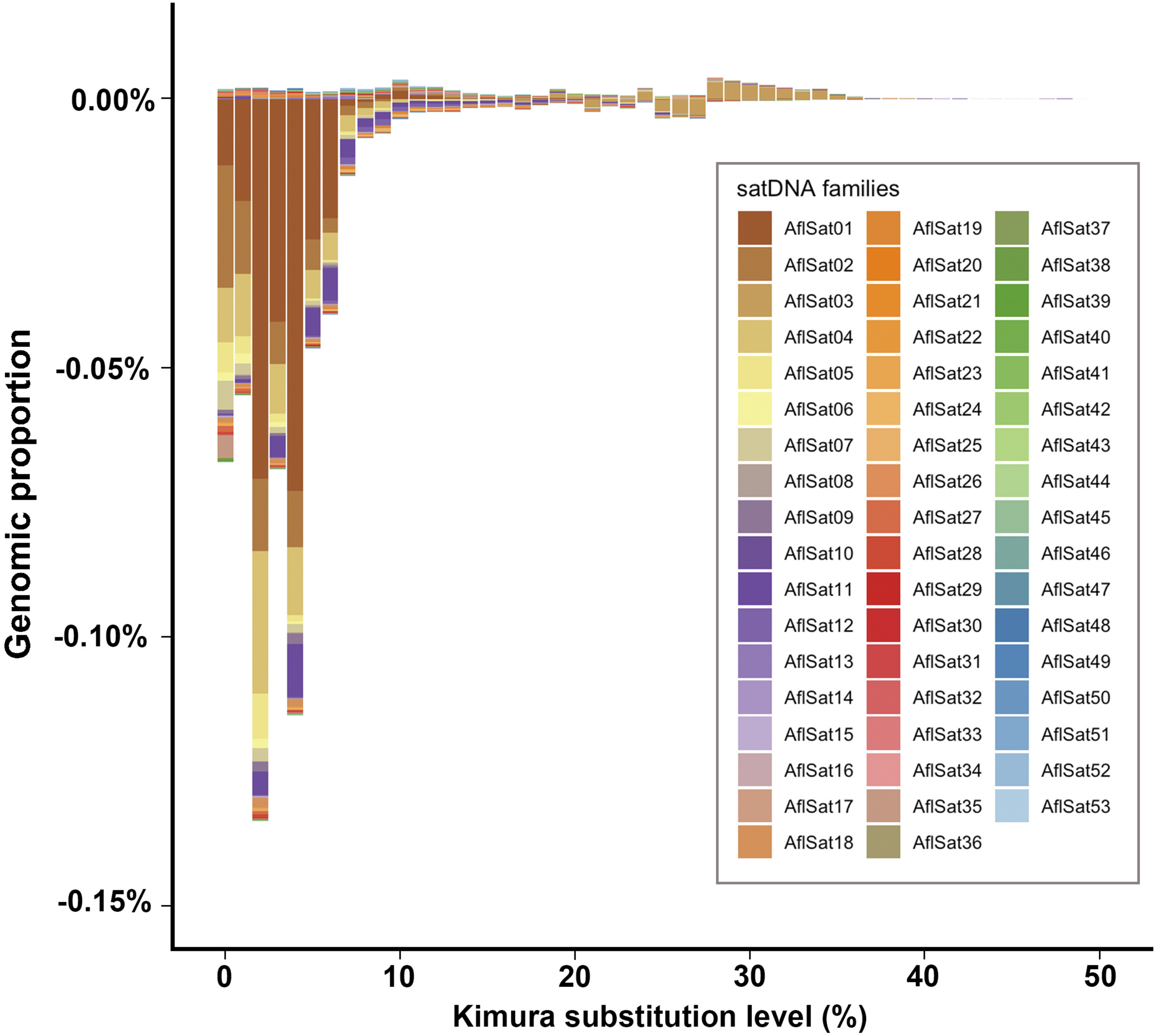
LTR/Gypsy*
DNA/Maverick
LINE/Penelope
LINE/L1
DNA/Tc1*
DNA/DNA
LINE/Jockey*
RC/Helitron
LINE/I
DNA/Mariner
DNA/hAT
DNA/Sola
LTR/Pao
LINE/R1



2B
1B
0B

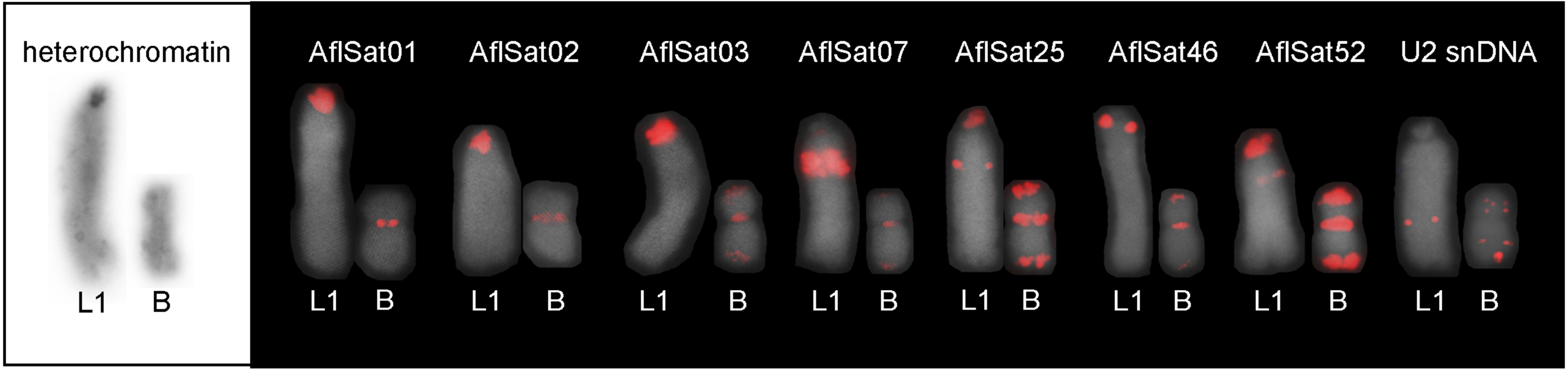
0.0%
0.5%
1.0%
1.5%
2.0%
2.5%

Genomic proportion


0.00%
-0.05%
-0.10%
-0.15%
Genomic proportion
0
10
20
30
40
50
Kimura substitution level (%)
satDNA families
AflSat01
AflSat02
AflSat03
AflSat04
AflSat05
AflSat06
AflSat07
AflSat08
AflSat09
AflSat10
AflSat11
AflSat12
AflSat13
AflSat14
AflSat15
AflSat16
AflSat17
AflSat18
AflSat19
AflSat20
AflSat21
AflSat22
AflSat23
AflSat24
AflSat25
AflSat26
AflSat27
AflSat28
AflSat29
AflSat30
AflSat31
AflSat32
AflSat33
AflSat34
AflSat35
AflSat36
AflSat37
AflSat38
AflSat39
AflSat40
AflSat41
AflSat42
AflSat43
AflSat44
AflSat45
AflSat46
AflSat47
AflSat48
AflSat49
AflSat50
AflSat51
AflSat52
AflSat53

AflSat01
AflSat02
AflSat03
AflSat07
AflaSat01
AflSat02
AflSat03
AflSat07
2B
1B
0B
0.0% 0.2% 0.4% 0.6% 0.8% 1.0% 1.2% 1.4% 1.6% 1.8% 2.0%
a
Genomic proportion

AflSat25
AflSat40
AflSat46
AflSat52
AflSat25
AflSat40
AflSat46
AflSat52
2B
1B
0B
0.000% 0.002% 0.004% 0.006% 0.008% 0.010% 0.012% 0 014% 0.016% 0.018% 0.020%
b
Genomic proportion


heterochromatin
L1
B
AflSat01
AflSat02
AflSat03
AflSat07
AflSat25
AflSat46
AflSat52
U2 snDNA
L1
B

SatDNA family	Pattern	Chromosome pair and location												
		L1	L2	L3	M4	M5	M6	M7	M8	S9	S10	S11	X	B
AflSat01-179	B	p	p	p	p	p	p	p	p	p	p	p	p	p
AflSat02-391	B	p	p	p	p	p	p	p	p	p	p		p	p
AflSat03-17	B	p	p	p	p	p	p	p	p	p		p	p	p
AflSat04-161	B		p*	p		p	p	p		p		p		
AflSat05-437	B	p	p	p	p	p	p	p	p	p		p	p	
AflSat06-265	B	d	d	d	d	d	d	d	d	d			d	
AflSat07-36	B	i	i	i	i	i	i+d	i+d	i	p			i	p+d
AflSat08-184	D													
AflSat09-187	B			d	d	d	d				d		d	
AflSat10-297	B	p	p	p	p	p	p	p	p	p	p		p	
AflSat11-407	NS													
AflSat12-237	B					pr	pr				p	p		
AflSat13-177	D													
AflSat14-110	B	i		i	i			d*					i	
AflSat15-299 / SF1	B	pr		pr										
AflSat16-298 / SF1	B						pr							
AflSat17-231	B		i											
AflSat18-7	B	d	d	d	d	d	d	d					i	
AflSat19-17	NS													
AflSat20-233 / SF2	B							pr	pr					
AflSat21-260	NS													
AflSat22-135	B	i	i	i	i	i		i		d		d	i	
AflSat23-286	B							p						
AflSat24-308	B								p					
AflSat25-40	B	p+i*	p+i	p+i	p+i	p+i	p+i	p+i	p+i+d	p	p		p+i	p+d
AflSat26-296 / SF1	B						pr							
AflSat27-609	B			i										
AflSat28-247 / SF2	B		i					i	i					
AflSat29-616	B										p			
AflSat30-197	B					d							d	
AflSat31-226	B						d							
AflSat32-429	B									d	p			
AflSat33-295	B						pr						i	
AflSat34-30	B		i*	i										
AflSat35-362	B										i			
AflSat36-41	NS													
AflSat37-134	B	i	i	i	i		i	i					pr	
AflSat38-377	B										p			
AflSat39-832	B		i											
AflSat40-218	B										d		i	p+i
AflSat41-186	B								p					
AflSat42-75	D													
AflSat43-131	B												i	
AflSat44-254	B										d			
AflSat45-414	NS													
AflSat46-153	B	p												p+d
AflSat47-75	B												d	

AflSat48-167	B												d	
AflSat49-138	B												d	
AflSat50-102	B												i	
AflSat51-30	B												i	
AflSat52-23	B	p+i	p+i	p+i	p+i	p+i	p+i	p+i	p+i	p+i	p+i	p+i	p+i	p+d
AflSat53-17	NS													
Total (p)		8	8	8	7	8	8	9	9	9	9	6	7	8
Total (pr)		1	0	1	0	1	4	1	1	0	0	0	1	0
Total (i)		6	9	8	6	4	4	6	4	1	2	1	11	1
Total (d)		2	2	3	3	4	4	4	2	3	3	1	6	4
Total		17	19	20	16	17	20	20	16	13	14	8	25	13
Shared with B		7	6	6	6	6	6	6	6	6	5	3	7	-

Supplementary Table 1. Primers used for amplification of repetitive DNAs mapped on chromosome spreads of *A. flavolineata* through FISH.

Sequence name	Primer forward	Primer reverse
AflSat01-179	CTCAGACCACTGGGCCATAT	AGACGCTAGTTATAATGAAAAACTGT
AflSat02-391	CCGTTTTAAACACTTCCATTACAGAGT	AGACGCTAGTTATAATGAAAAACTGT
AflSat03-17	TGCAAGAGTGCTGCATG	CATGCAGCACTCTTGCA
AflSat04-161	CTATCTCCAGACTACTAAATTCGAGT	AGTTTTAGCTATCTTTCGATGCA
AflSat05-437	TTGCCTTCTTTCACCTGCGA	TTTACGAGAAAGAAAACACATTGTT
AflSat06-265	GCCCGTTGCACTCTACCCACC	ACGCAGTATCACAAACAGCC
AflSat07-36	GGGAAAAAGGGTAATTTTTGGGAGT	CCCGAGTCATAGTTACTCCCA
AflSat08-184	AACGTGTGGTTTTGCATATACATGT	ACGACTGTTATGGGTAGCAT
AflSat09-187	CCGAATTGATAGGTTTTTAAGATGAT	CGGGCTCTTGCCAGTGATGGC
AflSat10-297	GGCAACAGTAAGTATTTCTACTATCT	GCCAAGAGGCTAGTTGAAATT
AflSat11-407	GTGAGGGCAGTTGACACTCT	CACTCCAGCAGACACTGGGG
AflSat12-237	ACCACTCCTTTCTAAACTCTTACA	AATGCTGGTGGTTTGAAAGT
AflSat13-177	TGTTAATGTATAACTGCATACTGTGT	ACACATTTTCTTTTCACTAGGCGT
AflSat14-110	CCAACTCGCAAGCAAACTGTTGCC	GGCGAAGCCAACGAGTGTGT
AflSat15-299	TTCCGTCAGGTGATGTCCTGTCT	CCCACATGTCCTCGATGGCGAG
AflSat16-298	TGTCCGACCACCCCATTCTACCA	CACGCCCCGATTCGAGGACA
AflSat17-231	AAAACAGTTGACTGTGGCGCT	CAGACCAGTGCAGTCAGTGCG
AflSat18-7	CCCAACACCCAACACCC	TGTTGGGTGTTGGGTGT
AflSat19-17	TATGCGTTTTTCGCATATATGCG	GCATATATGCGAAAAACGCATAT
AflSat20-233	CACTCAGAAAAACGGGCCTCTT	TGTCAGTTATTGGACCCCCTCCGT
AflSat21-260	AGCAATCGCATGTAAGTATAGACA	GCTGGTGCAGTATGCCCGACA
AflSat22-35	AGCTGAGATCCAACATTTCAG	CACCATGGAGCAAACATTCAT
AflSat23-286	CCACCCCCACCCCCTTATTTGC	TGGAACTGGGCCAGGACAGAGA
AflSat24-308	GCCATACCTCCACTGCTCCAGC	GCCTTGGGAGCGGACTAAAGCT
AflSat25-40	CCGCTATATTTTGTTTGTAAATTAGCA	CGGCACAAGGCCTCCATGCT
AflSat26-296	GGTGTTCCATCAGGACCTTTCT	TCCTTTGTGGAAAAATGCAGCAAT
AflSat27-609	TCCATCGTTAAATACAGAACGCT	TGAAGCAACATCCTATTAATCACGT
AflSat28-247	TACACATGTCGACAAACAGGTCTA	TATGGCCACTTGCAGTTGTC
AflSat29-616	AGAGGACTTATAGATGTAGGC	CATTCGTCTTGTCTTTGCCAA
AflSat30-197	CCTAGCTTCATCATCAGGCACT	GGTAGCTTTGAAATCGATTGTAAT
AflSat31-226	GCACAGATGGTAAGAAACATTAAA	TGCTGTGTTGAAGCTTCCTT
AflSat32-429	CAGATTATCTACATGGAACTTG	TGTTTTATGTAGTGGCAGGG
AflSat33-295	ACAGCACTGGTGGGTCATCA	ACAAACAGGACGTTTAAGAATTCA
AflSat34-30	TGCCGAGGCAGTTGGCTAC	CTTAAGGTAGCCAACTGCCTCG
AflSat35-362	CCGTATGCAGGTTAGTATTAC	GTTAACGCAGACACTGATTACAT
AflSat36-41	GCGGATGTCCAATACACTAAA	CGCCTAGTTGCGATGTGATA
AflSat37-134	AGCCGCAATATCCACATCGTCA	GCTGATGAGTCGCCTATGCCAC
AflSat38-377	GCGGGAGCAGCAACACAACAC	CGCCGGTTTGTCATGGCTCAGT
AflSat39-832	ATGGCACTCGCACAGCCTGT	GCGTGGGAAGAAGAGGTAGCTGG
AflSat40-218	ACAGTGCACAACAAACTGTAGTTGTCC	TAGACATGCTGTTGCCATGTC
AflSat41-186	GTCCGGCGCCCAAGTGATCG	TTAGTCGTCGTCAGCCGGCA
AflSat42-75	CCGATGCCGCAACCGCC	TGAATCCCATAATCCCCAGC
AflSat43-131	GAGAGAACTGGCTGTGTGCA	ACAGCAAGATATAGTTACATTGCAC
AflSat44-254	AACTGCGAGAAGTGTGGCCCA	GTCCAGTGCTGCAAGCCGCA
AflSat45-414	ACTTTTATATCTGGCAAGGGCGT	TGTTGATGTACTAGTTTCAGTATGT
AflSat46-153	CGTATCTACCAAGGGTATAGGGT	AGCAGGGAAGTGGTGGCACA
AflSat47-75	TGTGACCAACTACTTCCCATCTGT	ACAGTTGACCAGACGAAAGGAG
AflSat48-167	GCCCTGTGGTCGTCACTCACA	GCCCGAGTTGCTAGTTACAT
AflSat49-138	ACCTCCTCTCTTCTTTTCCGT	CGTGCAGTGGATGCCTTTAACGT
AflSat50-102	GGATGCTGGAGCCACACTACA	TCCAGTTGCGTTTACTGATAAGCT
AflSat51-30	GTAGTCTTCCATCACAGAATAGTC	ACTATTCTGTGATGGAAGACTACT
AflSat52-23	GTGCAACCCCCTTGGCTCGG	CGAGCCAAGGGGGTTGCACG
AflSat53-17	TTTTTCCAAATAATCGTTTTT	ACGATTATTTGGAAAAAAC
DNA/Tc1_74	GGGTGACTTTACGTATGCAG	AAATTGCGTCCACCATGAGC
LINE/Jockey_72	GGGTCTTGGTCGGGTTGAA	TGGCATGGAAACATGCTGAG
LTR/Gypsy_17	GTACCTGAAACAGATAGGCC	CCTCCAACTCTGTTATCGAG

Supplementary Table 2. Main features of 53 satDNA families found in *Abracris flavolineata* individuals with no B chromosome (0B, mean of three individuals), with one B chromosome (1B, mean of two individuals) and with two B chromosomes (2B, mean of two individuals) from Rio Claro/SP population. The asterisks (*) mark sequences that presented increase abundance correlated with increase of B chromosome number and the number signs (#) mark sequences that were observed on the B chromosome by FISH. ML (monomer length), SF (superfamily), K2P (Kimura 2-parameter divergence).

				0B individuals mean		1B individuals mean		2B individuals mean		Coefficient	
Family	**ML**	**AT**	**SF**	**K2P**	**Abundance (%)**	**K2P**	**Abundance (%)**	**K2P**	**Abundance (%)**	**1B/0B**	**2B/0B**
AflSat01#	179	55.9		4.09	1.73151	4.07	1.56725	4.20	1.46928	0.91	0.85
AflSat02#	391	53.7		2.34	0.68864	2.32	0.60298	2.32	0.61016	0.88	0.89
AflSat03#	17	47.1		24.46	0.54214	24.82	0.51374	24.65	0.55049	0.95	1.02
AflSat04	161	67.1		3.58	0.51584	3.50	0.40285	3.70	0.43007	0.78	0.83
AflSat05	437	55.4		2.16	0.18719	2.20	0.17848	2.18	0.16637	0.95	0.89
AflSat06	265	58.1		3.68	0.10558	3.52	0.09728	3.67	0.09723	0.92	0.92
AflSat07#	36	61.1		3.26	0.07981	3.11	0.06769	3.18	0.06359	0.85	0.80
AflSat08#	184	62.5		5.62	0.07437	5.72	0.07328	5.62	0.07350	0.99	0.99
AflSat09	187	59.9		3.17	0.06914	3.22	0.07014	3.21	0.06315	1.01	0.91
AflSat10	297	62.3		7.57	0.06888	7.95	0.06955	7.61	0.06915	1.01	1.00
AflSat11	407	54.5		18.12	0.06437	22.84	0.02512	22.93	0.02339	0.39	0.36
AflSat12	237	62.9		8.05	0.03959	8.57	0.02447	7.98	0.03172	0.62	0.80
AflSat13*#	177	75.1		8.06	0.02836	8.04	0.02908	7.98	0.03021	1.03	1.07
AflSat14*	110	50.0		20.72	0.02162	21.35	0.02206	20.43	0.02256	1.02	1.04
AflSat15	299	58.5	1	12.09	0.01881	11.60	0.01864	12.02	0.01715	0.99	0.91
AflSat16	298	58.4	1	10.56	0.01833	10.58	0.01695	10.56	0.01783	0.92	0.97
AflSat17	231	57.6		15.98	0.01806	17.43	0.01621	16.12	0.01756	0.90	0.97
AflSat18	7	42.9		3.36	0.01609	3.27	0.01055	3.33	0.00907	0.66	0.56
AflSat19	17	64.7		11.59	0.01581	11.81	0.01249	11.62	0.01349	0.79	0.85
AflSat20	233	57.5	2	5.72	0.01510	5.84	0.01394	5.69	0.01538	0.92	1.02

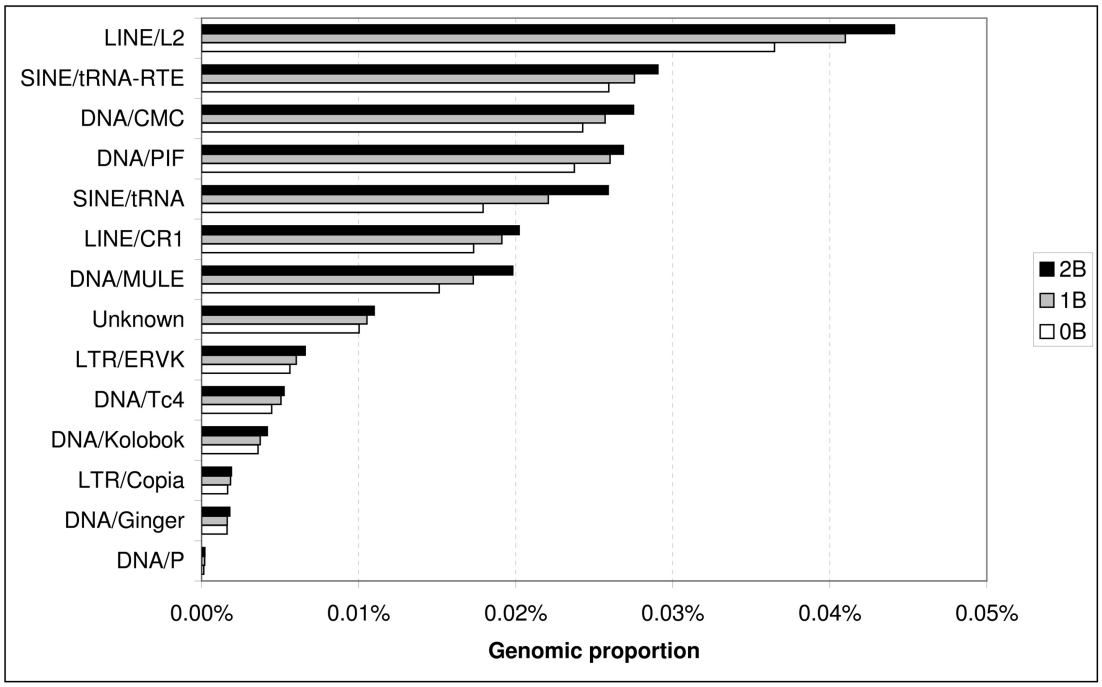
AflSat21	260	54.6		6.68	0.01458	6.33	0.01460	6.26	0.01527	1.00	1.05
AflSat22	135	59.3		16.36	0.01455	16.45	0.01364	16.15	0.01487	0.94	1.02
AflSat23	286	61.2		5.62	0.01357	5.80	0.01227	5.86	0.01158	0.90	0.85
AflSat24	308	57.1		12.19	0.01354	13.26	0.01310	12.07	0.01297	0.97	0.96
AflSat25[#]	40	60.0		9.02	0.01343	9.07	0.01077	9.14	0.00936	0.80	0.70
AflSat26	296	63.5	1	8.99	0.01268	9.62	0.00758	7.10	0.01485	0.60	1.17
AflSat27	609	53.0		2.44	0.01248	2.84	0.00899	2.72	0.00959	0.72	0.77
AflSat28	247	57.1	2	4.63	0.00975	4.90	0.00854	5.16	0.00821	0.88	0.84
AflSat29	616	57.2		7.69	0.00948	8.75	0.00866	7.17	0.00863	0.91	0.91
AflSat30	197	54.8		12.09	0.00852	11.70	0.00838	11.30	0.00861	0.98	1.01
AflSat31	226	61.9		7.98	0.00698	8.73	0.00659	8.74	0.00638	0.94	0.91
AflSat32	429	58.5		12.90	0.00646	10.72	0.00769	10.42	0.00770	1.19	1.19
AflSat33	295	61.0		20.00	0.00640	20.51	0.00630	20.76	0.00604	0.99	0.94
AflSat34	30	46.7		13.94	0.00612	17.16	0.00422	13.83	0.00675	0.69	1.10
AflSat35	362	65.5		0.15	0.00585	0.20	0.00304	0.13	0.00151	0.52	0.26
AflSat36	41	61.0		10.12	0.00501	10.30	0.00538	10.14	0.00504	1.08	1.01
AflSat37	134	67.2		6.16	0.00488	6.17	0.00380	6.18	0.00523	0.78	1.07
AflSat38	377	60.7		6.86	0.00483	8.41	0.00439	9.49	0.00447	0.91	0.92
AflSat39	832	50.8		3.09	0.00455	2.22	0.00690	3.33	0.00411	1.52	0.90
AflSat40[*#]	218	61.5		8.57	0.00431	9.15	0.00452	9.23	0.00500	1.05	1.16
AflSat41	186	46.8		15.46	0.00376	11.68	0.00603	12.87	0.00419	1.61	1.12
AflSat42[*#]	75	38.7		15.26	0.00374	15.73	0.00434	15.30	0.00458	1.16	1.22
AflSat43	131	57.3		7.45	0.00329	7.11	0.00410	7.55	0.00347	1.25	1.05
AflSat44	254	58.3		5.62	0.00297	6.70	0.00277	7.18	0.00300	0.93	1.01
AflSat45	414	62.1		14.31	0.00288	12.61	0.00439	14.04	0.00321	1.52	1.11
AflSat46[#]	153	54.9		7.49	0.00285	7.87	0.00272	8.54	0.00276	0.95	0.97
AflSat47	75	52.0		5.63	0.00224	5.88	0.00221	5.33	0.00249	0.99	1.11

AflSat48	167	56.9	7.03	0.00194	6.63	0.00248	6.22	0.00221	1.27	1.14
AflSat49	138	54.3	6.42	0.00133	6.22	0.00161	6.50	0.00142	1.21	1.07
AflSat50	102	42.2	10.36	0.00104	11.17	0.00097	9.35	0.00080	0.93	0.77
AflSat51	30	60.0	8.99	0.00089	9.23	0.00083	9.79	0.00111	0.93	1.24
AflSat52[*#]	23	30.4	10.02	0.00014	10.53	0.00111	9.89	0.00274	8.03	19.84
AflSat53	17	76.5	30.79	0.00001	22.89	0.00002	28.37	0.00003	1.60	2.29

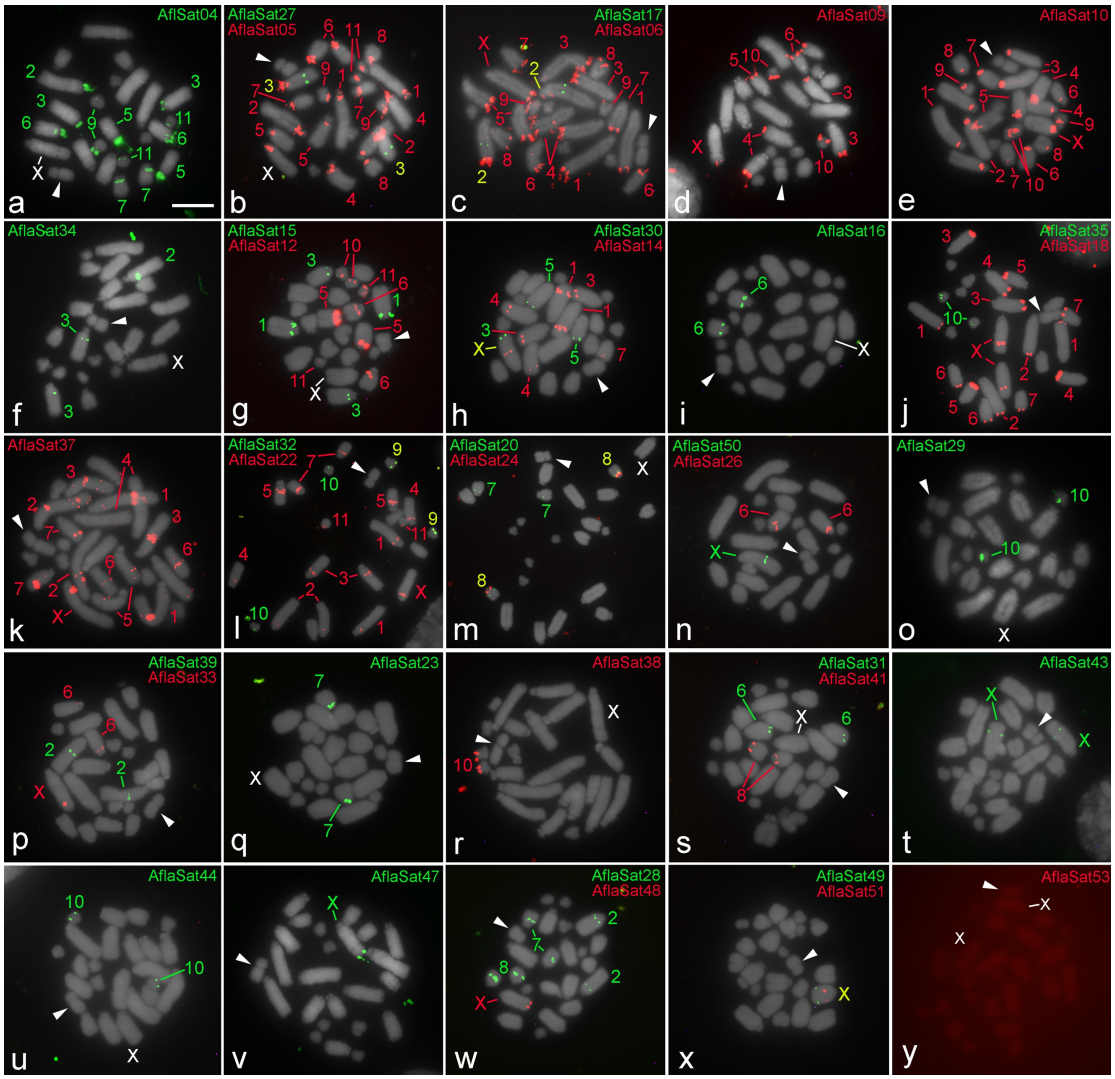
Supplementary Table 3. List of the 28 TEs showing higher abundance in presence of one or two B chromosomes, arranged in order of increasing abundance in the 0B genome. Asterisks indicate TEs selected for FISH mapping.

	Proportion in genomes with				
TE family	0B	1B	2B	1B/0B	2B/0B
LTR/Gypsy*	1.7730%	1.9341%	2.0689%	1.09	1.17
DNA/Maverick	0.2429%	0.2606%	0.2729%	1.07	1.12
LINE/Penelope	0.1951%	0.2165%	0.2266%	1.11	1.16
LINE/L1	0.1626%	0.1794%	0.1967%	1.10	1.21
DNA/Tc1*	0.1586%	0.1720%	0.1783%	1.08	1.12
DNA/DNA	0.1329%	0.1453%	0.1541%	1.09	1.16
LINE/Jockey*	0.0988%	0.1394%	0.1628%	1.41	1.65
RC/Helitron	0.0985%	0.1041%	0.1135%	1.06	1.15
LINE/I	0.0977%	0.1067%	0.1127%	1.09	1.15
DNA/Mariner	0.0713%	0.0791%	0.0822%	1.11	1.15
DNA/hAT	0.0580%	0.0626%	0.0669%	1.08	1.15
DNA/Sola	0.0537%	0.0598%	0.0681%	1.11	1.27
LTR/Pao	0.0471%	0.0524%	0.0549%	1.11	1.17
LINE/R1	0.0421%	0.0468%	0.0487%	1.11	1.16
LINE/L2	0.0365%	0.0410%	0.0441%	1.12	1.21
SINE/tRNA-RTE_3	0.0259%	0.0276%	0.0291%	1.06	1.12
DNA/CMC_2	0.0243%	0.0257%	0.0275%	1.06	1.13
DNA/PIF	0.0237%	0.0260%	0.0269%	1.10	1.13
SINE/tRNA	0.0179%	0.0221%	0.0259%	1.23	1.44
LINE/CR1	0.0173%	0.0191%	0.0202%	1.10	1.17
DNA/MULE	0.0151%	0.0173%	0.0198%	1.14	1.31
Unknown/Unknown_10	0.0100%	0.0105%	0.0110%	1.05	1.10
LTR/ERVK_1	0.0056%	0.0060%	0.0066%	1.07	1.17
DNA/Tc4_4	0.0045%	0.0051%	0.0053%	1.13	1.18
DNA/Kolobok_13	0.0036%	0.0037%	0.0042%	1.04	1.16
LTR/Copia_9	0.0017%	0.0018%	0.0019%	1.11	1.15
DNA/Ginger_3	0.0016%	0.0016%	0.0018%	1.00	1.10
DNA/P_1	0.0001%	0.0002%	0.0002%	1.46	1.57
Total	3.42%	3.77%	4.03%		

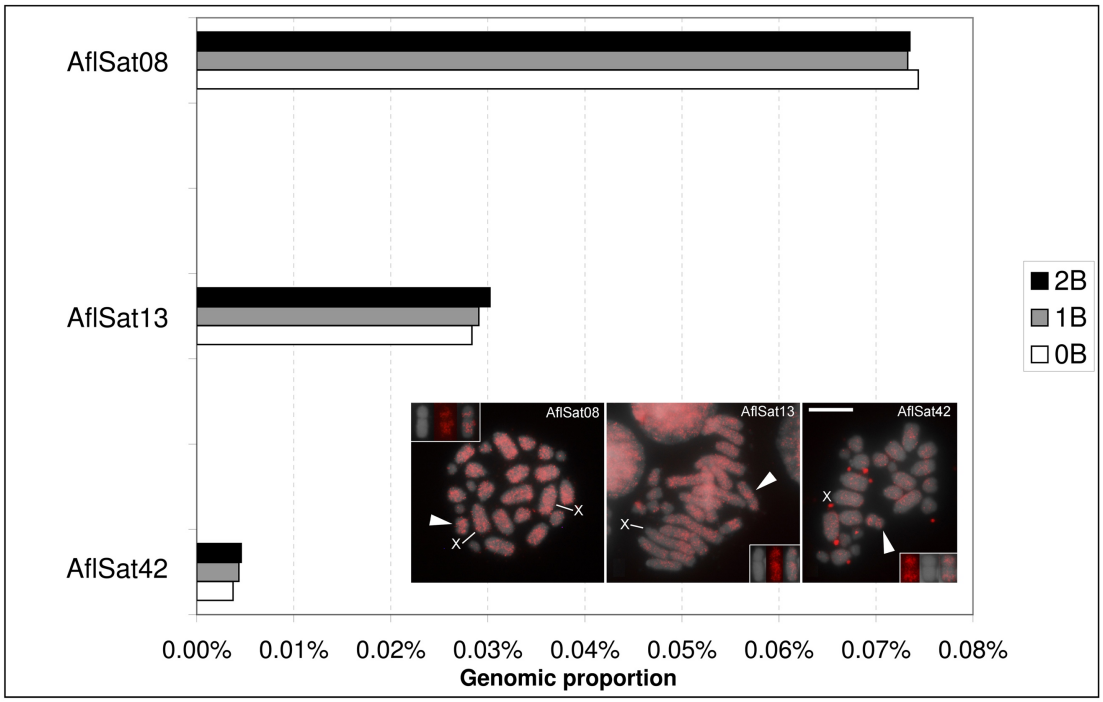
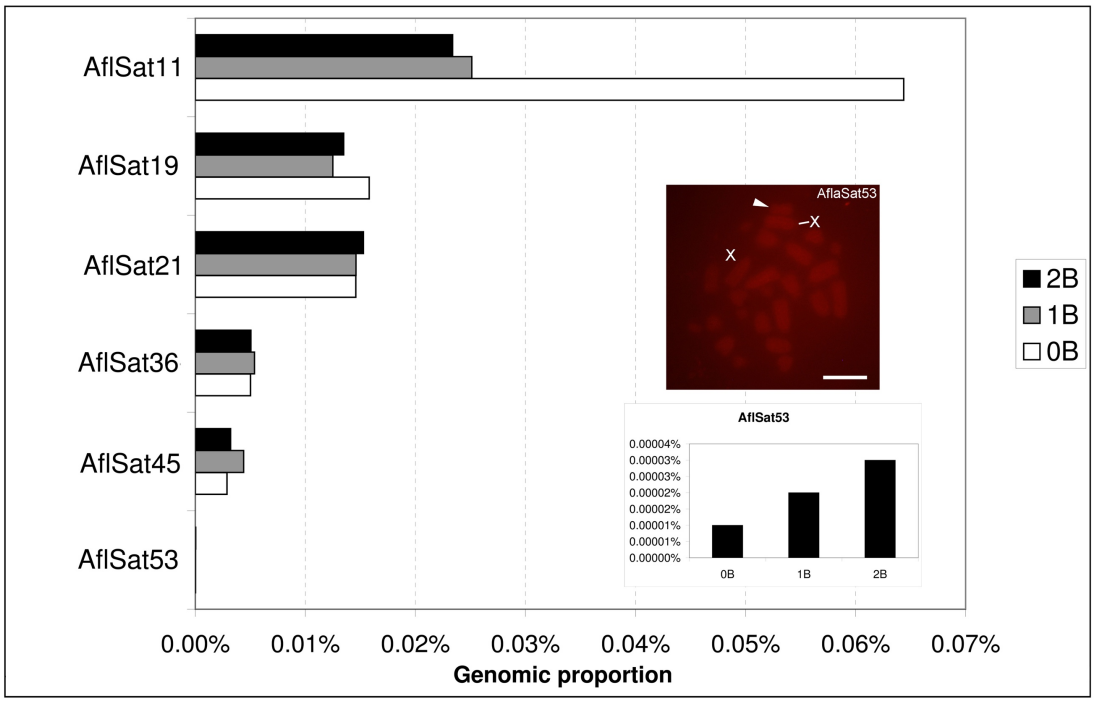
Supplementary Figure 1. Comparative genomic proportion between the 0B, 1B and 2B genomes for he TEs with low abundance in the *Abracris flavolineata* genome.



Supplementary Figure 2. (a-x) FISH mapping of satDNAs that displayed signals exclusively in the A complement (no signals on B chromosome) plus one example of satDNAs with no signal (y) in mitotic metaphase spreads from embryos. Chromosomes with signals are indicated in the corresponding color for each satDNA family that is shown for each metaphase spread. Chromosomes with signals for two satDNAs mapped in the same metaphase spread are labeled in yellow. The X chromosomes are identified: X (male embryos) and XX (female embryos). Chromosomes with FISH signals are identified in each image. Bar = 10 μm.



Supplementary Figure 3. Comparative genomic proportion and FISH mapping for three satDNAs showing the dotted pattern on the B chromosomes. Note the occurrence of signals on the long chromosomal arms corresponding to euchromatic regions of A complement. On the B chromosome (arrowheads), the signals are virtually distributed along its entire extension, but they are less abundant on pericentromeric and terminal regions for AflSat08-184 (a) and AflSat42-75 (c). For AflSat13-177 (b) it is less abundant in the proximal region of the short arm. Boxed are the B chromosomes showing DAPI channel (gray), signal channel (red) and merged channels. X chromosomes are indicated. Bar = 10μm.



Supplementary Figure 4. Comparative genomic proportion for six satDNAs with no FISH signals (NS). Due to low abundance in comparison to the other satDNAs, a separate graph for AflaSat53-17 is additionally shown. The mitotic chromosome spread shows the signal channel (red) for FISH mapping for this same element. The B chromosome is indicated by an arrowhead, and the X chromosomes are indicated by letters. Note the absence of visible signals. Bar = 10 μm.