

Comparative Analysis of Satellite DNA in the *Drosophila melanogaster* Species Complex

Madhav Jagannathan,^{*,†,1} Natalie Warsinger-Pepe,^{†,*,1} George J. Watase,^{*,†,1} and

Yukiko M. Yamashita^{*,†,§,2}

^{*}Howard Hughes Medical Institute, [†]Life Sciences Institute, [‡]Department of Molecular and Integrative Physiology, and [§]Department of Cell and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109

ORCID IDs: 0000-0003-3428-6812 (M.J.); 0000-0002-9375-8990 (N.W.-P.); 0000-0001-8250-9027 (G.J.W.); 0000-0001-5541-0216 (Y.M.Y.)

ABSTRACT Satellite DNAs are highly repetitive sequences that account for the majority of constitutive heterochromatin in many eukaryotic genomes. It is widely recognized that sequences and locations of satellite DNAs are highly divergent even in closely related species, contributing to the hypothesis that satellite DNA differences may underlie speciation. However, due to its repetitive nature, the mapping of satellite DNAs has been mostly left out of recent genomics analyses, hampering the use of molecular genetics techniques to better understand their role in speciation and evolution. Satellite DNAs are most extensively and comprehensively mapped in *Drosophila melanogaster*, a species that is also an excellent model system with which to study speciation. Yet the lack of comprehensive knowledge regarding satellite DNA identity and location in its sibling species (*D. simulans*, *D. mauritiana*, and *D. sechellia*) has prevented the full utilization of *D. melanogaster* in studying speciation. To overcome this problem, we initiated the mapping of satellite DNAs on the genomes of the *D. melanogaster* species complex (*D. melanogaster*, *D. simulans*, *D. mauritiana*, and *D. sechellia*) using multi-color fluorescent *in situ* hybridization (FISH) probes. Our study confirms a striking divergence of satellite DNAs in the *D. melanogaster* species complex, even among the closely related species of the *D. simulans* clade (*D. simulans*, *D. mauritiana*, and *D. sechellia*), and suggests the presence of unidentified satellite sequences in these species.

KEYWORDS

Drosophila
evolution
fluorescent *in situ*
hybridization
heterochromatin
satellite DNA

Short tandem repetitive or “satellite” DNAs are abundant and conserved features of eukaryotic genomes. Although typically considered “junk” DNA due to a lack of protein coding potential, decades of study have implicated satellite DNA function in cellular processes such as kinetochore/centromere function, meiotic chromosome segregation, and X chromosome recognition (Dernburg *et al.* 1996; Karpen *et al.* 1996; Guenatri *et al.* 2004; Bouzinba-Segard *et al.* 2006; Usakin *et al.* 2007; Wong *et al.* 2007; Menon *et al.* 2014; Rosic *et al.* 2014). Moreover, aberrant transcription of satellite DNAs has been associated with

human diseases such as cardiomyopathy and cancer, suggesting critical importance of the regulation of this underappreciated component of eukaryotic genomes (Gaubatz and Cutler 1990; Feber *et al.* 2011; Ting *et al.* 2011; Haider *et al.* 2012). Yet, other than these examples, the functions of the majority of satellite DNAs remain obscure.

The common fruit fly, *Drosophila melanogaster*, is an excellent model system with which to study satellite DNAs. Approximately 21% of the *D. melanogaster* genome is comprised of satellite DNA (Lohe and Brutlag 1986) and extensive efforts have mapped the location of 15 unique repeats on *D. melanogaster* chromosomes (Waring and Pollack 1987; Bonaccorsi and Lohe 1991; Abad *et al.* 1992; Lohe *et al.* 1993; Dernburg *et al.* 1996). Efforts to identify satellite repeats and map them onto chromosomes have been made in many species including *Abracris flavolineata* (Bueno *et al.* 2013), *Aegilops geniculata* and wheat (Koo *et al.* 2016), various *Arabidopsis* species (Kamm *et al.* 1995; Ito *et al.* 2007; Kawabe and Charlesworth 2007), maize (Lamb *et al.* 2007), *Turritis glabra* (Kawabe and Nasuda 2006), rodent species including *Phodopus sungorus* (Paco *et al.* 2014), *Cricetus cricetus* and *Microtus arvalis* (Louzada *et al.* 2015), *Xenopus* (Schmid and Steinlein 2015), and human (Altemose *et al.* 2014), revealing general patterns of

Copyright © 2017 Jagannathan *et al.*

doi: 10.1534/g3.116.035352

Manuscript received September 7, 2016; accepted for publication December 18, 2016; published Early Online December 22, 2016.

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

¹These authors contributed equally to this work.

²Corresponding author: Life Sciences Institute, University of Michigan, 210 Washtenaw Ave., Rm 5403, Ann Arbor, MI 48109. E-mail: yukikomy@umich.edu

■ Table 1 Combination of probes used for FISH mapping

Combination	488 (Green)	Red (Cy3)	Blue (Cy5)
C1	(AATAT) ₆	(AATAG) ₆	359
C2	(AATAGAC) ₆	(AATAC) ₆	(AATAAAC) ₆
C3	IGS	(AAGAG) ₆	Prod: (AATAACATAG) ₃
C4	IGS	372	359
C5	(AAGAG) ₆	Dodeca	(AACAC) ₆
C6	(AATAT) ₆	(GAACAGAACATGTTC) ₂	(AACAAAC) ₅
C7	(AATAT) ₆	(AAGAC) ₆	(AAAAC) ₆
C8	(AAGAG) ₆	(AAGAC) ₆	(AAGAGAG) ₅
C9	(AATAT) ₆	(GAACAGAACATGTTC) ₂	359
C10	(AATAT) ₆	(GAACAGAACATGTTC) ₂	Dodeca
C11	(AATAT) ₆	(TAGA) ₈	260

See Materials and Methods for the sequence details of the probes for 359, 372, 260, IGS, and dodeca.

centromeric, pericentromeric, and telomeric satellite distribution. The rich history of *D. melanogaster* genetics has resulted in the comprehensive identification and mapping of satellite DNA to individual chromosomes; *D. melanogaster* remains the only species with this resolution (Waring and Pollack 1987; Bonaccorsi and Lohe 1991; Abad *et al.* 1992; Lohe *et al.* 1993; Dernburg *et al.* 1996). Even in *Drosophila* sibling species such as *D. simulans*, *D. mauritiana*, and *D. sechellia* (together called the *D. simulans* clade), satellite composition and chromosome location has only been partially examined (Lohe and Brutlag 1987; Larracuento 2014).

Interestingly, it has been shown that even closely related species display significant divergence in the abundance and sequence of individual satellite DNA repeats (Lohe and Brutlag 1987; Lohe and Roberts 2000; Bosco *et al.* 2007). These observations led to the hypothesis that rapid divergence of satellite DNA may play an important role in speciation by causing reproductive isolation between closely related species (Yunis and Yasminah 1971; Gatti *et al.* 1976). In support of this idea, it was shown that a satellite DNA on the *D. melanogaster* X chromosome (*i.e.*, 359 bp repeats) causes hybrid incompatibility when crossed with its closest sibling species, *D. simulans* (Sawamura *et al.* 1993; Ferree and Barbash 2009). However, a lack of information regarding satellite DNAs in other species hinders efforts to identify whether there are more instances of hybrid incompatibility caused by satellite DNA among closely related species.

In this study, we have used FISH to map known *Drosophila* satellite DNA repeats on the mitotic chromosomes of the sibling species *D. melanogaster*, *D. simulans*, *D. mauritiana*, and *D. sechellia* (collectively categorized as “the *D. melanogaster* species complex”). We reveal a remarkable divergence in the abundance and location of specific satellite DNA repeats in these closely related sibling species, and provide this information as a resource for future work on *Drosophila* chromosome biology and speciation.

MATERIALS AND METHODS

Drosophila strains and fly husbandry

All fly stocks were raised on standard Bloomington medium at 25°, and male third instar wandering larvae were used. For better chromosome squash, larvae cultured at 18° were used. The following fly stocks were used: *D. melanogaster* yw, *D. simulans* w⁵⁰¹ (DSSC#14021-0251.195), *D. sechellia* w¹ (DSSC#14021-0248.30), and *D. mauritiana* w¹ (DSSC#14021-0241.60).

Larval brain squash, chromosome FISH, and microscopy

We adapted a simple FISH protocol against squashed chromosomes published by Larracuento and Ferree (2015) with small modifications.

Briefly, male third instar wandering larvae were collected and brains were dissected in PBS. Larval brains were fixed in 25 µl of acetic acid: 4% formaldehyde in PBS (45%:55%) for 4 min on Sigmacote-coated coverslips (Sigma: SL2 SIGMA). The whole sample was quickly applied to a clean Superfrost plus slide and the sample was manually squashed via thumb/stamp over coverslip, over sample, on top of the slide. The slide/coverslip/sample was immediately submerged in liquid nitrogen until it stopped boiling. Slides were removed from liquid nitrogen and coverslips were quickly flicked off the slide with a razor blade. Slides were then washed in 100% ethanol at room temperature for 5 min then dried in a dark, dust-free location.

Hybridization was performed in 50% formamide, 10% dextran sulfate, 2 × SSC buffer, 0.5 µM of each probe, and up to 20 µl of diH₂O. Hybridization buffer was added to the samples and covered with a coverslip. Slides were heated at 95° for 5 min, cooled briefly, wrapped in parafilm, and incubated in a humid chamber in the dark overnight at room temperature. Coverslips were removed and slides were washed three times for 15 min in 0.1 × SSC, removed of excess buffer, and mounted in Vectashield mounting medium containing DAPI.

Images were taken using an upright Leica TCS SP8 confocal microscope with a 63 × oil immersion objective (NA = 1.4) and processed using Adobe Photoshop software. Images were modified solely for the purpose of clarity. Modified images were not quantified.

Detailed sequences of probes used that are not provided in Table 1 are as follows: 359 (part of 359 bp unit, antisense direction), 5'-AGG-ATTTAGGGAAATTAATTTTTGGATCAATTTTCGCATTTTTTG-TAAG-3'; 372 (part of 372 bp unit), 5'-TATTTTGATCAAAACATT-GAAAATAATGGCCAAATATGGAATGTCATACCTCGTTGAG-TTTGTTTTTTA-3'; IGS (part of 240 bp unit), 5'-AGTGAAAAATG-TTGAAATATTCCCATATTCTCTAAGTATTATAGAGAAAAGC-CATTTTAGTGAATGGA-3'; dodeca, 5'-ACCGAGTACGGGACC-GAGTACGGGACCAGTACGGGACCAGTACGGG-3'; 260 (part of 260 bp unit), 5'-CATATTTGCAAATTTTGATGAATGCGAAAAT-TAAC-3'.

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article.

RESULTS

Based on previous studies (Simeone *et al.* 1985; Lohe and Brutlag 1987; Waring and Pollack 1987; Lohe and Roberts 1990; Abad *et al.* 1992, 2000; Carmena *et al.* 1993; Lohe *et al.* 1993; Hoskins *et al.* 2015), we designed 17 probes that are known to exist in *D. melanogaster* (some of which are also known to exist in *D. simulans*), and two probes that are

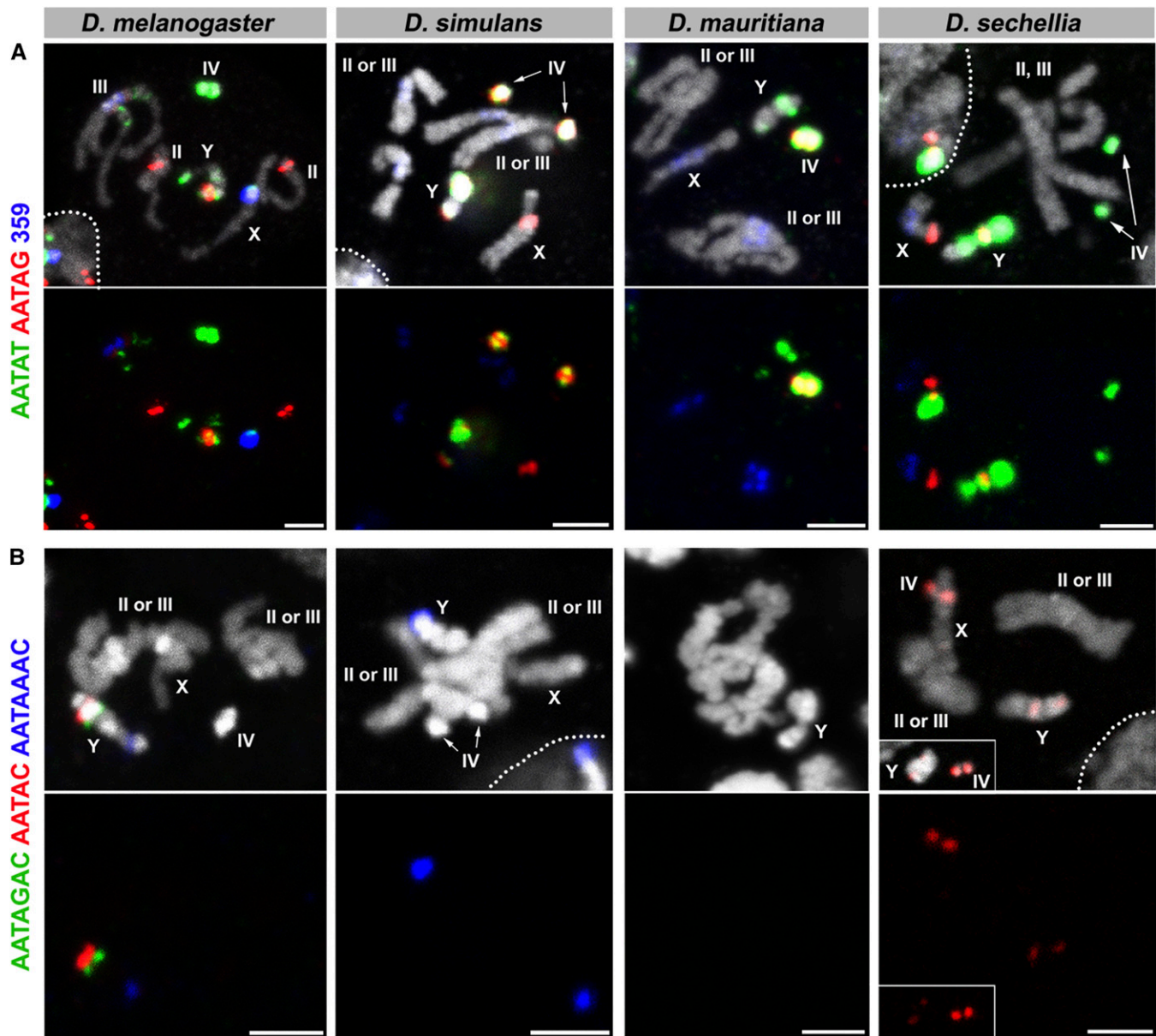


Figure 1 FISH on neuroblast chromosome spread from *D. melanogaster*, *D. simulans*, *D. mauritiana*, and *D. sechellia*. (A) C1 probes and (B) C2 probes. Probe sequences are indicated by the colored text. The top panels show three probes combined with DAPI, and the bottom panels show only probe hybridization signals. Bar, 2.5 μm . C, combination; DAPI, 4',6-diamidino-2-phenylindole; FISH, fluorescent *in situ* hybridization.

known to exist in *D. simulans* but are supposedly absent in *D. melanogaster* (Table 1). We used these 19 probes in 11 different combinations (C1–11, Table 1), with each combination consisting of Alexa488-, Cy3-, and Cy5-conjugated probes. We then performed FISH on the mitotic chromosomes of squashed larval brains from *D. melanogaster*, *D. simulans*, *D. mauritiana*, and *D. sechellia* using a method described previously (Larracunte and Ferree 2015). We selected male third instar larvae based on their gonad morphology such that the mitotic chromosome spread would always contain a Y chromosome.

The results are shown in Figure 1, Figure 2, Figure 3, Figure 4, and Figure 5, and the chromosomal location(s) of each satellite DNA repeat in all four species are summarized in Table 2 and Table 3. We observed a dramatic divergence of satellite repeat abundance and distribution

that clearly demarcates the chromosomes of these closely related species. It should be noted that the probe used for 372 bp repeats perfectly colocalized with the 359 bp probe on the mitotic chromosomes in all four species, suggesting that these two probes do not differentiate between the 359 bp satellite and the 372 bp repeat sequences, as well as other derivatives such as 353 and 356 bp repeats in *D. melanogaster* (Losada and Villasante 1996; Abad *et al.* 2000), and 360 family sequences in the *D. simulans* clade (Strachan *et al.* 1985). Indeed, our 70-bp-long probe designed for 372 bp repeats shows 74% identity to the 359 bp satellite consensus sequence. One puzzling observation is that the 372 bp repeats were originally identified as middle repetitive elements dispersed along the X chromosome euchromatin in *D. melanogaster* (Waring and Pollack 1987). Later, it was found that the 372 bp repeats are a part of the 1.688 g/cm³ class of satellite DNAs, which is

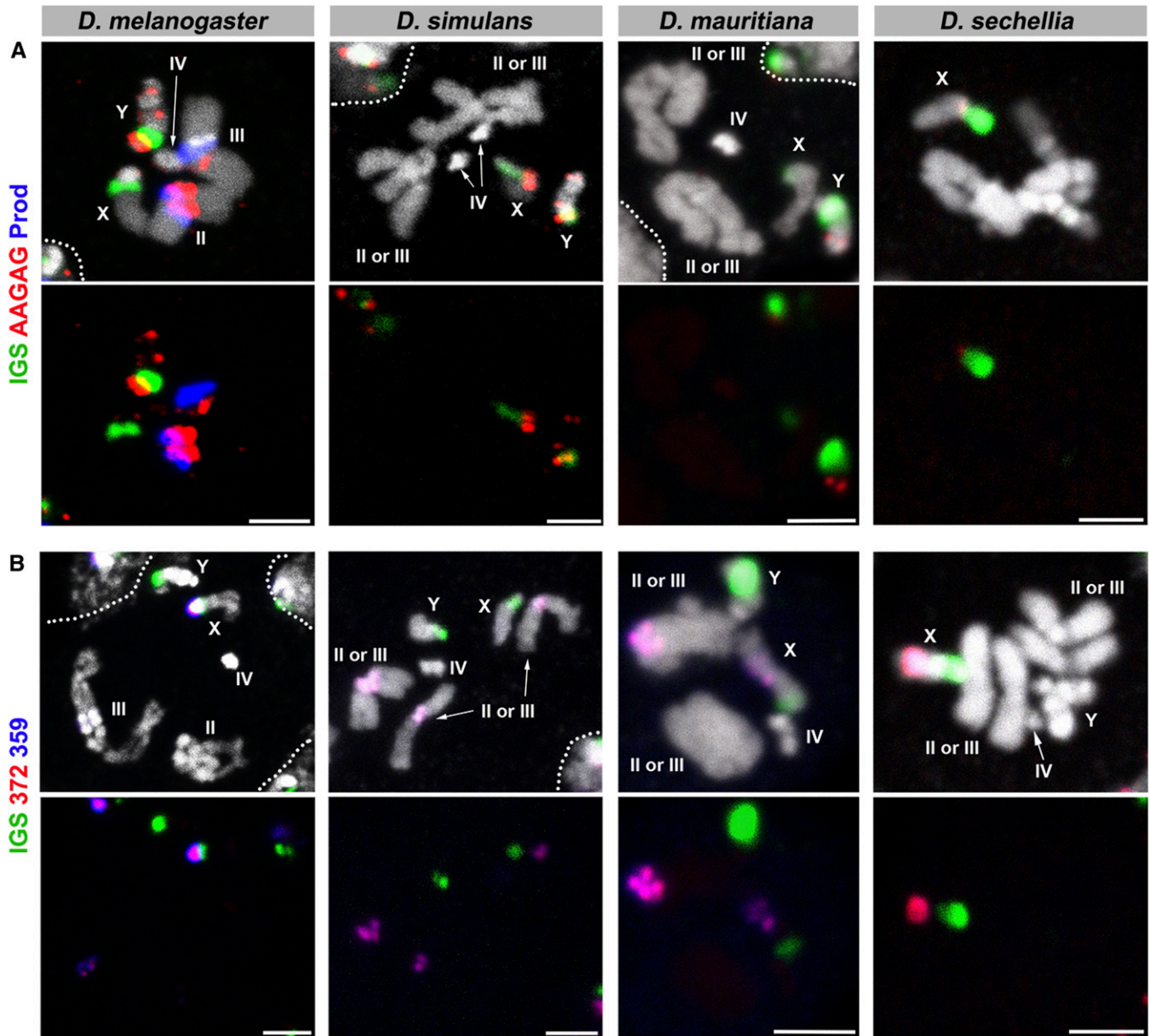


Figure 2 FISH on neuroblast chromosome spread from *D. melanogaster*, *D. simulans*, *D. mauritiana*, and *D. sechellia*. (A) C3 probes and (B) C4 probes. Probe sequences are indicated by the colored text. The top panels show three probes combined with DAPI, and the bottom panels show only probe hybridization signals. Bar, 2.5 μm . C, combination; DAPI, 4',6-diamidino-2-phenylindole; FISH, fluorescent *in situ* hybridization.

present at numerous locations throughout the X chromosome euchromatin (DiBartolomeis *et al.* 1992). Our mitotic chromosome FISH did not detect hybridization of 372 or 359 bp probes to these euchromatic regions. This might be due to the fact that euchromatic loci of the 1.688 g/cm^3 class of satellite DNAs (including 372 bp repeats) are below the detection limit of our FISH technique (DiBartolomeis *et al.* 1992; Kuhn *et al.* 2012). In addition, the detection of such low abundance sequences in euchromatic regions might require polytenization, where the euchromatic 372 bp repeat was originally identified (Waring and Pollack 1987). We additionally designed a 260 bp probe, as neither of our 359 or 372 bp probes was expected to hybridize to 260 bp repeats. This probe still cross-hybridized to 359/372 bp repeats, but successfully detected an additional locus (Figure 5), likely representing *bona fide* 260 bp repeats on the *D. melanogaster* second chromosome as reported previously

(albeit at a low signal intensity) (Abad *et al.* 2000). In *D. mauritiana*, the 260 bp probe hybridized to the same chromosomal location as 359/372 bp repeats, suggesting this might also represent cross-hybridization. However, in *D. simulans* and *D. sechellia*, the 260 bp repeat probe did not show any signal, suggesting either that (1) *D. simulans* and *D. sechellia* 359/372 bp repeats are more diverged in the region that is covered by our 260 bp repeat probe or that (2) cross-hybridization of the 260 bp probe to 359/372 bp repeats is very weak and slight differences in hybridization conditions yield different results.

While two *D. simulans*-specific satellites (AACAAAC and GAA-CAGAACATGTTTC) had been previously identified by cesium chloride centrifugation (Lohe and Brutlag 1987), their chromosomal locations were unknown. Here, we map their locations on the *D. simulans* chromosomes (Figure 3, Table 2, and Table 3) and extend this

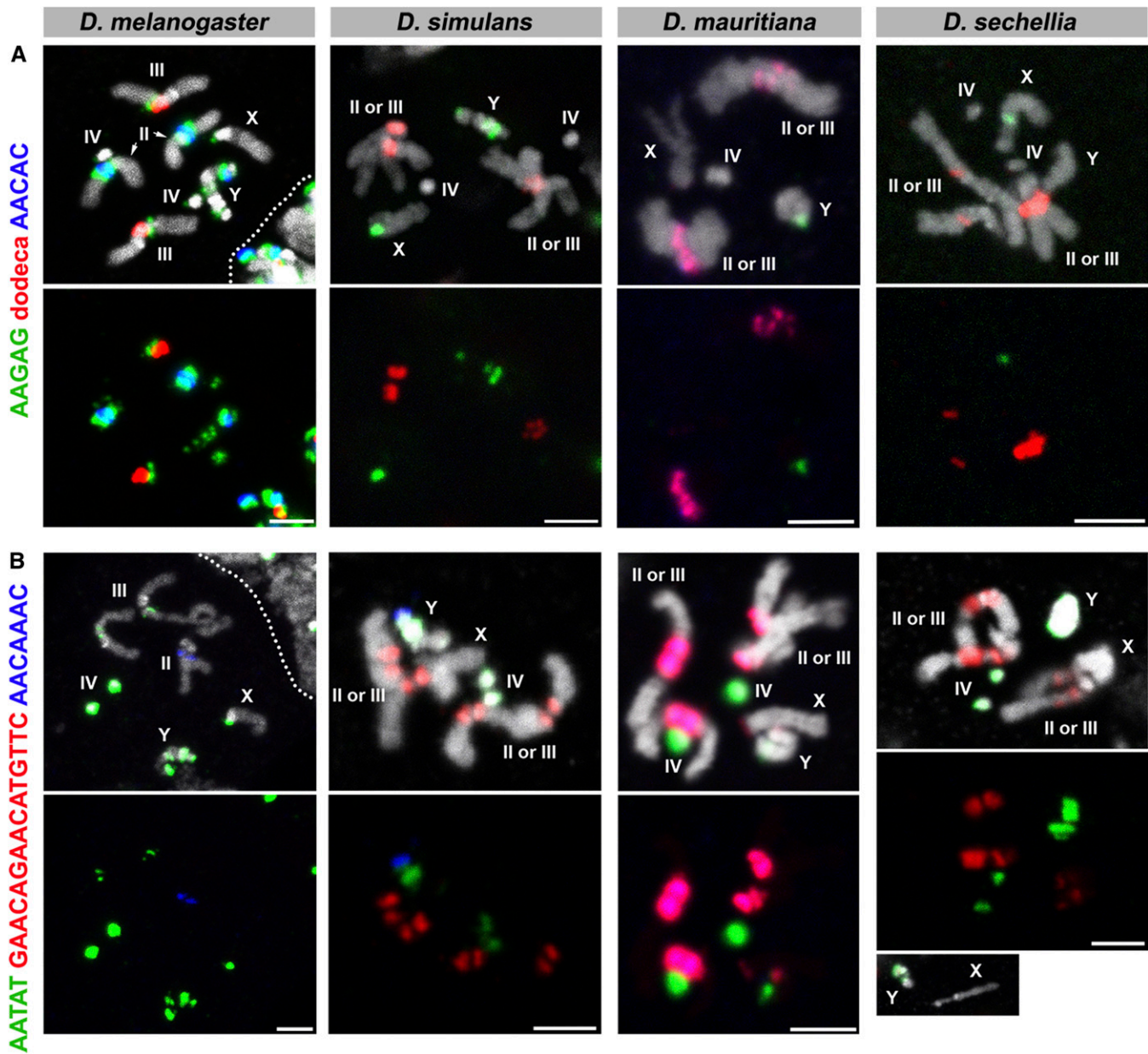


Figure 3 FISH on neuroblast chromosome spread from *D. melanogaster*, *D. simulans*, *D. mauritiana*, and *D. sechellia*. (A) C5 probes and (B) C6 probes. Probe sequences are indicated by the colored text. The top panels show three probes combined with DAPI, and the bottom panels show only probe hybridization signals. Bar, 2.5 μ m. C, combination; DAPI, 4',6'-diamidino-2-phenylindole; FISH, fluorescent *in situ* hybridization.

mapping to both *D. mauritiana* and *D. sechellia*. The GAACAGAA-CATGTT repeat is completely absent in *D. melanogaster* and is predominantly present on the autosomes of *D. simulans*, *D. mauritiana*, and *D. sechellia*. On the other hand, the AACAAAC satellite repeat is specifically present on the *D. simulans* Y chromosome and is absent from *D. mauritiana* and *D. sechellia*. Although AACAAAC was reported to be absent in *D. melanogaster*, we observed probe hybridization on the *D. melanogaster* second chromosome, albeit at a low level. It is possible that a slightly diverged sequence on the *D. melanogaster* chromosome is still able to hybridize to the AACAAAC probe. Alternatively, the *D. melanogaster* genome could have acquired/amplified this satellite repeat in the time between the previous report and this study (~30 yr).

An important aspect of the satellite DNA maps in *D. melanogaster* is the knowledge of chromosome-specific satellite sequences, a feature that has facilitated the study of chromosome-specific behaviors (Dernburg *et al.* 1996; Joyce *et al.* 2012, 2013; Christophorou *et al.* 2013). Although our FISH experiments revealed differences in the contents of satellite DNA repeats on the autosomes of the *D. simulans* clade, it is nevertheless difficult to determine which chromosome (second or third) corresponds to which satellite composition, as these two autosomes are morphologically indistinguishable in all species. We have attempted to overcome this issue in *D. simulans* using two stocks, Dsim\T(Y;2) and Dsim\T(Y;3), where a translocation event between the Y and second or third chromosomes allows us to unambiguously distinguish the second and third chromosomes. By using a probe

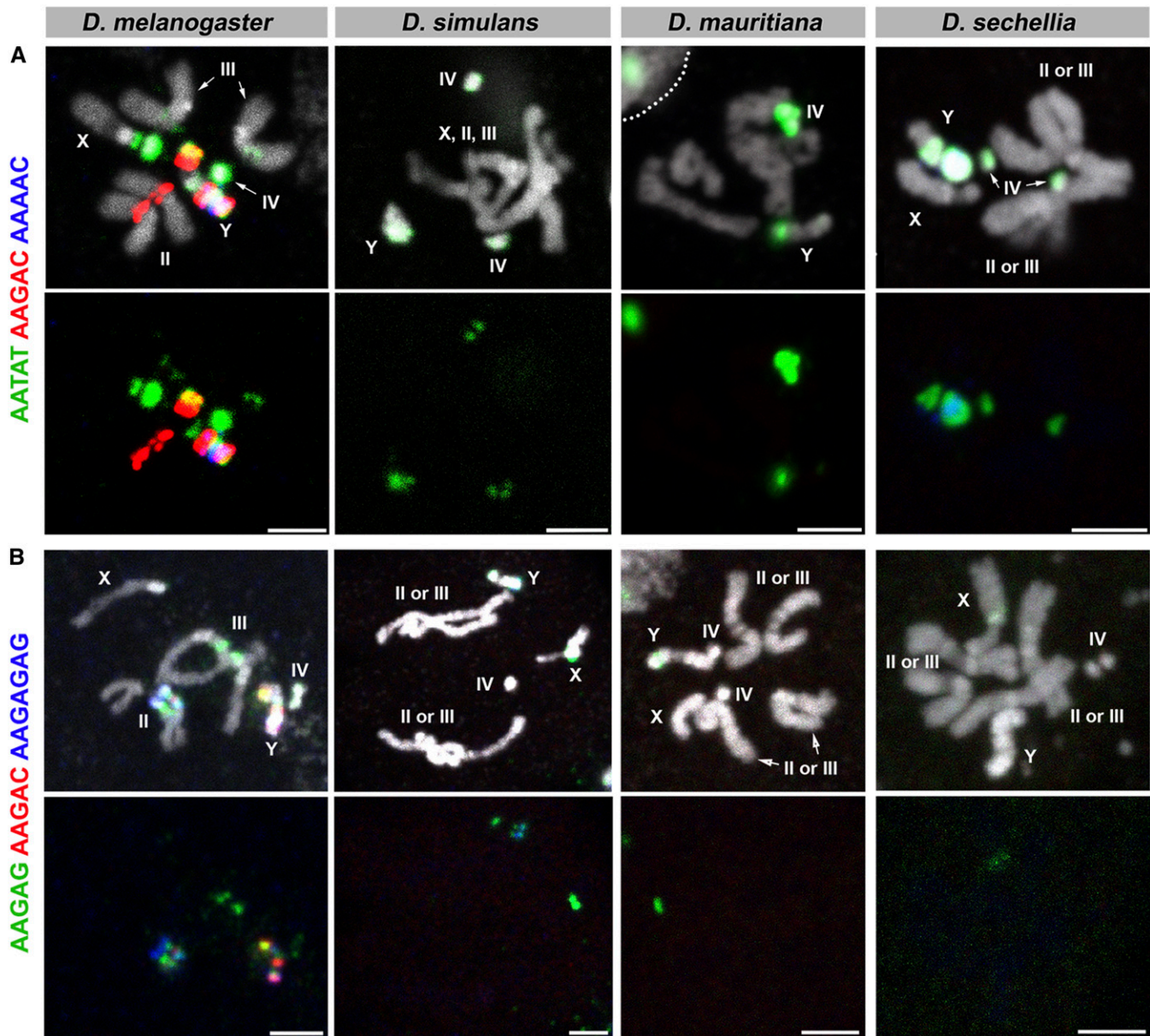


Figure 4 FISH on neuroblast chromosome spread from *D. melanogaster*, *D. simulans*, *D. mauritiana*, and *D. sechellia*. (A) C7 probes and (B) C8 probes. Probe sequences are indicated by the colored text. The top panels show three probes combined with DAPI, and the bottom panels show only probe hybridization signals. Bar, 2.5 μm . C, combination; DAPI, 4',6-diamidino-2-phenylindole; FISH, fluorescent *in situ* hybridization.

combination that focuses on *D. simulans* autosomes (C10, Table 1), we observed these translocation stocks of *D. simulans*. Our data show that the majority of the dodeca satellite repeats are present on the *D. simulans* second chromosome (Figure 6). Considering the fact that the dodeca satellite repeat is present on the third chromosome in *D. melanogaster*, it is clear that each autosome does not have “signature” sequences (*i.e.*, chromosome-specific sequences) that are conserved among the four species examined in this study. Due to the lack of translocation stocks in *D. mauritiana* and *D. sechellia*, the identity of the second and third chromosomes with respect to satellite DNA composition remains unclear in these species.

Finally, we observed that the GAACAGAACATGTTC and 359/372 satellite repeats were asymmetrically distributed between the *D. mauritiana* second and third chromosomes (Figure 2, C4 and Figure 3, C6);

that is, one chromosome (second or third) contained more GAACA-GAACATGTTC satellites than the other, and similarly, one chromosome contained more 359/372 repeats than the other chromosome. Also, we observed a similar asymmetric distribution of the GAACA-GAACATGTTC and the dodeca satellite repeats on the *D. sechellia* second and third chromosomes (Figure 3, C5 and C6). To determine the relationship between these satellite repeats on the autosomes of *D. mauritiana* and *D. sechellia*, we used two probe combinations (C9 and C10) (Figure 7 and Table 1). This revealed that the abundances of 359 bp and GAACAGAACATGTTC satellite repeats are inversely correlated in *D. mauritiana*; the autosome that contains more 359 bp repeats contains less GAACAGAACATGTTC compared to the other autosome. Also, in *D. mauritiana*, the abundances of GAACAGAACATGTTC and dodeca satellite repeats are positively correlated. In

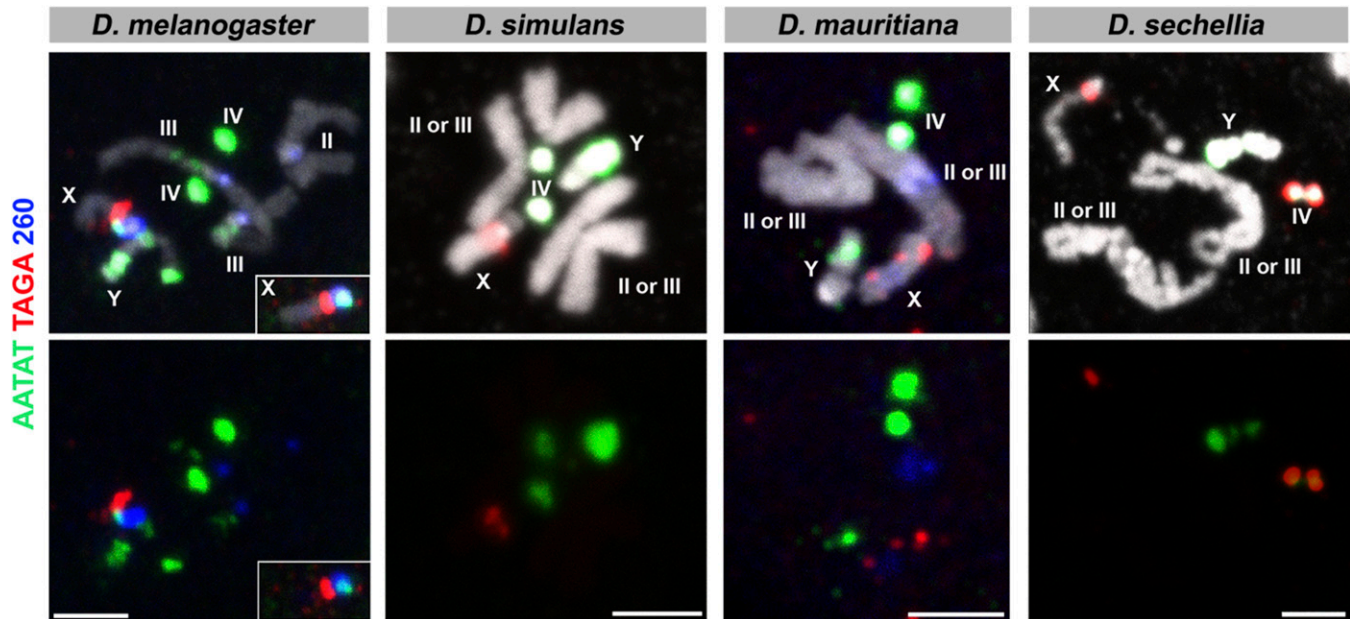


Figure 5 FISH on neuroblast chromosome spread from *D. melanogaster*, *D. simulans*, *D. mauritiana*, and *D. sechellia*. C11 probes were used. Probe sequences are indicated by the colored text. The top panels show three probes combined with DAPI, and the bottom panels show only probe hybridization signals. Bar, 2.5 μm . C, combination; DAPI, 4',6-diamidino-2-phenylindole; FISH, fluorescent *in situ* hybridization.

D. sechellia, the abundance of GAACAGAACATGTTC and dodeca are positively correlated; the autosome that contained more GAACAGAACATGTTC also contained more dodeca compared to the other autosome.

DISCUSSION

D. melanogaster is one of the few species where satellite DNA has most comprehensively been mapped (Lohe *et al.* 1993). In previous studies,

the mapping effort was carried out by the use of radioactive probes, allowing precise quantitation of satellite DNA repeats. In addition, the extensive and unique set of chromosomal aberrations in this historical model organism facilitated the mapping of these repeats onto specific chromosomal loci. However, these mapping experiments cannot be extended to the sibling species of *D. melanogaster* due to the lack of a similar set of chromosomal aberrations. Therefore, other approaches are required to resolve

Table 2 Location of the satellite DNAs on the chromosomes of the *D. melanogaster* species complex

	<i>D. melanogaster</i>	<i>D. simulans</i>	<i>D. mauritiana</i>	<i>D. sechellia</i>
(AATAT) _n	Y, X, III, IV	Y, IV	Y, IV	Y, IV
(AATAG) _n	Y, II, III ^a	Y, X, IV	IV	Y, X
359/372	X ^b , III ^b	II ^b , III ^b	X ^b , II or III ^b	X ^b
260	X ^b , III ^b , II	None	X ^b , II or III ^b	None
(AATAGAC) _n	Y	None	None	None
(AATAC) _n	Y	None	None	Y, IV
(AATAAAC) _n	Y	Y	None	None
IGS	X, Y	X, Y ^c	X, Y	X (not on Y ^{c,d})
(AAGAG) _n	X, Y, II, III, IV	X, Y	Y	X
Prod	II, III	None	None	None
Dodeca	Y ^e , III	II, III	II, III	II, III
(AACAC) _n	Y, II	None	II ^a , III ^a	None
(GAACAGAACATGTTC) _n	None	II, III	II, III, X ^a , Y ^a	II, III
(AACAAAC) _n	II ^a	Y	II ^a , III ^a	None
(AAGAC) _n	Y, II	None	None	None
(AAAAAC) _n	Y	None	None	Y
(AAGAGAG) _n	Y, II	Y	None	None
(TAGA) _n	X	X	X	X, IV

rDNA, ribosomal DNA.

^aIndicates the sequences that are obviously low in abundance.

^bLikely cross-hybridization among 359 bp, 372 bp, 260 bp, and other related repeats.

^cWithout rDNA.

^dRoy *et al.* (2005) reports the presence of rDNA (using Py12 and 240 bp probes) on *D. sechellia* Y, although much weaker than X rDNA. This may be because our IGS-240 probe does not cover the entire length of the 240 bp sequence, and our probe sequence does not hybridize well to *D. sechellia* IGS.

^eYadlapalli and Yamashita (2013) described its presence on Y but this chromosome spread did not find dodeca on Y, possibly due to differences in hybridization conditions or strain.

■ Table 3 Satellite composition of each chromosome in the *D. melanogaster* species complex

	<i>D. melanogaster</i>	<i>D. simulans</i>	<i>D. mauritiana</i>	<i>D. sechellia</i>
X	AATAT 359/372/260 IGS AAGAG	IGS AAGAG AATAG	359/372/260 IGS GAACAGAACATGTTC TAGA	359/372 IGS AAGAG AATAG TAGA
	TAGA	TAGA		
Y	AATAT AAGAC AATAC AATAG AATAGAC AATAAAC IGS AAGAG AACAC AAAAC AAGAGAG	AATAT AATAG AATAAAC IGS AAGAG AAGAGAG AACAAAC	AATAT IGS AAGAG GAACAGAACATGTTC ^a	AATAT AAAAC
	II	AATAG AAGAG Prod AACAC AAGAC AACAAAC AAGAGAG 260	 359/372 Dodeca GAACAGAACATGTTC	 Dodeca GAACAGAACATGTTC
	III	AATAT AATAG AAGAG 359/372/260 Prod Dodeca	359/372 Dodeca GAACAGAACATGTTC	359/372/260 (only II or III) Dodeca GAACAGAACATGTTC AACAC
	IV	AATAT AAGAG	AATAT AATAG	AATAT AATAC TAGA

^aIndicates a sequence that is obviously low in abundance.

the relationship of multiple satellite repeats on the chromosomes of sibling species.

Here, by using multi-color FISH, we have mapped previously identified satellite DNAs onto the genomes of the *D. melanogaster* species complex. To our knowledge, this is the most comprehensive mapping of satellite DNAs across all the species comprising the *D. melanogaster* species complex to date.

We estimate that the detection limit of our method is between 110 and 140 kb. By comparing our FISH results with the previous mapping effort by Lohe *et al.* (1993), which used ³H-labeled radioactive probes and is more accurate in quantification, the detection limit

of our method can be inferred (Table 4). Except for (AATAG)_n on the III, which was estimated to be 30 kb and detected with our method, all satellite repeats that were estimated below 110 kb by Lohe *et al.* (1993) were not detected with our method, whereas all satellite repeats that were estimated above 140 kb by Lohe *et al.* (1993) were detectable with our method (Table 4). Therefore, it is possible that the satellite repeats that exist below 140 kb in the *D. simulans* clade species were missed in this study.

Our work provides a few critical insights into the arrangement and evolution of satellite DNA and reveals the need for future mapping efforts in these species.

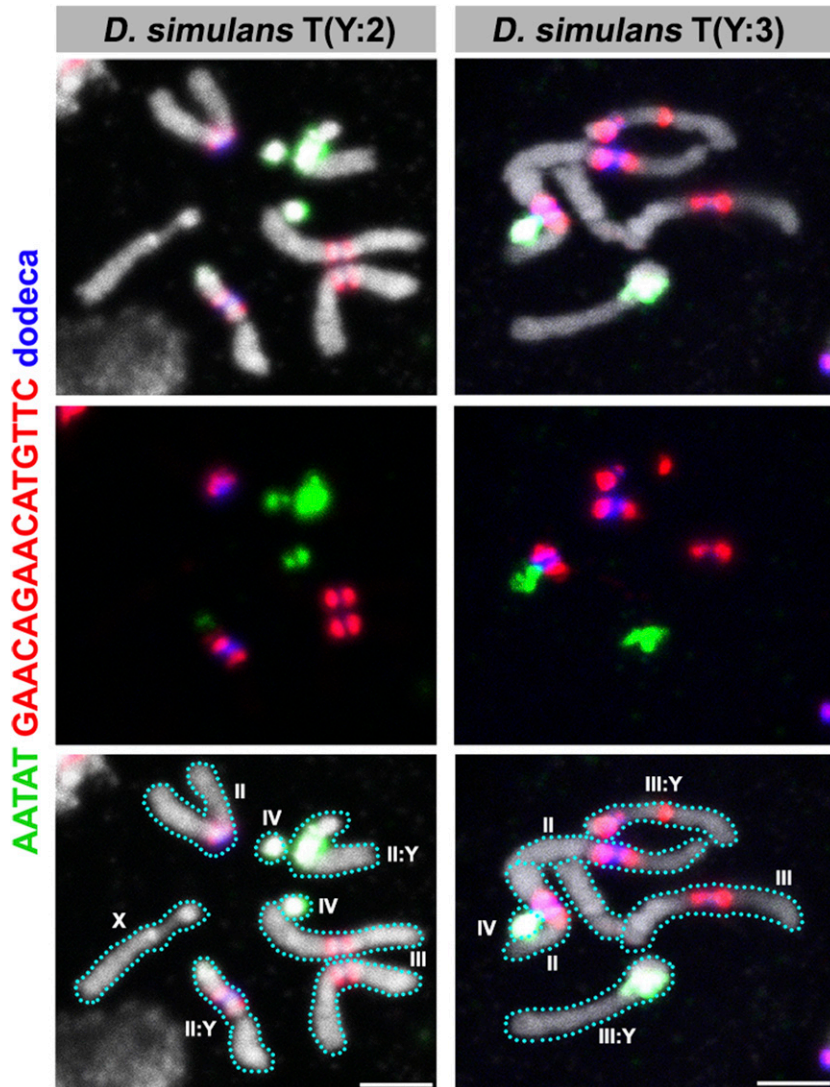


Figure 6 FISH on neuroblast chromosome spread from *D. simulans* T(Y:2) and T(Y:3) translocation strains. C10 probes were used. Probe sequences are indicated by the colored text. The top panels show three probes combined with DAPI, the middle panels show only probe hybridization signals and the bottom panels indicate the interpretation. Bar, 2.5 μm . C, combination; DAPI, 4',6-diamidino-2-phenylindole; FISH, fluorescent *in situ* hybridization.

Satellite DNA divergence in the *D. melanogaster* species complex

Lohe and Brutlag (1987) suggested that after *D. melanogaster* and *D. simulans* diverged from a common ancestor, the satellite DNA sequences between both of these species remained similar in sequence, although different in abundance. Although quantification is not possible with our method of analysis, our study confirms that some satellite DNAs are visibly different in abundance between these species (e.g., 359 bp, AATAAAC, AATAC, AAGAG, AACAC, AAGAC, AAGAGAG, GAACAGAACATGTTC, and prod satellite). Furthermore, our mapping in this study reveals that *D. melanogaster* and *D. simulans* differ from each other in the patterning and location of these satellite DNA repeats. Among four species of the *D. melanogaster* species complex, a few satellite DNA repeats are found to be absent in one or more species, suggesting species specificity of certain satellites. Overall, it is clear that satellite DNA distribution/composition in *D. melanogaster* is most distant within the *D. melanogaster* species complex, whereas those of three species in the *D. simulans* clade (*D. simulans*, *D. mauritiana*, and *D. sechellia*) are much closer to each other, just like their evolutionary relationship. However, it is apparent that there are drastic differences in satellite DNA contents even among species of the *D. simulans* clade. This

suggests that changes in satellite DNA sequence/patterning occurs rapidly over the short evolutionary timespan that separates sibling species, as suggested by Lohe and Brutlag (1987) and Hey and Kliman (1993).

In addition to distinct composition of satellite DNA repeats in these species, multi-color FISH revealed that satellites that are present in distinct locations in one species are overlapping in another species. For instance, the dodeca satellite and AACAC are present at distinct loci on separate autosomes in *D. melanogaster*, but completely overlap in *D. mauritiana* (Figure 3, C5). In another example, GAACAGAACATGTTC and AACAAAC overlap in *D. mauritiana* but are separate in *D. simulans*, where GAACAGAACATGTTC remains on the autosomes and AACAAAC is only present on the Y chromosome (Figure 3, C6). Overlapping signals likely reflect that those satellite sequences are interspersed among each other without forming large blocks of repeats composed of a single sequence. The fact that the repeats are interspersed in one species, but completely separate in another species, indicates that satellite evolution cannot be explained by a single simple translocation event that cuts/copies and pastes a large segment of a chromosome to another. Instead, multiple events of one or many mechanisms must occur to account for the diverse patterning and abundance of satellite DNA repeats in closely related species.

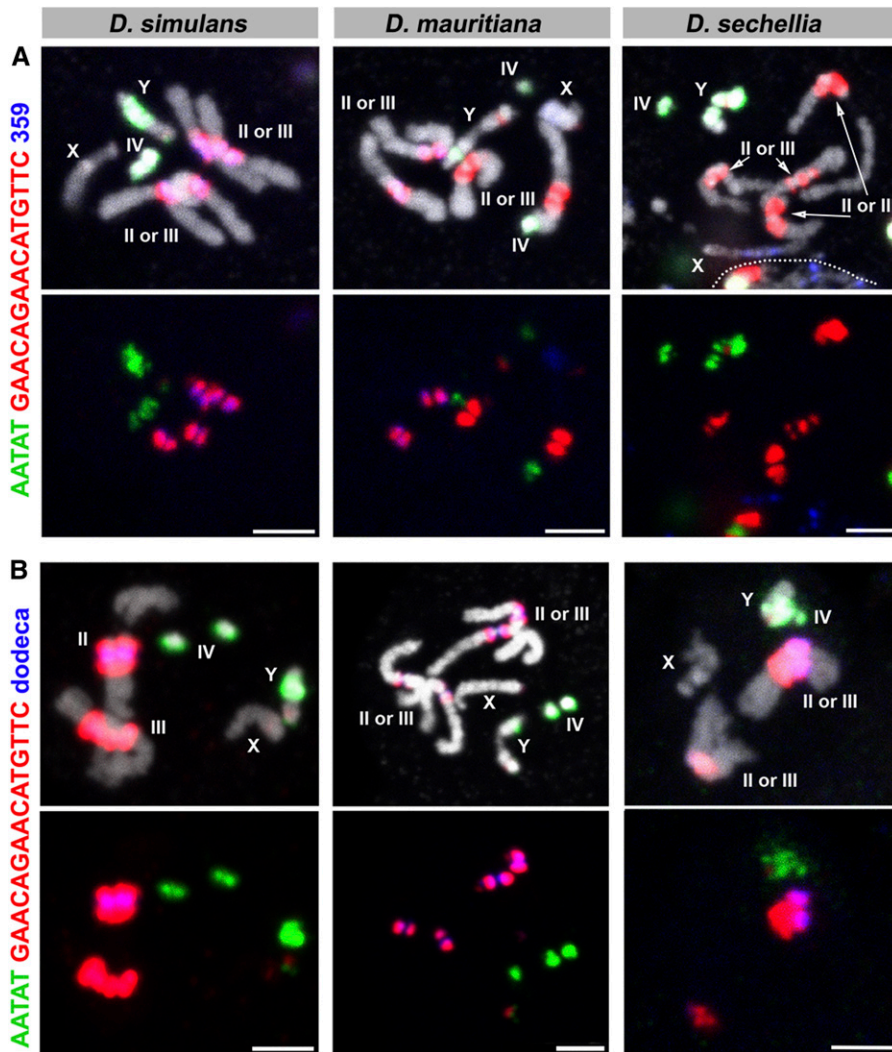


Figure 7 FISH on neuroblast chromosome spread from *D. simulans*, *D. mauritiana*, and *D. sechellia*. (A) C9 probes and (B) C10 probes. Probe sequences are indicated by the colored text. The top panels show three probes combined with DAPI, and the bottom panels show only probe hybridization signals. Bar, 2.5 μ m. C, combination; DAPI, 4',6-diamidino-2-phenylindole; FISH, fluorescent *in situ* hybridization.

Potential underlying mechanisms include transposition and non-homologous recombination among satellites (Ugarkovic and Pohl 2002; Pohl *et al.* 2012).

Possible presence of unidentified satellite DNA sequences in the *D. simulans* clade species

In this study, we mainly utilized satellite DNA probes that were previously identified in *D. melanogaster* and performed comparative mapping in the entire *D. melanogaster* species complex. This leaves the possibility that there are unidentified satellite repeats in *D. simulans* clade species that were left out of our mapping effort. Although we were able to cover the majority of the Y chromosomes of *D. melanogaster*, *D. simulans*, and *D. sechellia* using our satellite DNA probes, only a minor portion of the *D. mauritiana* Y chromosome was visualized by our probe set. As it was shown that the *D. mauritiana* Y chromosome contains a large amount of transposons (*HeT-A* and *TART*) (Berloco *et al.* 2005), it is possible that these transposons are the major constituent of Y chromosome heterochromatin in *D. mauritiana*. It will be of future interest to identify unknown satellite repeats in *D. simulans* clade species. It is possible that those unidentified satellites belong to the classes of “complex repeats,” such as 180 and 500 bp repeats, as indicated previously (Strachan *et al.* 1982).

In summary, our comparative mapping of satellite DNA using multi-color FISH probes has confirmed the highly divergent nature of satellite DNAs among the *D. melanogaster* species complex, and identified

Table 4 Estimating the detection limit of FISH method used in this study

Satellite	X	Y	II	III	IV
(AATAT) _n	600	5800	10	630	2700
(AATAG) _n	8.1	310	200	30^a	78
(AATAC) _n	0	3500	0	0	0
(AAAAC) _n	0	400	0	0	0
(AAGAC) _n	81	8500	1800	110	0
(AAGAG) _n	1200	7200	5500	1100	170
(AATAAAC) _n	0	1600	0	0	0
(AATAGAC) _n	0	1600	0	0	0
(AAGAGAG) _n	270	1800	1700	140	100
(AATAACATAG) _n	0	0	1900	1600	0
359 bp	11000	0	0	0	0

The numbers indicate the amounts (kilobases) of satellite repeats in the *D. melanogaster* genome estimated by Lohe *et al.* (1993). Bold numbers indicate the satellite repeats that were successfully detected with our FISH method. FISH, fluorescent *in situ* hybridization. ^aExcept for 30 kb of (AATAG)_n, which was detected on the third chromosome, the detection limit with our method is likely between 110 and 140 kb [110 kb of (AAGAC)_n being undetectable and 140 kb of (AAGAGAG)_n being detectable].

species-specific and chromosome-specific satellites. We hope that this study will serve as a resource for future work on chromosome biology and speciation in these model organisms.

ACKNOWLEDGMENTS

We thank Patricia Wittkopp, the Bloomington Stock Center, and the *Drosophila* Species Stock Center for *Drosophila* strains, and the Yamashita laboratory for discussions and comments on the manuscript. Research in the Yamashita laboratory is supported by the Howard Hughes Medical Institute. Y.M.Y. is supported by the MacArthur Foundation.

Author contributions: The project was conceived following discussion by all authors. All authors conducted experiments, interpreted the results, and wrote and edited the manuscript.

LITERATURE CITED

- Abad, J. P., M. Carmena, S. Baars, R. D. Saunders, D. M. Glover *et al.*, 1992 Dodeca satellite: a conserved G+C-rich satellite from the centromeric heterochromatin of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 89: 4663–4667.
- Abad, J. P., M. Agudo, I. Molina, A. Losada, P. Ripoll *et al.*, 2000 Pericentromeric regions containing 1.688 satellite DNA sequences show anti-kinetochore antibody staining in prometaphase chromosomes of *Drosophila melanogaster*. *Mol. Gen. Genet.* 264: 371–377.
- Altemose, N., K. H. Miga, M. Maggioni, and H. F. Willard, 2014 Genomic characterization of large heterochromatic gaps in the human genome assembly. *PLoS Comput. Biol.* 10: e1003628.
- Berlaco, M., L. Fanti, F. Sheen, R. W. Levis, and S. Pimpinelli, 2005 Heterochromatic distribution of *HeT-A*- and *TART*-like sequences in several *Drosophila* species. *Cytogenet. Genome Res.* 110: 124–133.
- Bonaccorsi, S., and A. Lohe, 1991 Fine mapping of satellite DNA sequences along the Y chromosome of *Drosophila melanogaster*: relationships between satellite sequences and fertility factors. *Genetics* 129: 177–189.
- Bosco, G., P. Campbell, J. T. Leiva-Neto, and T. A. Markow, 2007 Analysis of *Drosophila* species genome size and satellite DNA content reveals significant differences among strains as well as between species. *Genetics* 177: 1277–1290.
- Bouzinba-Segard, H., A. Guais, and C. Francastel, 2006 Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc. Natl. Acad. Sci. USA* 103: 8709–8714.
- Bueno, D., O. M. Palacios-Gimenez, and D. C. Cabral-de-Mello, 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.
- Carmena, M., J. P. Abad, A. Villasante, and C. Gonzalez, 1993 The *Drosophila melanogaster* dodecasatellite sequence is closely linked to the centromere and can form connections between sister chromatids during mitosis. *J. Cell Sci.* 105(Pt. 1): 41–50.
- Christophorou, N., T. Rubin, and J. R. Huynh, 2013 Synaptonemal complex components promote centromere pairing in pre-meiotic germ cells. *PLoS Genet.* 9: e1004012.
- Dernburg, A. F., J. W. Sedat, and R. S. Hawley, 1996 Direct evidence of a role for heterochromatin in meiotic chromosome segregation. *Cell* 86: 135–146.
- DiBartolomeis, S. M., K. D. Tartof, and F. R. Jackson, 1992 A superfamily of *Drosophila* satellite related (SR) DNA repeats restricted to the X chromosome euchromatin. *Nucleic Acids Res.* 20: 1113–1116.
- Feber, A., G. A. Wilson, L. Zhang, N. Presneau, B. Idowu *et al.*, 2011 Comparative methylome analysis of benign and malignant peripheral nerve sheath tumors. *Genome Res.* 21: 515–524.
- Ferree, P. M., and D. A. Barbash, 2009 Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in *Drosophila*. *PLoS Biol.* 7: e1000234.
- Gatti, M., S. Pimpinelli, and G. Santini, 1976 Characterization of *Drosophila* heterochromatin. I. Staining and decondensation with Hoechst 33258 and quinacrine. *Chromosoma* 57: 351–375.
- Gaubatz, J. W., and R. G. Cutler, 1990 Mouse satellite DNA is transcribed in senescent cardiac muscle. *J. Biol. Chem.* 265: 17753–17758.
- Guenatri, M., D. Bailly, C. Maison, and G. Almouzni, 2004 Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* 166: 493–505.
- Haider, S., L. Cordeddu, E. Robinson, M. Movassagh, L. Siggins *et al.*, 2012 The landscape of DNA repeat elements in human heart failure. *Genome Biol.* 13: R90.
- Hey, J., and R. M. Kliman, 1993 Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. *Mol. Biol. Evol.* 10: 804–822.
- Hoskins, R. A., J. W. Carlson, K. H. Wan, S. Park, I. Mendez *et al.*, 2015 The Release 6 reference sequence of the *Drosophila melanogaster* genome. *Genome Res.* 25: 445–458.
- Ito, H., A. Miura, K. Takashima, and T. Kakutani, 2007 Ecotype-specific and chromosome-specific expansion of variant centromeric satellites in *Arabidopsis thaliana*. *Mol. Genet. Genomics* 277: 23–30.
- Joyce, E. F., B. R. Williams, T. Xie, and C. T. Wu, 2012 Identification of genes that promote or antagonize somatic homolog pairing using a high-throughput FISH-based screen. *PLoS Genet.* 8: e1002667.
- Joyce, E. F., N. Apostolopoulos, B. J. Beliveau, and C. T. Wu, 2013 Germline progenitors escape the widespread phenomenon of homolog pairing during *Drosophila* development. *PLoS Genet.* 9: e1004013.
- Kamm, A., I. Galasso, T. Schmidt, and J. S. Heslop-Harrison, 1995 Analysis of a repetitive DNA family from *Arabidopsis arenosa* and relationships between *Arabidopsis* species. *Plant Mol. Biol.* 27: 853–862.
- Karpen, G. H., M. H. Le, and H. Le, 1996 Centric heterochromatin and the efficiency of achiasmate disjunction in *Drosophila* female meiosis. *Science* 273: 118–122.
- Kawabe, A., and D. Charlesworth, 2007 Patterns of DNA variation among three centromere satellite families in *Arabidopsis halleri* and *A. lyrata*. *J. Mol. Evol.* 64: 237–247.
- Kawabe, A., and S. Nasuda, 2006 Chromosome-specific satellite sequences in *Turritis glabra*. *Genes Genet. Syst.* 81: 287–290.
- Koo, D. H., V. K. Tiwari, E. Hribova, J. Dolezel, B. Friebe *et al.*, 2016 Molecular cytogenetic mapping of satellite DNA sequences in *Aegilops geniculata* and wheat. *Cytogenet. Genome Res.* 148: 314–321.
- Kuhn, G. C., H. Kuttler, O. Moreira-Filho, and J. S. Heslop-Harrison, 2012 The 1.688 repetitive DNA of *Drosophila*: concerted evolution at different genomic scales and association with genes. *Mol. Biol. Evol.* 29: 7–11.
- Lamb, J. C., J. M. Meyer, B. Corcoran, A. Kato, F. Han *et al.*, 2007 Distinct chromosomal distributions of highly repetitive sequences in maize. *Chromosome Res.* 15: 33–49.
- Larracuente, A. M., 2014 The organization and evolution of the *Responder* satellite in species of the *Drosophila melanogaster* group: dynamic evolution of a target of meiotic drive. *BMC Evol. Biol.* 14: 233.
- Larracuente, A. M., and P. M. Ferree, 2015 Simple method for fluorescence DNA *in situ* hybridization to squashed chromosomes. *J. Vis. Exp.* 95: 52288.
- Lohe, A. R., and D. L. Brutlag, 1986 Multiplicity of satellite DNA sequences in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 83: 696–700.
- Lohe, A. R., and D. L. Brutlag, 1987 Identical satellite DNA sequences in sibling species of *Drosophila*. *J. Mol. Biol.* 194: 161–170.
- Lohe, A. R., and P. A. Roberts, 1990 An unusual Y chromosome of *Drosophila simulans* carrying amplified rDNA spacer without rRNA genes. *Genetics* 125: 399–406.
- Lohe, A. R., and P. A. Roberts, 2000 Evolution of DNA in heterochromatin: the *Drosophila melanogaster* sibling species subgroup as a resource. *Genetica* 109: 125–130.
- Lohe, A. R., A. J. Hilliker, and P. A. Roberts, 1993 Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Genetics* 134: 1149–1174.

- Losada, A., and A. Villasante, 1996 Autosomal location of a new subtype of 1.688 satellite DNA of *Drosophila melanogaster*. *Chromosome Res.* 4: 372–383.
- Louzada, S., A. Vieira-da-Silva, A. Mendes-da-Silva, S. Kubickova, J. Rubes *et al.*, 2015 A novel satellite DNA sequence in the *Peromyscus* genome (PMSat): evolution via copy number fluctuation. *Mol. Phylogenet. Evol.* 92: 193–203.
- Menon, D. U., C. Coarfa, W. Xiao, P. H. Gunaratne, and V. H. Meller, 2014 siRNAs from an X-linked satellite repeat promote X-chromosome recognition in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 111: 16460–16465.
- Paco, A., F. Atega, N. Mestrovic, M. Plohl, and R. Chaves, 2014 Evolutionary story of a satellite DNA from *Phodopus sungorus* (Rodentia, Cricetidae). *Genome Biol. Evol.* 6: 2944–2955.
- Plohl, M., N. Mestrovic, and B. Mravinac, 2012 Satellite DNA evolution. *Genome Dyn.* 7: 126–152.
- Rosic, S., F. Kohler, and S. Erhardt, 2014 Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J. Cell Biol.* 207: 335–349.
- Roy, V., L. Monti-Dedieu, N. Chaminade, S. Siljak-Yakovlev, S. Aulard *et al.*, 2005 Evolution of the chromosomal location of rDNA genes in two *Drosophila* species subgroups: *ananassae* and *melanogaster*. *Heredity* (Edinb) 94: 388–395.
- Sawamura, K., M. T. Yamamoto, and T. K. Watanabe, 1993 Hybrid lethal systems in the *Drosophila melanogaster* species complex. II. The Zygotic hybrid rescue (*Zhr*) gene of *D. melanogaster*. *Genetics* 133: 307–313.
- Schmid, M., and C. Steinlein, 2015 Chromosome banding in Amphibia. XXXII. The genus *Xenopus* (Anura, Pipidae). *Cytogenet. Genome Res.* 145: 201–217.
- Simeone, A., A. La Volpe, and E. Boncinelli, 1985 Nucleotide sequence of a complete ribosomal spacer of *D. melanogaster*. *Nucleic Acids Res.* 13: 1089–1101.
- Strachan, T., E. Coen, D. Webb, and G. Dover, 1982 Modes and rates of change of complex DNA families of *Drosophila*. *J. Mol. Biol.* 158: 37–54.
- Strachan, T., D. Webb, and G. A. Dover, 1985 Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. *EMBO J.* 4: 1701–1708.
- Ting, D. T., D. Lipson, S. Paul, B. W. Brannigan, S. Akhavanfard *et al.*, 2011 Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* 331: 593–596.
- Ugarkovic, D., and M. Plohl, 2002 Variation in satellite DNA profiles—causes and effects. *EMBO J.* 21: 5955–5959.
- Usakin, L., J. Abad, V. V. Vagin, B. de Pablos, A. Villasante *et al.*, 2007 Transcription of the 1.688 satellite DNA family is under the control of RNA interference machinery in *Drosophila melanogaster* ovaries. *Genetics* 176: 1343–1349.
- Waring, G. L., and J. C. Pollack, 1987 Cloning and characterization of a dispersed, multicopy, X chromosome sequence in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 84: 2843–2847.
- Wong, L. H., K. H. Brettingham-Moore, L. Chan, J. M. Quach, M. A. Anderson *et al.*, 2007 Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res.* 17: 1146–1160.
- Yadlapalli, S., and Y. M. Yamashita, 2013 Chromosome-specific nonrandom sister chromatid segregation during stem-cell division. *Nature* 498: 251–254.
- Yunis, J. J., and W. G. Yasmin, 1971 Heterochromatin, satellite DNA, and cell function. Structural DNA of eucaryotes may support and protect genes and aid in speciation. *Science* 174: 1200–1209.

Communicating editor: B. J. Andrews