Multiple recent sex chromosome fusions in *Drosophila virilis* associated with elevated satellite DNA abundance

Jullien M. Flynn*, Kevin B. Hu, Andrew G. Clark

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY

*Corresponding author: jmf422@cornell.edu

ABSTRACT

Repetitive satellite DNA is highly variable both within and between species, and is often located near centromeres. However, the abundance or array length of satellite DNA may be constrained or have maximum limits. *Drosophila virilis* contains among the highest relative satellite abundances, with almost half its genome composed of three related 7 bp satellites. We discovered a strain of *D. virilis* that has 15% more pericentromeric satellite DNA compared to other strains, and also underwent two independent centromere-to-centromere sex chromosome fusion events. These fusions are presumably caused by DNA breakage near the pericentromeric satellites followed by repair using similar repetitive regions of nonhomologous chromosomes. We hypothesized that excess satellite DNA might increase the risk of DNA breaks and genome instability when stressed, which would be consistent with the apparent high rate of fusions we found in this strain. To directly quantify DNA breakage levels between strains with different satellite DNA abundances, we performed the comet assay after feeding flies gemcitabine and administering low-dose gamma radiation. We found a positive correlation between the rate of DNA breakage and satellite DNA abundance. This was further supported by a significant decrease in DNA breakage in an otherwise genetically identical substrain that lost the chromosome fusion and several megabases of satellite DNA. We find that the centromere-to-centromere fusions resulted in up to a 21% nondisjunction rate between the X and Y chromosomes in males, adding a fitness cost. Finally, we propose a model consistent with our

30 data that implicates genome instability as a critical evolutionary constraint to satellite

31 abundance.

32

INTRODUCTION

Satellite DNA consists of long arrays of tandemly repeated sequences, and is often located near centromeres in heterochromatin (reviewed in Thakur *et al.* 2021). Satellite DNA varies greatly in sequence and abundance within and between species (Subirana *et al.* 2015; Wei *et al.* 2018; Cechova *et al.* 2019), and this can be partially explained by high rates of copy number mutation (Flynn *et al.* 2017). Although previously assumed to be inert "junk," recent work has shown that satellite DNA is involved in essential processes in the cell, thus variation in it may be biologically important (Jagannathan *et al.* 2018; Mills *et al.* 2019). Although satellite DNA differences between some species have been linked to reproductive incompatibilities (Ferree and Barbash 2009; Jagannathan and Yamashita 2021), the biological implications of intraspecies abundance variation has not been explored. Satellite DNA can vary in abundance by several megabases among individuals of the same species, including in flies and humans (Miga *et al.* 2014; Wei *et al.* 2014; Flynn *et al.* 2020). Satellite DNA appears to be constrained by maximum limits, with no species studied so far having more than about half of their genome made up of satellite DNA (Gall and Atherton 1974a; Fry and Salser 1977; Petitpierre *et al.* 1995). The genomic abundance of transposable elements (TEs), the other highly pervasive type of repetitive DNA, seems to be less constrained than satellite DNA with many genomes over 50% and some containing up to 85% TE content (Anderson *et al.* 2019). The nature of satellite DNA with long tandem arrays of the same sequence, may impose instability that prevents it from expanding beyond a threshold, compared to more diverse sequences interspersed in the genome.

Drosophila virilis is an excellent model for studying satellite DNA variation. *D. virilis* has the highest relative abundance of simple satellite DNA (defined as satellites with unit length <=20 bp) compared to any other studied species. Three 7 bp satellites, AAACTAC, AAACTAT, and AAATTAC take up over 40% of the genome in *D. virilis*, and they form arrays tens of megabases long in the pericentromeric region (Gall *et al.* 1971; Gall and Atherton 1974; Flynn *et al.* 2020). The extremely high relative abundance of satellite DNA in *D. virilis* makes it an ideal system in which to ask whether there are constraints or maximum limits on satellite abundance. One

63 strain in particular, vir00 (15010-1051.00), contains 15% more pericentromeric satellite DNA
64 compared to other *D. virilis* strains (Flynn *et al.* 2020 Figure 4B). In past modeling efforts,
65 satellite DNA arrays have been proposed to be weakly deleterious until they reach a maximum
66 length beyond which they are not tolerated by selection (Charlesworth *et al.* 1986). Slow DNA
67 replication or development time have been suggested as mechanisms to enforce strong
68 negative selection against long satellite arrays, however empirical evidence for this has been
69 limited (but see Bilinski *et al.* 2018). Here, we propose genome instability as a constraint on
70 excessively long or abundant satellite DNA arrays.
71
72 Genome instability is characterized by DNA damage that often results in large-scale mutations
73 or rearrangements, and is a fundamental driver of karyotype evolution, chromosomal
74 disorders, and cancer rearrangements (Black and Giunta 2018; Mayrose and Lysak 2021). The
75 first step is spontaneous DNA damage such as double-stranded DNA breaks, one of the most
76 dire events to occur in a cell (Featherstone and Jackson 1999). Even if DNA breaks are repaired,
77 they often result in large-scale genome rearrangements. Most DNA breaks leading to
78 rearrangements occur near the centromere and within or near satellite DNA, including in
79 human genomes (Black and Giunta 2018; Balzano *et al.* 2020). This may be due to intrinsic
80 instability of satellite DNA arrays caused by replication stress of polymerases progressing
81 through highly repetitive sequences, or the formation of unstable DNA topology (Barra and
82 Fachinetti 2018). Robertsonian translocations, one of the most common rearrangements in
83 medical genetics and evolution, occur when there are breaks near the centromere of
84 acrocentric chromosomes and when they are repaired they are fused to each other (Mayrose
85 and Lysak 2021). Robertsonian translocations are associated with multiple miscarriages in
86 humans and aneuploidy disorders like Patau and Down syndromes (Braekeleer and Dao 1990),
87 with increased rates of aneuploidy driven by increased nondisjunction of Robertsonian or fused
88 chromosomes (Schulz *et al.* 2006). We hypothesize that variation in pericentromeric satellite
89 DNA abundance influences the risk of genome instability events. Specifically, excess satellite
90 DNA might increase the risk of genome instability and genome rearrangements.
91

Sex chromosome evolution has been studied for decades, mainly making use of sex chromosomes that have arisen in different time periods (Charlesworth and Charlesworth 2000). In Drosophila, when an autosome fuses to either an X or Y chromosome, a so-called neo-Y chromosome is formed. Because either the fused or unfused version of the chromosome will only be present in males and male Drosophila do not undergo recombination, mutations immediately begin to accumulate through Hill-Robertson interference and other linked-selection processes (Charlesworth and Charlesworth 2000). In the genus Drosophila, autosomes have fused to sex chromosomes multiple independent times (Nozawa *et al.* 2021): *D. pseudoobscura* (10 million years), *D. miranda* (1 million years; Bachtrog 2013), *D. albomicans* (0.24 million years; (Wei and Bachtrog), and *D. americana* (29 thousand years; (Vieira *et al.* 2006). Neo-sex chromosomes formed by *de novo* sex chromosome fusions are rare and actually under-represented compared to autosomal fusions in *Drosophila* (Anderson *et al.* 2020), and have never been discovered at their infancy before detectable divergence has occurred. How new sex chromosome fusions become stable, and how they compete with the ancestral karyotype within a species is unknown.

Here, we describe the discovery of two independent and extremely recent sex chromosome-autosome fusions in one *D. virilis* strain, vir00. We hypothesize that this strain has been more prone to DNA instability events, possibly caused by its excessive satellite DNA abundance. After applying DNA replication and physical stressors, we measured DNA damage levels directly and demonstrate that the DNA damage response is associated with satellite abundance. We use two genetically identical strains that differ only by a chromosome fusion and satellite DNA abundance, and demonstrate that the strain with more satellite DNA has significantly increased DNA damage when stressed. Finally, we propose a model that genome instability may impose a constraint on satellite abundance, a model that is entirely consistent with our data.

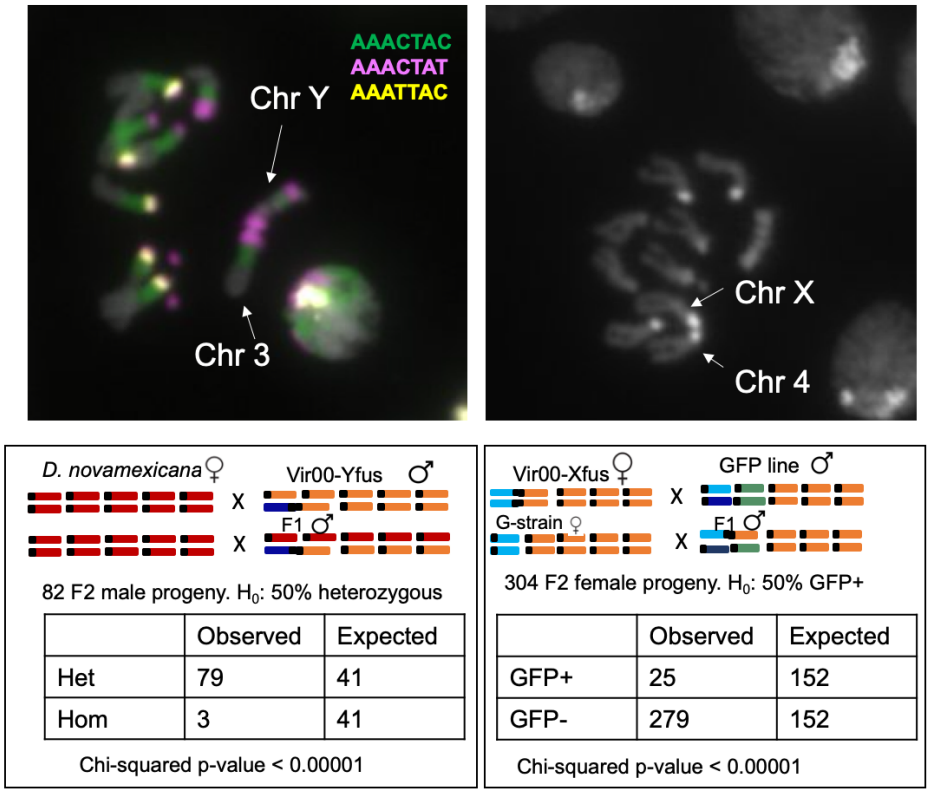
RESULTS

Two novel sex chromosome Robertsonian translocations in *D. virilis* strain vir00

In summer 2019, we performed DNA fluorescence in-situ hybridization (FISH) on larval neuroblast nuclei of the vir00 strain that we had obtained several months earlier from the National Drosophila Species Stock Center. We discovered it contained a Y-autosome fusion (Figure 1A). The Y chromosome is recognizable in *D. virilis* because it has a distinct DAPI staining intensity pattern, and contains a distinct arrangement of satellite DNA (Flynn *et al*. 2020). However, all other chromosomes are difficult to distinguish in metaphase spreads, so we could not immediately determine the fusion partner. We observed that the fused chromosome contained the same centromere-proximal satellite as the Y chromosome, AAACTAT. To determine if this chromosome fusion was present in other strains from similar geographical locations, we imaged larvae from strains vir08, vir86, and vir48, which were collected from different localities in California and Mexico (vir00 was collected from California). No larvae screened in these other strains contained any chromosome fusions. We designate the vir00 substrain with the Y fusion as vir00-Yfus.

In Fall 2019, we obtained a second stock of vir00 from the National Drosophila Species Stock Center. We set up single pair crosses and performed larval neuroblast squashes to karyotype multiple male progeny of each cross. Surprisingly, we found that the Y-autosome fusion was not present in any of the larvae karyotyped. However, 3/10 crosses karyotyped contained a different fusion. This new fusion contained the brightly-dapi-staining AAATTAC satellite as the centromere-proximal satellite, indicating that it is completely distinct from the Y fusion and involved different chromosomes. We hypothesized that this new fusion involved the X chromosome for several reasons; 1) based on centromere satellite identity, it had a 67% probability (Flynn *et al.* 2020); 2) it was found only in single copy in male larvae, but sometimes two copies in female larvae; 3) it was associated with observed X-Y nondisjunction events, such as the presence of an XXYY female (Figure S1). We performed single-pair crosses and screened the resulting progeny until we isolated a substrain fixed for the X fusion, and called this substrain vir00-Xfus. We maintained one of the cross descendants from the 2019 stock that did not have any evidence of the X fusion, which we designate vir00-Nofus. We inferred that both X and Y fusions likely represent canonical Robertsonian translocations, in which two acrocentric

chromosomes that underwent DNA breakage were fused together at the centromere during repair.



	Observed	Expected
Het	79	41
Hom	3	41

Chi-squared p-value < 0.00001

	Observed	Expected
GFP+	25	152
GFP-	279	152

Chi-squared p-value < 0.00001

Figure 1. Discovery and genetic validation of two independent fusion events in vir00. A) DNA-FISH in metaphase chromosome squashes demonstrating the Y fusion. B) DAPI staining of metaphase chromosome squashes demonstrating the X-4 fusion. C) Genetic validation of the Y-3 fusion and D) the X-4 fusion. Red chromosomes: *D. novamexicana*; orange: wildtype *D. virilis* autosomes; green: autosome containing a GFP marker; light blue: X chromosome; dark blue: Y chromosome.

Genetic validation revealed Y-3 and X-4 fusions

We designed two separate two-generation crossing experiments to validate the fusions genetically and identify the autosome each sex chromosome is fused to. Both experiments exploited autosomal markers which we could determine if they were segregating non-independently of sex. The first experiment to validate the Y-autosome fusion used crosses between vir00-Yfus and *D. novamexicana* and scoring of microsatellite loci on each candidate autosome. We scored 82 F2 male progeny for the Chr3 marker, and 79/82 contained both alleles, whereas the null hypothesis was 50% should contain both alleles (Figure 1C, Figure S2).

171 The three progeny that did not contain both alleles were determined to have nondisjunction
172 events and did not contain the Y chromosome (Figure S3). We concluded that the Y
173 chromosome is fused to Chr3, and that the fusion was fixed in this subline. The other markers
174 on Chr2, Chr4, and Chr5 segregated independently of sex and acted as negative controls (Table
175 S1). We also did a negative control with the same crossing scheme except with vir08 instead of
176 vir00 (Chr3 chi-square p = 0.39, N=22, Table S1).
177
178 The second experiment to validate the X-autosome fusion used crosses to *D. virilis* transgenic
179 lines containing GFP markers on one of each of the candidate autosomes. For the crosses to
180 Chr4-GFP (vir95), we phenotyped 304 F2 female progeny, and found that progeny containing a
181 GFP signal were significantly depleted compared to the Mendelian expectation of 50% (Figure
182 1D). This indicated that the X chromosome was fused to Chr4. This crossing scheme with two
183 other candidate autosomes did not show association of GFP signal with sex (Chi-square p-value
184 > 0.1, Table S2). We also performed a negative control with the *D. virilis* genome strain instead
185 of vir00-Xfus crossed to the Chr4-GFP line and found the GFP signal was independent of sex
186 (Chi-square p-value = 0.92, Table S2).
187
188 We validated that vir00-Yfus and vir00-Xfus are the same genetic line and not the result of
189 contamination from other lines either in our lab or the stock center. We made use of medium-
190 coverage whole genome sequencing from Flynn et al. (2020) to design primers to amplify
191 singleton insertion/deletion variants present only in vir00 (the version that was sequenced was,
192 in hindsight, vir00-Yfus) and not in any other wildtype *D. virilis* strain present in the stock center
193 (Table S3). We designed primers to amplify four loci on chromosomes 2, 3, 5, and 6 which
194 contain a homozygous 12-13 bp deletion in vir00 compared to the reference and the other
195 strains (Table S4). We found that both versions of vir00 contained the deletion at each of these
196 loci, confirming that these substrains are indeed the same strain and not a contamination
197 (Figure S4). Although we did the above indel experiment first, the whole genome resequencing
198 SNP analysis (below) was also concordant with these substrains being the same genetic line.
199

Satellite DNA decreased in vir00-Nofus compared to vir00-Yfus

We used Illumina to resequence the three vir00 substrains: vir00-Yfus, vir00-Xfus, and vir00-Nofus. We then used k-Seek to quantify the satellite abundance in each substrain (Wei *et al.* 2014). Assuming the most ancestral strain was vir00-Yfus, we found that satellite abundances decreased in the two derived lines vir00-Xfus and vir00-Nofus (Table S5). In vir00-Xfus, there was an 8% decrease only in the centromere proximal satellite of ChrX and Chr4 (AAATTAC). vir00-Nofus had an overall 12% loss of satellite DNA compared to vir00-Yfus (Figure 3b). vir00-Nofus had a similar 8% decrease in AAATTAC, in addition to a 10% loss in the pericentromeric satellite (AAACTAC) and a 13% loss in the centromere-proximal satellite of ChrY and Chr3 (AAACTAT). These data are consistent with our interpretation that vir00-Nofus is the result of breaking apart of the fusion chromosomes, and we conclude that significant satellite DNA was lost from all three pericentromeric and centromeric satellites in vir00-Nofus.

DNA damage levels in response to stress is related to satellite abundance in *D. virilis*

The vir00 strain, with the highest abundance of satellite DNA, contained two recent independent Robertsonian translocations. There are multiple steps required to detect a Robertsonian translocation: DNA breakage near the centromere, repair and fusion with another acrocentric chromosome, retention of a functional centromere, and intergenerational retention. To isolate the first step and the fundamental process in genome instability, we sought to measure DNA breakage directly. Since vir00-Yfus also contained the highest abundance of satellite DNA compared to the other strains, we wanted to test if variation in DNA breakage level after stress was associated with satellite DNA abundance. We therefore measured DNA damage levels in 7 different *D. virilis* strains, including vir00-Yfus and vir00-Nofus, with varying abundances of satellite DNA in response to replication stress and low-level radiation.

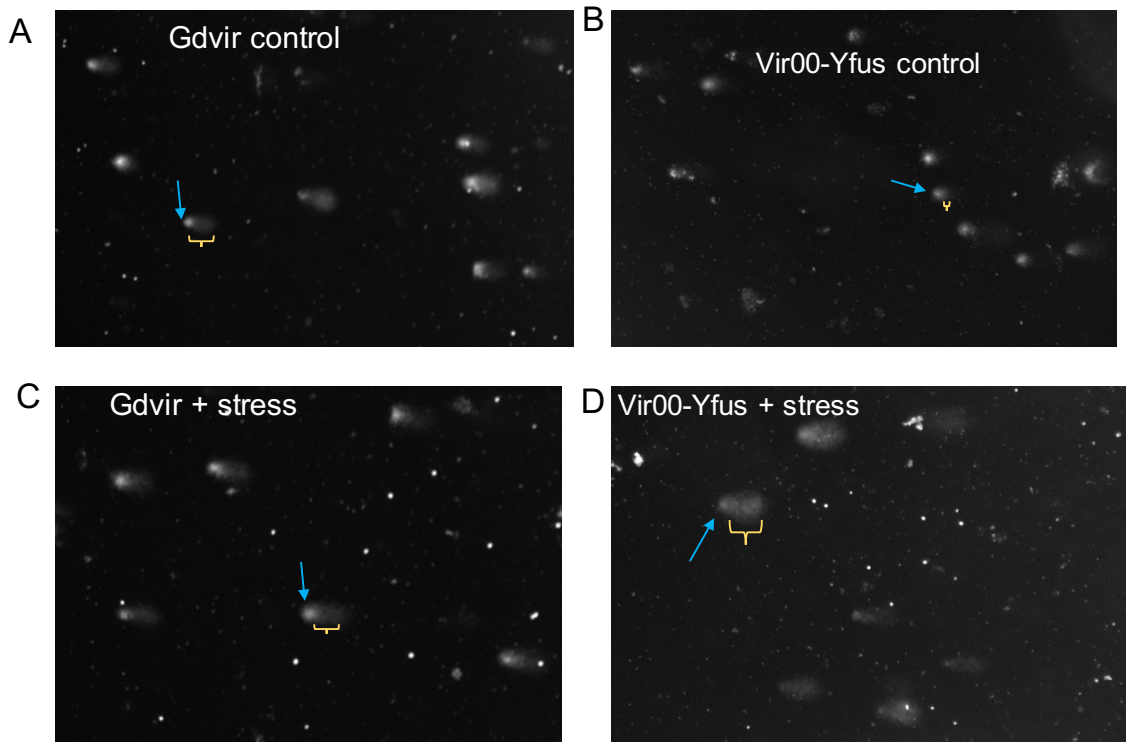


Figure 2. Comet assay images of nuclei isolated from *D. virilis* adult testes of the genome strain vir87 (A/C) and from vir00-Yfus (B/D) with and without stress (gemcitabine + radiation). Blue arrows point to the comet head and yellow brackets indicate the comet tails. The comet head and tail distribution for GDvir are of similar size and intensity. In vir00-Yfus, the comet tail is largely diffuse and expanded in the stress conditions compared to the control. These strains represent the lowest and highest satellite DNA abundances which have the least and greatest increase in DNA damage in response to stress, respectively.

Since spontaneous DNA breakage events are rare and we did not want to confound our data with breaks that occur as part of meiotic crossing over, we used stressors to elevate the rate of DNA damage. This would allow us to potentially detect a difference in the phenotype of interest between strains and also amplify possible types of stress imposed by excess satellite DNA content. We fed 0-1 day old adult flies the nucleoside analog gemcitabine for 8 days, which stalls replication forks and acts as a sensitizer for radiation via the Rad51 pathway (Kobashigawa *et al.* 2015). We then irradiated these sensitized flies with gamma rays at low level radiation (10 Gy). For each of seven strains tested, we included a control which was fed with the same liquid food with no gemcitabine and did not receive radiation treatment. We used the comet assay or single-cell gel electrophoresis to measure DNA damage in the male

248 germline in each line (Figure 2). We quantified comets using OpenComet software and used the
249 olive moment measurement as the statistic representing DNA damage (Gyori *et al.* 2014) (Table
250 S6). To detect differences in DNA damage in response to stress between strains, we took the
251 mean difference in olive moments between the control and stress treatments (see Methods).
252
253 The strain with the lowest satellite abundance, the genome strain, contained the lowest DNA
254 damage response and this was significantly lower than all other strains tested (Figure 2A,C,
255 Figure 3A). The strain with the highest satellite abundance, vir00-Yfus, had among the highest
256 DNA damage responses (Figure 2B,D), but was not significantly different from two other strains
257 vir8 and vir48 (Figure 3A). Like most other phenotypes, there are likely multiple genetic factors
258 contributing to the variation we found. Ideally, to demonstrate that satellite DNA is a causal
259 factor, we would manipulate satellite DNA abundance and test the DNA damage phenotypes,
260 but multi-megabase long arrays of satellite DNA cannot be manipulated with traditional
261 genome editing. However, vir00-Yfus and vir00-Nofus differ only by a chromosome fusion and
262 12% satellite abundance. Thus, if there is a difference in DNA damage response between these
263 substrains, it may be caused by differing satellite abundance. vir00-Nofus, which contained 12%
264 less satellite DNA than vir00-Yfus, had a significantly reduced DNA damage response, which is
265 concordant with our expectations that satellite DNA plays a causal role (Figure 3B).
266

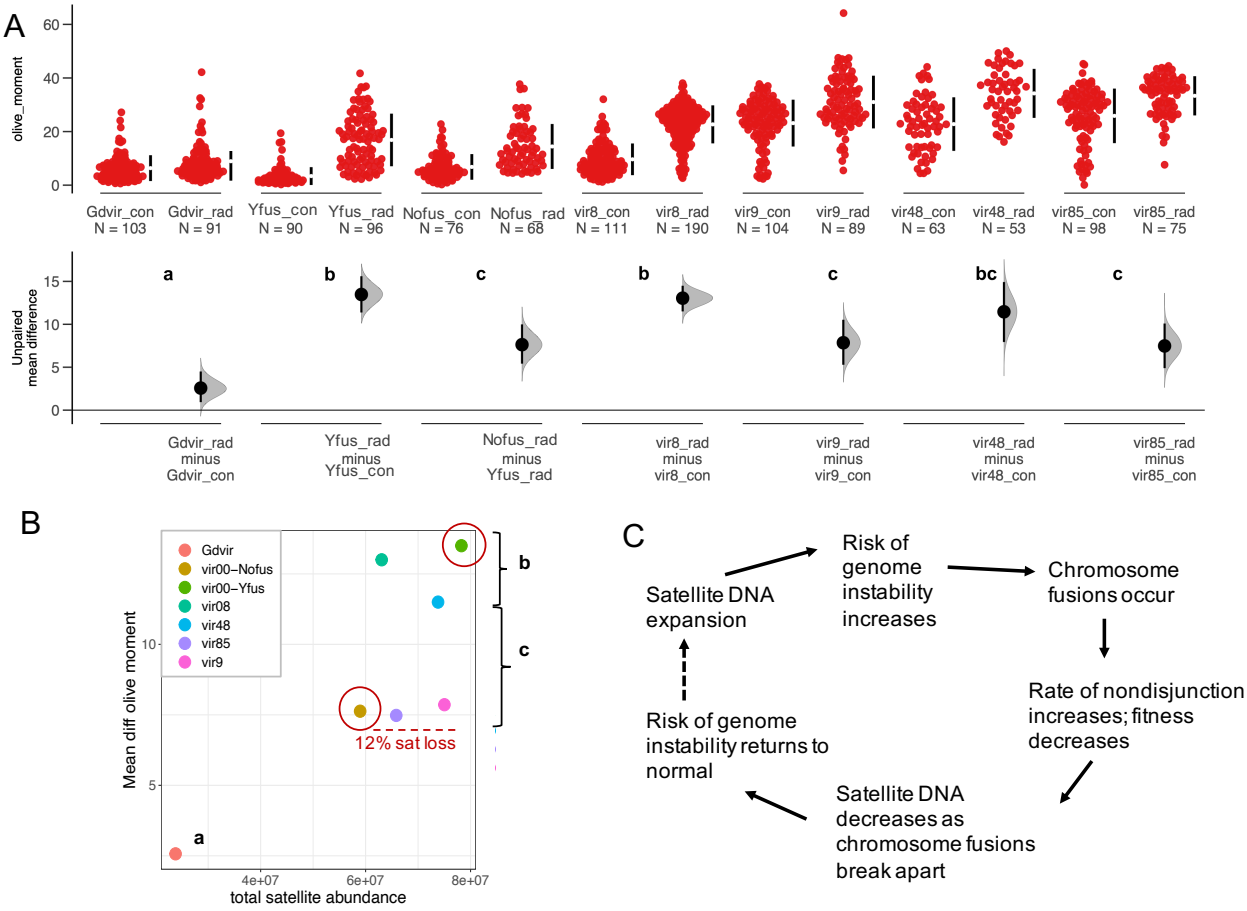


Figure 3. Satellite DNA is associated with DNA breakage in response to stress. A) Olive moment, a statistic of the comet assay, measuring DNA damage for 7 different strains paired with a control and stress treatment. N indicates the number of nuclei analyzed for each treatment. Flies treated with gemcitabine and radiation are suffixed with "rad" and control flies are suffixed with "con". The lower panel shows the unpaired mean difference between "rad" and "con" for each fly strain (black dot) and 95% confidence intervals (black line) produced from dabestr (5000 bootstrap method). Groups a, b, and c indicate samples with overlapping confidence intervals. B) The satellite DNA abundance of each *D. virilis* strain (x axis) and its unpaired mean difference between treatment and control (y axis). vir00-Nofus experienced a 12% loss in satellite DNA compared to vir00-Yfus, which is associated with a significantly decreased DNA damage in response to stress. C) Model that is consistent with our observations and data.

Extremely minimal degradation of the neo-Y chromosome

The Y-fused version of Chr3 is expected to accumulate mutations independently of the autosomal version of Chr3 over time because of the halt in recombination in male flies. This

would be represented in the mapping of short-read sequencing data by elevated heterozygosity on Chr3. Specifically, if the mutations arose after the fusion of Chr3 to the Y, they would not be present in any other strains of virilis, assuming no recurrent mutation. Thus, we used GATK genotyping and found heterozygous singletons unique to vir00-Yfus compared to other non-vir00 strains on each autosome (Table S3). We found that the number of heterozygous singletons was modestly but significantly enriched on Chr3 in vir00-Yfus: 2.45 SNPs/Mb more than other autosomes (permutation test, $p < 0.001$). We suggest this elevated density of singleton heterozygous sites may be due to the fusion with the Y chromosome and lack of recombination over several generations. Assuming the enrichment was caused only by the Y-fusion and that the neo-Y (Chr3) evolved clonally in a single lineage, we roughly estimate the Y-3 fusion occurred 1000-2000 generations ago.

Because the X-4 fusion was segregating with the no-fusion karyotype, recombination in heterozygotes would prevent degeneration of the neo-Y version of Chr4. Furthermore, we believe the X-4 fusion occurred between 2018-2019 since it was not present in the stock we obtained in 2018. We did not observe an enrichment of heterozygous singletons on Chr4 in vir00-Xfus. However, there was still a slight enrichment on Chr3 in vir00-Xfus (1.45 SNPs/Mb, permutation test $p = 0.031$), supporting our assumption that vir00-Xfus is the result of the breaking apart of the Y fusion and forming a new fusion.

Nondisjunction between the X and Y chromosomes is highly elevated in chromosome fusion lines

Nondisjunction occurs when homologous chromosomes fail to separate at meiosis, and results in aneuploidy in the progeny. Autosomal and X chromosome aneuploidy is lethal in flies; but Y chromosome aneuploidy is viable: females that have a Y chromosome (XXY) are fertile, males with no Y chromosome (XO) are sterile, and males with two Y chromosomes (XYY) are fertile. Elevated rates of X-Y nondisjunction represent a fitness cost because zygotes with infertile or lethal karyotypes will form at increased frequency. We found some evidence of nondisjunction

316 in the vir00-Yfus genetic validation (Figure S3), and also common Y chromosome aneuploidy in
317 the stock of vir00-Xfus (Figure S1). Since both fusions involve the sex chromosomes, we tested
318 the rate of primary X-Y nondisjunction in males in all three versions of vir00, along with the *D.*
319 *virilis* genome strain as a control. We first fully isolated the vir00-Nofus strain to ensure no
320 fusion chromosomes were segregating, which even at low frequency could increase the rate of
321 nondisjunction in the line.
322
323 We crossed individual males of each strain to genome strain females and genotyped for the
324 presence or absence of the Y chromosome in progeny. Female progeny containing a Y
325 chromosome indicate XY sperm from the father, and male progeny lacking a Y chromosome
326 indicate nullisomic sperm from the father. Since nondisjunction was extremely high in vir00-
327 Xfus, 2/7 fathers tested were of XYY karyotype, which we could infer if more than half of his
328 female progeny contained a Y chromosome (Maggert 2014). We eliminated these fathers'
329 progeny from the primary nondisjunction rate calculation. No fathers tested from other
330 sublines were determined to be XYY. Males without a Y chromosome would not produce
331 progeny. We found that vir00-Yfus had a slightly elevated nondisjunction rate of 4.5%,
332 compared to the genome strain control of 1.2%, but it was not statistically significant with the
333 sample sizes we used (Table 1). vir00-Xfus had an extremely high primary nondisjunction rate of
334 21% (e.g. Figure S5), which was significantly higher than that of all other substrains ($p < 0.004$,
335 pairwise proportion test, Table 1). Surprisingly, vir00-Nofus had a significantly elevated
336 nondisjunction rate compared to the genome strain control, at 5.7%. This was not statistically
337 different from vir00-Yfus.
338

Table 1: Nondisjunction of the X and Y chromosomes in males is elevated in chromosome fusion lines.

Line	Nondisjunction rate	Fathers with nondisjunction/ total fathers	Aneuploid progeny/ total progeny	Significance group
GDvir	1.2%	1 / 10	3 / 248	a
vir00-Yfus	4.5%	4 / 7	9 / 198	ab
vir00-Nofus	5.7%	4 / 6	11 / 192	b
vir00-Xfus*	21%	5 / 5	27 / 127	c

*only including primary nondisjunction

Discussion

D. virilis has the highest relative satellite DNA abundance of any studied species. Is there a maximum limit of satellite DNA a genome can tolerate before there are negative consequences? We found that vir00, the strain with 15% more pericentromeric satellite DNA than other strains recently underwent two independent chromosome fusion events. vir00-Yfus has the highest satellite DNA abundance and among the highest DNA damage level in response to stress. vir00-Nofus, with 12% less satellite DNA, which was presumably lost in the breaking apart of the chromosome fusions, has significantly lower DNA damage level in response to stress. We propose a model that places a constraint on satellite DNA abundance and can explain our findings with the vir00-line (Figure 3C). If satellite DNA expands, the risk of DNA breakage and genome instability increases and chromosome fusions may occur. These chromosome fusions may increase the nondisjunction rate and make the affected line less fit. The chromosome fusions may later break apart; and as this occurs pericentromeric satellite DNA may be lost (Figure 3C). This in turn results in a decreased risk of genome instability. The

question of whether satellite DNA abundance influences genome instability has had considerable interest recently (Arora *et al.* 2021), however a convenient system in which to test this question has been lacking.

There are several possible mechanisms that might cause excess satellite DNA to increase the risk of DNA breakage. Satellite DNA may form complex structures, loops, or non-B DNA, and increasing the length of the arrays may make these regions more unstable in cis (Barra and Fachinetti 2018). It may be challenging for polymerases to replicate many megabases of tandem satellite sequence, and longer arrays may have a higher risk of stalling polymerases and thus DNA breaks in cis (Barra and Fachinetti 2018). Finally, increased satellite DNA may titrate away binding proteins that maintain genome stability, in trans (Francisco and Lemos 2014; Brown *et al.* 2020a; Giunta *et al.* 2021). Whether there is a maximum threshold of satellite array length or abundance, as suggested in Charlesworth *et al.* (1986), or whether risk of genome instability varies continuously with satellite abundance will require further study.

Chromosome rearrangements such as Robertsonian translocations are rare occurrences, but are commonly the cause of karyotype evolution between species (Mayrose and Lysak 2021). In humans, rearrangements resulting from breaks near the centromere are associated with miscarriages and in the case of somatic rearrangements, cancer (Barra and Fachinetti 2018). Previous studies have found repetitive element involvement in chromosome rearrangements in multiple species (Paço *et al.* 2015; Reis *et al.* 2018). However, genetic variation of repetitive DNA within species has never been associated with increased risk of genome instability. Here our data suggests, at least in *D. virilis* strain vir00, that elevated satellite DNA may play a causal role in increasing the risk of DNA breakage, which can lead to deleterious rearrangements.

We identified elevated nondisjunction rates as a cost of chromosome fusions. Surprisingly, the rate of nondisjunction was 5-fold higher in vir00-Xfus compared to vir00-Yfus. With a nondisjunction rate of 21%, a significant proportion of abnormal karyotypes will be produced, such as XO (sterile), XYY (viable and fertile), and XXY or XXYY (viable and fertile), all of which we

found in cytological samples. Further mating between these abnormal karyotypes will produce significant proportions of sterile or inviable karyotypes like XYYY or XXX, which will further decrease the fitness of this line. Furthermore, karyotypes with extra Y chromosomes such as XYY and XXY have been found to have decreased lifespan (Brown *et al.* 2020b). Although we did not assay the nondisjunction rate of the fused autosome, it is possible the nondisjunction rate of Chr4 is also elevated in vir00-Xfus, which would further decrease the fitness of the line because Chr4 aneuploidy is expected to be lethal (Lindsley *et al.* 1972). The extreme nondisjunction rate indicates X-Y pairing is severely disrupted due to the X-4 fusion, but only slightly if at all due to the Y-3 fusion. When the Y fusion presumably broke apart and the X fusion formed, it is possible other rearrangements on the X and/or Y occurred that disrupted pairing. Furthermore, the nondisjunction rate did not decrease to a level similar to the genome strain in vir00-Nofus. We believe this indicates a remaining rearrangement in vir00-Nofus affecting the pairing and disjunction of the X-Y. We only found a modest difference in estimated rDNA copy number between the substrains (Table S7), which has been found to mediate pairing of the X and Y (McKee and Karpen 1990). Detailed analysis of structural rearrangements in the heterochromatin will be required to determine the mechanism of the elevated nondisjunction rates.

Our study has several limitations. First, the vir00-Nofus flies we used for resequencing and for DNA damage assays had the X-4 fusion segregating at low frequency (<15%), which was unknown to us until we were able to correct it for the nondisjunction assays. However, we believe our results hold firmly because the main comparison was with vir00-Yfus, the substrain with the most satellite DNA. For testing our hypothesis, the fusion status in the other substrains matters less than the satellite DNA abundance, which was markedly lower in the vir00-Nofus substrain we used. Furthermore, the increase in damage level between the control and stressed flies may not be directly applicable to the risk of DNA breaks and genome instability in natural conditions. Although we find a difference in the DNA damage in response to stress between vir00 substrains with different abundances of satellite DNA, their fusion status is also different. We cannot eliminate the possibility that the presence of the Y fusion itself increased the rate of

DNA damage instead of the abundance of satellite DNA. Finally, we cannot eliminate the possibility that satellite DNA increased in vir00-Yfus after the fusion occurred and not prior to as we suggest in our model.

We believe the system we discovered will be useful for a variety of future studies. The vir00 fusion substrains will be useful for studying centromere identity. In both the Y-3 and X-4 fusions, two spherical regions of satellite DNA are present at the centromere of these fusions, representing one from each acrocentric chromosome (Figure 1). We note that the X-4 fusion in vir00 is homologous to an independent X-4 fusion in *D. americana* 29 thousand years ago, a species only 4.5 million years diverged. In the X-4 fusion of *D. americana*, there is only one discrete region of centromeric satellite (Flynn *et al.* 2020), unlike what we found here. In female meiosis where chromosomes can compete to get into the oocyte rather than the polar body, "stronger" centromeres with more satellite may have an advantage. A "supercentromere" resulting from a centromere-centromere fusion is one way to do this, and in *D. americana* the X-4 fusion has biased transmission into the egg (Stewart *et al.* 2019). Finally, since both centromere-centromere fusions had matching centromere-proximal satellites (AAACTAT for Y-3 and AAATTAC for X-4), we suggest that the centromeric satellite identity is important for DNA repair for Robertsonian translocations, or for stability and retention of the centromere.

METHODS

Scripts required to reproduce the computational results are available here: https://github.com/jmf422/D-virilis-fusion-chromosomes

Neuroblast squashes and satellite DNA FISH

We dissected brains from wandering 3rd-instar larvae and performed the fixation steps as in (Larracuente and Ferree 2015). Specifically, we placed brains in sodium citrate solution for 6 minutes before fixation. After fixation and drying of slides, we applied Vectashield dapi

mounting medium. We performed DNA-FISH on vir00-Yfus, which allowed us to confidently identify the Y chromosome based on its unique satellite DNA composition. We used the same fixation and staining protocol as Flynn *et al.* (2020). We imaged metaphase cells using a 100x oil objective on an Olympus fluorescent microscope and Metamorph capture system at the Cornell Imaging Facility.

Y-autosome fusion validation

We designed an experiment that would both validate the Y chromosome fusion and to distinguish which autosome is fused. We first designed primers flanking microsatellite loci on all four autosomes that met the following criteria: 1) had 100% conserved non-repetitive and unique priming sites between *D. virilis* and *D. novamexicana*; 2) amplicon length differed between the species by at least 15 bp as to be distinguished on an agarose gel; 3) locus contained a mono or tri nucleotide repeat; 4) locus length ~200 bp. We next set up a two-generation crossing scheme (Figure 1A). We crossed *D. novamexicana* virgin females with vir00-Yfus males and selected the male progeny, which we backcrossed to *D. novamexicana* virgin females. We then genotyped the male F2 progeny from this cross at the 4 sets of primers corresponding to the four non-dot autosomes (Chr2, 3, 4, 5) (Table S4). We performed single-fly DNA extraction in strip tubes with Tris-EDTA buffer and 0.2 mg/mL proteinase K. We did 12 uL standard PCR reactions (3 min at 95, 30 cycles of 30 sec 95, 30 sec 55, 50 sec 72, final extension 5 min). Each primer on each PCR plate had a homozygous (*D. novamexicana*) and heterozygous (*D. novamexicana-D.virilis* F1 hydrid) control. We then ran the PCR product on 2.5% agarose gels. If there was indeed an autosome fused to the Y chromosome, we would expect to see 100% of the male progeny being heterozygous for the *virilis* and *novamexicana* alleles (except for rare cases of non-disjunction). For the autosomes that are not fused, we would expect to see 50% of the progeny being homozygous for the *novamexicana* allele, and half heterozygous, due to Mendel's law of random segregation. We successfully validated the existence of the Y fusion, and found that it is fused to chromosome 3 (Muller D) (Table S1, Figure 1). There were 3 male progeny that were homozygous for the Chr3 *novamexicana* allele. We verified that these

were cases of nondisjunction (opposed to the Y fusion not being fixed in this subline) by finding that the Y chromosome was absent in controlled Y chromosome PCR assays (Figure S3).

Isolation of vir00-Xfus and vir00-Nofus

The X fusion was found to be segregating with a no fusion substrain in the 2019 stock of vir00. We wanted to isolate these into two separate substrains where the karyotype is fixed. From the progeny of the three original crosses in which we found the X fusion, we made 10 single pair crosses and did neuroblast squashes of 6-8 larval progeny per cross, including both sexes. By chance, we should be able to find a cross in which the mother had two copies of the fusion and the father had a single copy – in which the derived line would be fixed for the fusion. If all progeny imaged contained the fusion (and females contained two copies of the fusion), then it is likely that this was the case. We created this line, and call it vir00-Xfus. We maintained a line isolated from the 2019 stock that had no evidence of the X fusion and called it vir00-Nofus. We later found that vir00-Nofus still had the X fusion segregating at low frequency. We isolated a fixed Nofus version in the same way as above for nondisjunction assays, because a low frequency fusion could increase the nondisjunction rate greatly.

X-autosome fusion validation

We obtained transgenic strains with GFP (or Blue) insertions which are expressed in the eye and larval brain from the National Drosophila Species Stock Center (vir95, vir121, vir117). Stern *et al.* (2017) found the insertion sites of these lines. We chose lines which contained the GFP marker on candidate autosomes Chr2, Chr4, and Chr5. Chr3 was not a candidate because it is fused to the Y in vir00-Yfus and contains a different centromeric satellite. Before setting up crosses, we screened 10-20 larvae of each line with a fluorescent microscope to ensure the transgene had not drifted to low frequency. Larvae containing the transgene demonstrated the GFP signal in their brain. We chose to phenotype at the larval stage since we would be crossing GFP strains to wildtype red-eyed flies and the visibility of GFP in the adult eye would be low. We

then designed a crossing scheme which would allow us to both validate that the X chromosome was fused and distinguish which autosome it was fused to (Figure 1B). We crossed GFP-line males to vir00-Xfus virgin females. We then selected the male F1 progeny and crossed them to virilis genome strain virgin females. We then phenotyped F2 larvae, classifying each as either GFP positive or negative. When the phenotyped flies emerged, we sexed and counted them. If the candidate autosome is fused to the X, we would expect sex to segregate with the GFP marker: all female progeny will be GFP negative, and all male progeny will be GFP positive (except for phenotyping errors or rare nondisjunction events). For all other lines, sex should not segregate with GFP status. We performed negative control crosses in which the parental cross was replaced by genome strain virgin females, to ensure the crossing scheme produced the expected results (Table S2). We found that the X chromosome is fused to Chr4 (Muller B).

Validation of the three versions of vir00 with private fixed indels

We used GATK recommended practices to do genotyping of our low-coverage whole genome sequencing data from Flynn et al. (2020). We used vcftools to subset singletons present only in vir00, which was, in hindsight, vir00-Yfus. We then used GATK's SelectVariants to select only non-reference homozygous indels 12 bp or more with a depth of at least 10 in vir00 and at least 2 in the other strains. We then manually inspected each potential candidate in IGV to ensure: no reads in other strains supported the indel, all reads in vir00 supported the indel, and there were no nearby indels in other strains. We then designed primers for the four loci (on Chr 2, 3, 5, 6) that met these criteria and also had enough SNP-free sequence flanking the indel in order to design primers that would amplify a locus 100-200 bp equally in all strains. We performed PCR and gel electrophoresis (2.5% gel, 98 V, 90 min).

DNA damage assays

We chose two stressors that would moderately increase the rate of DNA breaks and allow us to potentially detect differences between strains. Gemcitabine is a nucleoside analog that induces

538 replication stress by stalling polymerases, and also sensitizes cells to radiation via the RAD51
539 pathway (Kobashigawa *et al.* 2015). We selected dosage and a fly-feeding regime based on
540 (Kislukhin *et al.* 2012). Ionizing radiation has long been used to increase the rates of DNA breaks
541 in flies for mutagenesis. We chose a dose ¼ - ½ of what has been typically used in mutagenesis
542 (Carlson and Southin 1962).
543
544 We collected male flies 0-1 days old and fed them gemcitabine (0.718 mM) mixed with liquid
545 food in vials with 8-12 adult flies as in (Kislukhin *et al.* 2012). Liquid food consisted of 12.5g
546 sucrose, 17.5 g dry yeast, 5 mL corn syrup, and 95 mL PBS (autoclaved for 30 min immediately
547 after adding the yeast). Flies were fed the drug for 7-9 days before radiation. Controls were fed
548 with the same liquid food for 7-9 days, except no gemcitabine was added. Flies were moved to
549 fresh vials every 3-4 days. For radiation treatment, we transferred flies into 50 mL conical tubes
550 with 5 mL agar because these tubes were compatible with the radiation source. Control flies
551 were also transferred to new tubes. We used a J.L. Shepherd & Associates Mark I Irradiator with
552 1,100 Ci of Cs-137, and flies were irradiated at approximately 400 rad/min for a total of 10 Gy.
553 In one case, for the GDvir stress treatment, the radiation was not stopped on time so 4 extra Gy
554 were applied. We believe this did not affect our results, especially because the GDvir strain had
555 the lowest DNA damage increase with gemcitabine and radiation stress. We did comet assays
556 to measure DNA damage (Angelis *et al.* 1999) over three different dates (Table S6), but ensured
557 experimental conditions were practically identical each time. For some samples we had to
558 combine results from two different dates to have enough nuclei for statistical analysis (Table
559 S6). We dissected testes from approximately 8 flies from each treatment within one hour of
560 radiation treatment to minimize the opportunity for breakage repair (Shetty *et al.* 2017). We
561 then homogenized the testes tissue using a dounce, filtered the homogenate through a 40
562 micron filter to remove debris, and centrifuged and resuspended the cells to approximately 10^5
563 cells/mL.
564
565 We next performed the alkaline comet assay as directed by the Enzo comet kit (ADI-900-166),
566 which provides higher sensitivity than the neutral comet assay (Angelis *et al.* 1999). We imaged

slides on a metamorph imaging system at 10x magnification using a fluorescent green filter to detect the CyGreen dye included in the comet kit. We quantified damage levels using the software OpenComet as a plugin in ImageJ (Gyori *et al.* 2014). We filtered called nuclei that were not comet shapes or contained background interference. We used the measure of "olive moment," which is the product of the percent of DNA in the tail and distance between intensity-weighted centroids of head and tail (Gyori *et al.* 2014) as the statistic to compare between strains and treatments.

Resequencing sublines to determine differences in their satellite abundance.

Pools of 6 male flies were DNA extracted with Qiagen DNeasy blood and tissue kit. PCR-free libraries were then prepared with Illumina TruSeq PCR-free library prep. Libraries were sequenced on a NextSeq 500 single end 150 bp. We removed adapters and poly-G signal with fastp and then ran k-Seek to count satellite abundances (Wei *et al.* 2014). We used average read depth to normalize the kmer counts. We also mapped the reads to the *D. virilis* rDNA consensus sequence (http://blogs.rochester.edu/EickbushLab/?page_id=602) to estimate the rDNA copy number in the three vir00 substrains as well as a vir08 as a control (Table S3, S7).

Using sequencing data to estimate the age of the Y-3 fusion

Scripts for this section are available here: https://github.com/jmf422/D-virilis-fusion-chromosomes/tree/main/simulate_degradation. We used the sequencing data from Flynn et al. (2020) in addition to data produced here for vir00-Yfus, vir00-Xfus, and vir00-Nofus and 10 other *D. virilis* strains. We mapped the data to the RS2 genome assembly using bowtie2. We then genotyped with GATK following standard procedures (McKenna *et al.* 2010). We extracted heterozygous singleton sites for vir00, and counted how many occurred on each autosome. We calculated the enrichment on Chr3 in vir00-Yfus based on the difference from the average SNP density on the other autosomes (excluding the dot chromosome Chr6). To determine whether this enrichment of SNPs was significant based on the size of the chromosome and the number

of mutations, we randomly permuted the total number of heterozygous singleton SNPs on all autosomes and calculated the proportion falling on Chr3, and repeated this 1000 times.

We then performed simple simulations to determine approximately how many generations of mutation accumulation without recombination or selection would result in the enrichment we observed. Since heterozygous singletons are challenging for the genotyper to call with moderate coverage sequencing data, we incorporated this into our simulation. First, we made the genome assembly diploid then used mutation-simulator (Kühl *et al.* 2020) to simulate random mutations (transition/transversion ratio 2.0) at a rate of 2×10^{-9} per bp per generation for 500, 1000, 2000, and 5000 generations on one copy of Chr3 only. We then simulated Illumina reads with ART (Huang *et al.* 2012) at the same depth as we have for vir00-Yfus in our real data (23 x haploid or 11.5 x diploid). We next used standard GATK genotyping and selected out heterozygous singletons on Chr3. We repeated the simulation 10 times for each number of generations to get a range of values. The empirical enrichment fell in between what we found in the simulations for 1000 and 2000 generations.

Nondisjunction assays

We crossed a single male from vir00-Yfus, vir00-Xfus, vir00-Nofus (fixed), and GDvir (control) to one or two GDvir (genome strain vir87) females. We collected the virgin progeny from each cross, extracted DNA with a squish-proteinase K prep, and genotyped with PCR and gel electrophoresis for the presence or absence of the Y chromosome in up to 16 female and 16 male progeny. We amplified a locus unique to the Y chromosome (primers designed by Yasir Ahmed-Braimah for a different project, Table S4). For a subset of individuals, we also performed multiplex controls with an autosomal locus. Otherwise, we performed DNA extractions in large batches with the same proteinase K mixture to minimize the chance of DNA extraction failure. A very small quantity of DNA is required for a standard PCR with robust primers. To control for the completeness of the PCR mastermix, we included male samples in the same batch as female samples. If a male lacked a Y chromosome, we inferred the father's

sperm was missing the Y chromosome (nullisomic), and if a female contained a Y chromosome, we inferred the father's sperm contained both X and Y. We used R prop.test to evaluate whether there were any differences between nondisjunction proportions for the different strains. After finding this highly significant, we used pairwise.prop.test in R with Holm-Bonferroni multiple test correction to determine which pairs of substrain nondisjunction rates were significantly different from each other.

Acknowledgements

We thank Yasir Ahmed-Braimah for discussions and use of primers. We greatly appreciate the use of the Cs-137 irradiator from the Robert Weiss lab and to Amanda Loehr for operating the device. Asha Jain prepared sequencing libraries for this project, and Yassi Hafezi provided advice on the nondisjunction assays. We also thank members of the Clark lab for discussions and encouragement on this project.

References

Anderson S. N., M. C. Stitzer, A. B. Brohammer, P. Zhou, J. M. Noshay, *et al.*, 2019 Transposable elements contribute to dynamic genome content in maize. Plant J. 100: 1052–1065.

Anderson N. W., C. E. Hjelmen, and H. Blackmon, 2020 The probability of fusions joining sex chromosomes and autosomes. Biol. Lett. 16: 20200648.

Angelis K. J., M. Dusinská, and A. R. Collins, 1999 Single cell gel electrophoresis: detection of DNA damage at different levels of sensitivity. Electrophoresis 20: 2133–2138.

Arora U. P., C. Charlebois, R. A. Lawal, and B. L. Dumont, 2021 Population and subspecies diversity at mouse centromere satellites. BMC Genomics 22: 279.

Bachtrog D., 2013 Y-chromosome evolution: emerging insights into processes of Y-chromosome

degeneration. Nat. Rev. Genet. 14: 113–124.

Balzano E., F. Pelliccia, and S. Giunta, 2020 Genome (in)stability at tandem repeats. Semin. Cell Dev. Biol. https://doi.org/10.1016/j.semcdb.2020.10.003

Barra V., and D. Fachinetti, 2018 The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. Nat. Commun. 9: 4340.

Bilinski P., P. S. Albert, J. J. Berg, J. A. Birchler, M. N. Grote, *et al.*, 2018 Parallel altitudinal clines reveal trends in adaptive evolution of genome size in Zea mays. PLoS Genet. 14: e1007162.

Black E. M., and S. Giunta, 2018 Repetitive Fragile Sites: Centromere Satellite DNA As a Source of Genome Instability in Human Diseases. Genes 9. https://doi.org/10.3390/genes9120615

Braekeleer M. D., and T.-N. Dao, 1990 Cytogenetic studies in couples experiencing repeated pregnancy losses. Hum. Reprod. 5: 519–528.

Brown E. J., A. H. Nguyen, and D. Bachtrog, 2020a The Drosophila Y Chromosome Affects Heterochromatin Integrity Genome-Wide. Mol. Biol. Evol. 37: 2808–2824.

Brown E. J., A. H. Nguyen, and D. Bachtrog, 2020b The Y chromosome may contribute to sex-specific ageing in Drosophila. Nat Ecol Evol 4: 853–862.

Carlson E. A., and J. L. Southin, 1962 Comparative mutagenesis of the dumpy locus in Drosophila melanogaster. I. X-ray treatment of mature sperm--frequency and distribution. Genetics 47: 321–336.

668 Cechova M., R. S. Harris, M. Tomaszkiewicz, B. Arbeithuber, F. Chiaromonte, *et al.*, 2019 High
669 satellite repeat turnover in great apes studied with short- and long-read technologies. Mol.
670 Biol. Evol. https://doi.org/10.1093/molbev/msz156

671 Charlesworth B., C. H. Langley, and W. Stephan, 1986 The evolution of restricted recombination
672 and the accumulation of repeated DNA sequences. Genetics 112: 947–962.

673 Charlesworth B., and D. Charlesworth, 2000 The degeneration of Y chromosomes. Philosophical
674 Transactions of the Royal Society of London. Series B: Biological Sciences 355: 1563–1572.

675 Featherstone C., and S. P. Jackson, 1999 DNA double-strand break repair. Curr. Biol. 9: R759–
676 R761.

677 Ferree P. M., and D. A. Barbash, 2009 Species-specific heterochromatin prevents mitotic
678 chromosome segregation to cause hybrid lethality in Drosophila. PLoS Biol. 7: e1000234.

679 Flynn J. M., I. Caldas, M. E. Cristescu, and A. G. Clark, 2017 Selection Constrains High Rates of
680 Tandem Repetitive DNA Mutation in Daphnia pulex. Genetics 207: 697–710.

681 Flynn J. M., M. Long, R. A. Wing, and A. G. Clark, 2020 Evolutionary Dynamics of Abundant 7-bp
682 Satellites in the Genome of Drosophila virilis. Mol. Biol. Evol. 37: 1362–1375.

683 Francisco F. O., and B. Lemos, 2014 How Do Y-Chromosomes Modulate Genome-Wide
684 Epigenetic States: Genome Folding, Chromatin Sinks, and Gene Expression. Journal of
685 Genomics 2: 94–103.

686 Fry K., and W. Salser, 1977 Nucleotide sequences of HS-alpha satellite DNA from kangaroo rat

687 Dipodomys ordii and characterization of similar sequences in other rodents. Cell 12: 1069–
688 1084.

689 Gall J., E. Cohen, and M. Polan, 1971 Repetitive DNA sequences in Drosophila. Chromosoma 33.
690 https://doi.org/10.1007/bf00284948

691 Gall J. G., and D. D. Atherton, 1974 Satellite DNA sequences in Drosophila virilis. J. Mol. Biol. 85:
692 633–664.

693 Giunta S., S. Hervé, R. R. White, T. Wilhelm, M. Dumont, *et al.*, 2021 CENP-A chromatin prevents
694 replication stress at centromeres to avoid structural aneuploidy. Proc. Natl. Acad. Sci. U. S.
695 A. 118. https://doi.org/10.1073/pnas.2015634118

696 Gyori B. M., G. Venkatachalam, P. S. Thiagarajan, D. Hsu, and M.-V. Clement, 2014 OpenComet:
697 an automated tool for comet assay image analysis. Redox Biol 2: 457–465.

698 Huang W., L. Li, J. R. Myers, and G. T. Marth, 2012 ART: a next-generation sequencing read
699 simulator. Bioinformatics 28: 593–594.

700 Jagannathan M., R. Cummings, and Y. M. Yamashita, 2018 A conserved function for
701 pericentromeric satellite DNA. Elife 7. https://doi.org/10.7554/eLife.34122

702 Jagannathan M., and Y. M. Yamashita, 2021 Defective satellite DNA clustering into
703 chromocenters underlies hybrid incompatibility in Drosophila. bioRxiv 2021.04.16.440167.

704 Kislukhin G., M. L. Murphy, M. Jafari, and A. D. Long, 2012 Chemotherapy-induced toxicity is
705 highly heritable in Drosophila melanogaster. Pharmacogenet. Genomics 22: 285–289.

706 Kobashigawa S., K. Morikawa, H. Mori, and G. Kashino, 2015 Gemcitabine Induces
707 Radiosensitization Through Inhibition of RAD51-dependent Repair for DNA Double-strand
708 Breaks. Anticancer Res. 35: 2731–2737.

709 Kühl M. A., B. Stich, and D. C. Ries, 2020 Mutation-Simulator: Fine-grained simulation of
710 random mutations in any genome. Bioinformatics.
711 https://doi.org/10.1093/bioinformatics/btaa716

712 Larracuente A. M., and P. M. Ferree, 2015 Simple method for fluorescence DNA in situ
713 hybridization to squashed chromosomes. J. Vis. Exp. 52288.

714 Lindsley D. L., L. Sandler, B. S. Baker, A. T. Carpenter, R. E. Denell, *et al.*, 1972 Segmental
715 aneuploidy and the genetic gross structure of the Drosophila genome. Genetics 71: 157–
716 184.

717 Maggert K. A., 2014 Reduced rDNA Copy Number Does Not Affect "Competitive" Chromosome
718 Pairing in XYY Males of Drosophila melanogaster. G3 Genes|Genomes|Genetics 4: 497–
719 507.

720 Mayrose I., and M. A. Lysak, 2021 The Evolution of Chromosome Numbers: Mechanistic Models
721 and Experimental Approaches. Genome Biol. Evol. 13.
722 https://doi.org/10.1093/gbe/evaa220

723 McKee B. D., and G. H. Karpen, 1990 Drosophila ribosomal RNA genes function as an X-Y pairing
724 site during male meiosis. Cell 61: 61–72.

McKenna A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, *et al.*, 2010 The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20: 1297–1303.

Miga K. H., Y. Newton, M. Jain, N. Altemose, H. F. Willard, *et al.*, 2014 Centromere reference models for human chromosomes X and Y satellite arrays. Genome Res. 24: 697–707.

Mills W. K., Y. C. G. Lee, A. M. Kochendoerfer, E. M. Dunleavy, and G. H. Karpen, 2019 RNA from a simple-tandem repeat is required for sperm maturation and male fertility in Drosophila melanogaster. Elife 8. https://doi.org/10.7554/eLife.48940

Nozawa M., Y. Minakuchi, K. Satomura, S. Kondo, A. Toyoda, *et al.*, 2021 Evolutionary trajectories of three independent neo-sex chromosomes in Drosophila. Cold Spring Harbor Laboratory 2021.03.11.435033.

Paço A., F. Adega, N. Meštrović, M. Plohl, and R. Chaves, 2015 The puzzling character of repetitive DNA in Phodopus genomes (Cricetidae, Rodentia). Chromosome Res. 23: 427–440.

Petitpierre E., C. Juan, J. Pons, M. Plohl, and D. Ugarkovic, 1995 Satellite DNA and constitutive heterochromatin in tenebrionid beetles, pp. 351–362 in *Kew Chromosome Conference IV. London: Royal Botanic Gardens*,.

Reis M., C. P. Vieira, R. Lata, N. Posnien, and J. Vieira, 2018 Origin and Consequences of Chromosomal Inversions in the virilis Group of Drosophila. Genome Biol. Evol. 10: 3152–3166.

745 Schulz R., L. A. Underkoffler, J. N. Collins, and R. J. Oakey, 2006 Nondisjunction and transmission
746 ratio distortion ofChromosome 2 in a (2.8) Robertsonian translocation mouse strain.
747 Mamm. Genome 17: 239–247.

748 Shetty V., N. J. Shetty, S. R. Ananthanarayana, S. K. Jha, and R. C. Chaubey, 2017 Evaluation of
749 gamma radiation-induced DNA damage in Aedes aegypti using the comet assay. Toxicol.
750 Ind. Health 33: 930–937.

751 Stern D. L., J. Crocker, Y. Ding, N. Frankel, G. Kappes, *et al.*, 2017 Genetic and Transgenic
752 Reagents for Drosophila

753 Stewart N. B., Y. H. Ahmed-Braimah, D. G. Cerne, and B. F. McAllister, 2019 Female meiotic
754 drive preferentially segregates derived metacentric chromosomes in Drosophila. bioRxiv
755 638684.

756 Subirana J. A., M. M. Albà, and X. Messeguer, 2015 High evolutionary turnover of satellite
757 families in Caenorhabditis. BMC Evol. Biol. 15: 218.

758 Thakur J., J. Packiaraj, and S. Henikoff, 2021 Sequence, Chromatin and Evolution of Satellite
759 DNA. Int. J. Mol. Sci. 22. https://doi.org/10.3390/ijms22094309

760 Vieira C. P., A. Almeida, J. D. Dias, and J. Vieira, 2006 On the location of the gene(s) harbouring
761 the advantageous variant that maintains the X/4 fusion of Drosophila americana. Genet.
762 Res. 87: 163–174.

763 Wei K. H.-C., J. K. Grenier, D. A. Barbash, and A. G. Clark, 2014 Correlated variation and

population differentiation in satellite DNA abundance among lines of Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. 111: 18793–18798.

Wei K. H.-C., S. E. Lower, I. V. Caldas, T. J. S. Sless, D. A. Barbash, *et al.*, 2018 Variable Rates of Simple Satellite Gains across the Drosophila Phylogeny. Mol. Biol. Evol. 35: 925–941.

Wei K. H.-C., and D. Bachtrog, Male recombination produced multiple geographically restricted neo-Y chromosome haplotypes of varying ages that correlate with onset of neo-Y decay in Drosophila albomicans