

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26

The sex locus is tightly linked to factors conferring sex-specific lethal effects in the mosquito *Aedes aegypti*

Elzbieta Krzywinska¹, Vladimir Kokoza², Marion Morris³, Elena de la Casa-Esperon⁴, Alexander S. Raikhel², Jaroslaw Krzywinski¹

¹The Pirbright Institute, Pirbright, GU24 0NF, UK, ²Department of Entomology and Institute for Integrative Genome Biology, University of California, Riverside, California 92521, ³Vector Group, Liverpool School of Tropical Medicine, Liverpool, L3 5QA, UK, ⁴ Castilla-La Mancha Science and Technology Park and University of Castilla-La Mancha, 02008 Albacete, Spain

Correspondence: Jaroslaw Krzywinski, The Pirbright Institute, Ash Road, Pirbright, GU24 0NF, UK. Tel.: +44 148 323 1211. Fax: +44 148 323 2448. E-mail: krzywin@pirbright.ac.uk

Running title: Genetic analysis of the *Aedes* sex locus neighborhood

Word count: 6,151 (including abstract)

27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

ABSTRACT

In many taxa, sex chromosomes are heteromorphic and largely non-recombining. Evolutionary models predict that spread of recombination suppression on the Y chromosome is fuelled by the accumulation of sexually antagonistic alleles in close linkage to the sex determination region. However, empirical evidence for the existence of sexually antagonistic alleles is scarce. In the mosquito *Aedes aegypti*, the sex-determining chromosomes are homomorphic. The region of suppressed recombination, which surrounds the male-specific sex determining gene, remains very small, despite ancient origin of the sex chromosomes in the *Aedes* lineage. We conducted a genetic analysis of the *A. aegypti* chromosome region tightly linked to the sex locus. We used a strain with an EGFP-tagged transgene inserted near the male-determining gene to monitor crossing-over events close to the boundary of the sex determining region, and to trace the inheritance pattern of the transgene in relation to sex. In a series of crossing experiments involving individuals with a recombinant sex chromosome we found developmental abnormalities leading to 1:2 sex biases, caused by lethality of half of the male or female progeny. Our results suggest that various factors causing sex-specific lethal effects are clustered within the neighbourhood of the sex determining region, which in the affected sex are likely lost or gained through recombination, leading to death. These may include genes that are recessive lethal, vital for development, and/or sexually antagonistic. The sex chromosome fragment in question represents a fascinating test case for the analysis of processes that shape stable boundaries of a non-recombining region.

Key words: sex chromosome, sex linkage, sex-determining locus, recombination, sex-specific lethality

INTRODUCTION

53

54 Evolutionary theory predicts that selection will favor mechanisms that reduce recombination
55 between the primary sex determination gene and genes with sexually antagonistic alleles
56 arising near the sex locus (FISHER 1931; CHARLESWORTH AND CHARLESWORTH 1980; BULL 1983;
57 RICE 1987). In the XY systems with the Y carrying the primary sex determiner, it will promote the
58 accumulation, near the male determining locus, of genes that are beneficial to males, but
59 detrimental to females. Tight enough linkage to the sex locus would allow accumulation on the
60 Y, and spread in a population, of alleles with selective advantage in males, even if they are
61 highly deleterious or lethal to females (RICE 1987). Complete recombination suppression
62 between the X and Y in the region involved would ensue, creating the sex determining region
63 (SDR). Close linkage to the SDR would promote the accumulation of sexually antagonistic
64 alleles in the regions adjacent to the SDR boundary, fuelling expansion of a non-recombining
65 portion of the Y chromosome. This process could continue until the X and Y fail to recombine
66 over their entire lengths, leading to mutation-driven erosion of Y-linked genes, accumulation of
67 repetitive sequences, and, eventually, heteromorphism of sex chromosomes (BACHTROG 2013).
68 However, in many taxa cessation of recombination in sex chromosomes has not gone to
69 completion. In eutherian mammals sex chromosomes are largely non-recombining and
70 heteromorphic, but homologous pairing and exchange of chromosomal arms has been
71 maintained within the pseudoautosomal regions for over 140 million years (VEYRUNES *et al.*
72 2008). In some groups, such as ratite birds and boid snakes, the SDR has not expanded and
73 the sex-determining chromosomes remain homomorphic, despite being nearly as ancient as
74 those in Eutheria (BACHTROG *et al.* 2011). The question why such homomorphic chromosomes
75 persist remains unresolved. Attempts to explain it include lack of sexually antagonistic mutations
76 in some species and resolution of sexual conflict by sex-specific or sex-biased expression
77 (VICOSO *et al.* 2013).

78 All but one of mosquito species studied have three pairs of chromosomes; the only
79 known exception, *Chagasia bathana*, has four pairs (KUMAR and RAI 1993). In *Anopheles*, the
80 karyotype comprises two pairs of freely-recombining autosomes and a pair of non-recombining
81 heteromorphic sex chromosomes, with males being heterogametic (XY) and females
82 homogametic (XX). In culicines, such as *Aedes* and *Culex*, chromosomes are homomorphic and
83 each pair, traditionally referred to as chromosomes 1, 2 and 3, undergoes recombination. In
84 *Culex* and *Aedes*, chromosome 1 is sex-determining; its p and q arms are largely syntenic to,
85 respectively, the X chromosome and chromosomal arm 2R of *An. gambiae* (NENE *et al.* 2007;
86 ARENSBURGER *et al.* 2010). In *A. aegypti* males, one chromosome of the chromosome 1 pair
87 (hereafter called the M-chromosome, as opposed to the non-sex-specific m-chromosome)
88 carries the sex-determining region that does not recombine and harbors a dominant male-
89 determining gene *Nix* (HALL *et al.* 2015). Thus, similar to anophelines, the *A. aegypti* males are
90 heterogametic (Mm) and females homogametic (mm). The SDR, located in that species in the
91 chromosomal region 1q21 (TIMOSHEVSKIY *et al.* 2013), appears to be very short. Its physical
92 location was delineated using two *A. aegypti* transgenic strains (called 'sensor' and J2), each
93 tagged with a different fluorescent marker integrated into the m-chromosome in the vicinity of
94 the sex locus ('sensor' in the rDNA array) (HALL *et al.* 2014). Crossing experiments yielded
95 males with the M-chromosome that sequentially acquired both transgenes, each on the opposite
96 flanks of the SDR, through a low rate recombination ('sensor': 0.4%, J2: 1.24%). Very little is
97 known about the content of the SDR and its neighborhoods. In addition to *Nix*, a male-
98 specifically expressed gene *myo-sex* was identified between the two transgenic markers (HALL
99 *et al.* 2014). However, *myo-sex* can be transferred together with the J2 transgene onto the m-
100 chromosome through recombination, and thus it is not located within the SDR. Fluorescence *in*
101 *situ* hybridization (FISH) to mitotic chromosome spreads showed that in double recombinant
102 males the hybridization signal from the transgenes practically co-localizes with the signal from
103 *Nix* and *myo-sex* (HALL *et al.* 2014; HALL *et al.* 2015). An estimated resolution of FISH probes

104 mapping to mitotic chromosomes (TIMOSHEVSKIY *et al.* 2013) indicates that the SDR may be
105 shorter than 1 Mb. Despite a short non-recombining region, the *Aedes* sex determining
106 chromosomes likely have an ancient origin. *Aedes* may have diverged from *Culex* not earlier
107 than 170 Mya (REIDENBACH *et al.* 2009), yet genetic markers that flank the maleness locus are
108 conserved in both taxa (MORI *et al.* 1999), which suggests that their SDRs may share the same
109 ancestry.

110 *A. aegypti* is known for naturally occurring departures from equal sex ratio toward
111 excess of males (CRAIG *et al.* 1961). In some laboratory strains the proportion of females can
112 vary from 15% to 30%, while for some inter-strain crosses it is close to 0%. Skewed sex ratios
113 are caused by a system of segregation distorter genes of unknown nature. In these cases the
114 levels of distortion are not faithfully inherited between generations or by individuals of the same
115 family [(HICKEY and CRAIG 1966), and references therein]. Stably inherited 1:2 sex biases are
116 also known. Wood (WOOD 1975) reported 1:2 sex biases toward either females or males in
117 broods from single pairs of individuals drawn from a population with an overall parity between
118 the sexes; the distorted sex ratios were hypothesized to result from the action of recessive lethal
119 genes. McGivern and Rai (MCGIVERN and RAI 1974) observed the 1♂:2♀ ratio in a progeny of
120 wild-type females crossed to males carrying a large irradiation-induced translocation of a
121 segment of chromosome 2 onto the M-chromosome and a large paracentric inversion
122 embedded within the translocated region. The sex bias was attributed to a single crossover
123 within the inversion loop during male meiosis, which led to a loss of half of the M-chromosomes
124 in dicentric bridges and to an inviability of the affected gametes.

125 Here we conducted a genetic analysis of the chromosome 1 region tightly linked to the
126 SDR in *A. aegypti*. We generated a transgenic strain, in which insertion of a transgene close to
127 the SDR boundary resulted in a predominantly male inheritance of an eye promoter-controlled
128 EGFP tag. Rare crossing-over events during male meiosis break the linkage between the
129 marker and the sex locus. We observed developmental abnormalities leading to 1:2 sex ratios

130 caused by lethality of either males or females in families derived from individuals carrying a
131 recombinant chromosome 1. These results indicate that within the SDR neighbourhood there
132 are several factors, which in the affected sex are likely lost or gained through recombination.
133 The *A. aegypti* genome is not assembled in that chromosomal region, which complicates
134 identification of the molecular background of these intriguing phenomena.

135

136

MATERIALS AND METHODS

137

138 **Mosquitoes**

139 *A. aegypti* wild-type Rockefeller/UGAL strain was reared at 27 °C and 80% humidity. Three-day
140 old previtellogenic females were fed on anesthetized rats. Three days after blood meal the eggs
141 were collected and prepared for microinjection as described previously (BIAN *et al.* 2005). All
142 mosquitoes used in this study were maintained at the same insectary conditions and their
143 females fed either on rats or on expired human transfusion blood using the Hemotek membrane
144 feeding system (Discovery Workshops, Accrington, UK).

145

146 **Molecular construct and transformation**

147 The pBac[3xP3-EGFP, *afm*] vector containing an enhanced GFP transformation marker under
148 the control of the 3xP3 eye-specific promoter (HORN and WIMMER 2000) was used to subclone a
149 DNA fragment containing the vitellogenin (*Vg*) gene promoter linked to the cecropin A (*CecA*)
150 gene cDNA and SV40 polyadenylation element (KOKOZA *et al.* 2010). The resulting pBac[3xP3-
151 EGFP *afm*, *Vg*-*CecA*] donor plasmid was mixed with the *phsp*-pBac helper plasmid at final
152 concentrations of 0.35 µg/mL and 0.25 µg/mL, respectively, in a 5 mM KCl and 0.1 mM
153 NaH₂PO₄ (pH 6.8) buffer. Injections with the above plasmid mixture into preblastoderm *A.*
154 *aegypti* Rockefeller/UGAL strain embryos and development of transgenics were performed as
155 described (KOKOZA *et al.* 2001).

156

157 **Microscopy**

158 Larvae and pupae of transgenic mosquitoes were screened for EGFP fluorescence in their eyes
159 using a Leica MZ FLIII stereo fluorescence microscope (Leica, Wetzlar, Germany) equipped
160 with a GFP-B filter (GFP Band pass, cat # C/6455, Ex 470/40 DM 495 BA 525/50). The intensity
161 of EGFP fluorescence was used to score progeny of line B-derived females as hetero- or
162 homozygotes. Images of larvae were captured with a Nikon DXM2100 camera (Nikon Inc.,
163 Melville, NY, USA) on a Leica MZ FLIII microscope. Emerging mosquitoes were photographed
164 using Canon D1000 camera (Canon, Tokyo, Japan) with a close-up lens.

165

166 **Genetic analysis**

167 To determine the linkage group of the piggyBac insertion, individual EGFP-positive males were
168 crossed with 2-3 females of *A. aegypti* RED strain, in which a recessive mutation associated
169 with a distinct visible marker is present on each of the three chromosomes: the red-eye (*re*)
170 locus on chromosome 1, the spot-abdomen (*s*) locus on chromosome 2 and the black-tarsus
171 (*blt*) locus on chromosome 3 (SEVERSON *et al.* 1993). The resulting male EGFP-positive
172 progeny were similarly backcrossed with RED strain females. The F₂ segregating populations
173 were scored for the RED phenotype and the presence of EGFP expression.

174 To evaluate the frequency of crossing-over between the sex determining locus and the
175 transgene integration site, the F₁ progeny of the EGFP-positive males and EGFP-negative
176 females from either the C42 transgenic strain (generated in this study) or the RED strain were
177 produced by single pair matings or mass matings and scored for gender and EGFP expression
178 at the pupal stage.

179 Sex ratios were evaluated in families represented by progeny of individual females
180 mated with single males. Up to 10 females were used for crosses with the same male individual,

181 if no more males were available. For families indicating excess of males or females, chi-square
182 analysis was used to test for statistical significance of deviation from the expected 1:1 ratio.

183 To evaluate mortality in the postembryonic stages, numbers of newly hatched larvae
184 were compared with the numbers of pupae in 10 families (progeny of 10 single females). This
185 procedure was done for the families with the 1:2 sex ratio bias, as well as the families of the
186 C42 strain and a wild-type Rockefeller strain mosquitoes.

187

188 **Molecular analysis**

189 DNA was extracted from individual mosquitoes using DNeasy Blood & Tissue Kit (Qiagen,
190 Manchester, UK) according to manufacturer's protocol. PCR was performed in 50 µl containing
191 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 0.2 mM each dNTP, 2.5 U Platinum *Taq*
192 polymerase (Invitrogen, Carlsbad, CA, USA), 10–25 pmol each primer, and 1 µl template DNA
193 (1/100th of the DNA extracted from a single mosquito). PCR thermal cycling included 3 min
194 initial denaturation at 94 °C, followed by 35 cycles of 40 s at 94 °C, 45 s at 52–65 °C, and 30-
195 240 s at 72 °C, and a final elongation for 10 min at 72 °C. Inverse PCR was conducted using
196 either Platinum *Taq* or *LA Taq* polymerase (Takara Bio, Otsu, Japan) and, as templates,
197 genomic DNA of EGFP-positive males digested singly with selected restriction endonucleases
198 and then circularized by ligation. The endonucleases and sequences of the corresponding
199 primers used in inverse PCR are listed in Supplementary Table S1; the primer sequences were
200 either published earlier, or designed manually and screened for a negligible potential to form
201 dimers and hairpin structures using OligoAnalyzer
202 (<http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/>). PCR products were purified using
203 QIAquick Gel Extraction Kit (Qiagen, Manchester, UK), ligated into pGEM-T Easy vectors
204 (Promega Corporation, Madison, WI, USA), and electroporated into *E. coli* ElectroMAX DH10B
205 cells (Invitrogen, Carlsbad, CA, USA). Cloned templates were PCR amplified and sequenced

206 using ABI BigDye terminator chemistry (PE Applied Biosystems, Foster City, CA, USA) on an
207 ABI 3130xl Genetic Analyzer (PE Applied Biosystems, Foster City, CA, USA). The sequences
208 were used as queries to search GenBank and *A. aegypti* genome databases using BLAST
209 (ALTSCHUL *et al.* 1990).

210 Integrity of the junctions between genomic DNA and the transgene was tested by PCR
211 using the following primers: 5' junction, near_pBac5'_2 (5'-CTTCGATGTCGGCTCTTCCT-3')
212 located in the genomic flank and pBc5'_invR2 (5'-CCTCGTGCGCTCTCCTGTTC-3') located
213 within the transgene; the 3' junction, pBac3'_invFN (5'-CAGTAGGAAGACGAATAGGTGG-3')
214 located within the transgene and near_pBac3'_2 (5'-GTTGTCTTCCATTGAATACGCA-3')
215 located within the genomic flank. Loss of the transgene in males and gain thereof in females as
216 a consequence of crossing over during male meiosis was tested by PCR using primers listed
217 above, and primers EGFP_end (5'-CTTGTACAGCTCGTCCATGCC-3') and EGFP_invF (5'-
218 GGGCATCGACTTCAAGGAGGAC-3') targeting a portion of the EGFP gene. Presence of the
219 *myo-sex* gene in the recombinant and non-recombinant individuals of the C42 strain was tested
220 by PCR using primers Myo_1_F (5'-CCTTCAAGCACACCGTTACA-3') and Myo_1_R (5'-
221 TCACTATGCAGGAGTTGTTTCG-3'). In addition, to evaluate quality of the DNA templates,
222 PCR was done using primers *AmsF2* (5'-TTCGAGACGCTCAAGTACGA-3') and *AmsR2* (5'-
223 CTCACGGTCCTTTTCGATGT-3') targeting a fragment of the *Ams* gene, which is located in
224 superconting 1.64 mapped to chromosome 3 (KRZYWINSKA and KRZYWINSKI 2009;
225 TIMOSHEVSKIY *et al.* 2014).

226 For Southern blot analysis, genomic DNA from single males and females of the C42
227 strain and from males of the wild-type Rockefeller strain mosquitoes was digested with the
228 respective restriction endonucleases. The DNA fragments were separated by electrophoresis on
229 a 0.8% agarose gel and transferred by capillary blotting onto Hybond-N+ membranes
230 (Amersham Biosciences, Buckinghamshire, UK) in 10X standard saline citrate (SSC) buffer

231 (SAMBROOK *et al.* 1998). Southern blots were hybridized overnight as described previously
232 (SEVERSON 1997) with a 653 bp long EGFP probe radioactively labeled using the gene-specific
233 primers. Membranes were washed twice in 2X SSC, 0.1% SDS, and then three times at high
234 stringency in 0.1X SSC, 0.1% SDS at 65 °C for 15 min each.

235

236

RESULTS

237 **Generation of a transgenic line**

238 Approximately 1,000 injected pre-blastoderm embryos yielded 100 G₀ adult individuals, which
239 were crossed with the wild-type mosquitoes. Out of 43 resulting families, one with a G₀ male
240 founder was used to establish a transgenic strain (C42), in which only males expressed EGFP,
241 as indicated by screening of randomly selected individuals from several consecutive
242 generations. The C42 males were shown to carry a single transposon insertion in their genomes
243 (Figures 1A and 1B). Females lacked the EGFP expression altogether (Supplementary Figure
244 S1); no transgene sequence was detected in their genome by PCR (Figure 2) or Southern blot
245 analysis (data not shown). The transgene construct has remained stably integrated into the
246 male genome for more than 60 generations. The insertion seemed not to markedly affect fitness
247 of the C42 males.

248

249 **Characterization of the insertion site**

250 Junctions between the transgene and the mosquito genomic DNA were isolated using inverse
251 and standard PCR approaches (Figure 1A; see also Supplementary Information). The sequence
252 identified at the 5' junction contained a nearly 2 kb long fragment of flanking genomic DNA, an
253 unexpected 873 bp-long fragment of a donor plasmid, and a *piggyBac* inverted terminal repeat.
254 The fragment isolated at the 3' junction consisted of an unexpectedly truncated transposon arm
255 (missing 488 bp from the 3' *piggyBac* terminus) abutting a 367 bp long flanking genomic DNA.

256 Thus, the integration of the element occurred either by a non-canonical transposition, or by an
257 illegitimate recombination independent of a transposase activity. Integrity of the transgene and
258 its contiguity with genomic flanks was confirmed by PCR (Figure 1C). Unlike in other dipterans,
259 *piggyBac* transposons integrated into the *A. aegypti* genome may not be able to remobilize
260 (PALAVESAM *et al.* 2013). Fortuitous truncation of the 3' arm only strengthens that notion; it
261 made the insertion highly stable, because remobilization of *piggyBac* would require both arms
262 intact (LI *et al.* 2001).

263 The genomic DNA flanking the transgene from the 5' end has several, periodically
264 occurring, perfect matches to genomic supercontig 1.836 (Genbank GI number: 78216866) and
265 represents a fragment of the 28S ribosomal RNA (rRNA) gene, which forms a part of the rRNA
266 array of cistrons. In *A. aegypti* the rRNA genes are tandemly repeated ~500 times and located
267 on chromosome 1 in a cluster adjacent to the SDR (GALE and CRAMPTON 1989; KUMAR and RAI
268 1990; TIMOSHEVSKIY *et al.* 2013). The genomic sequence flanking the 3' end of the transgene
269 represents a class of non-coding DNA organized in short tandem repeats, with matches to
270 sequences irregularly scattered between clusters of rRNA genes within supercontig 1.836.
271 Similar to the C42 transgene, the 'sensor' transgene is also inserted within the array of rRNA
272 genes (HALL *et al.* 2014).

273 Using primers targeting both genomic flanks, we isolated by PCR and sequenced a
274 region from the wild-type genome (Rockefeller and Liverpool strains of *A. aegypti*), into which
275 *piggyBac* apparently became inserted in the C42 strain. Sequences identical or similar to two
276 parts (corresponding to each flank of the C42 transgene) of such an "empty" site are
277 interspersed in supercontig 1.836, but rather than forming an expected contiguous DNA stretch,
278 they are spaced by a minimum 4 kb. Thus, the supercontig, or other currently available *A.*
279 *aegypti* genome data (NENE *et al.* 2007) do not provide information on a larger genomic context
280 of the transgene integration, perhaps because of instability of shotgun clones containing

281 tandemly repeated sequences (SONG *et al.* 2001) and/or inadequate genome sampling during
282 the *A. aegypti* genome project. Lack of the SDR-linked contigs within the *A. aegypti* genome has
283 been highlighted earlier (HALL *et al.* 2014).

284 A comparison of the “empty” site sequence with the sequences flanking the transgene
285 revealed a deletion in a repetitive DNA fragment that likely occurred during the transposon
286 integration process (Supplementary Figure S2). The deficiency is apparently small and
287 inconsequential for the viability and fertility of mosquitoes, because the C42 strain males mated
288 with the RED, Liverpool, or non-recombinant C42 strain females produce progeny with
289 balanced sex ratios and in numbers comparable to the progeny of the wild-type males (Table 1,
290 and data not shown).

291

292 ***piggyBac* integration into chromosome 1 and linkage to the sex locus**

293 During the strain maintenance we noted a rare occurrence of males lacking EGFP expression.
294 To explore the EGFP inheritance pattern and to confirm chromosomal location of the transgene,
295 we conducted crossing experiments between the EGFP-positive males and females of the RED
296 strain that carries visible recessive markers on each chromosome. Segregation of the
297 phenotypes in the back-cross F₂ generation was consistent with the integration of the transgene
298 in chromosome 1 and its linkage to the *re* locus (Table 1), which itself is linked to the sex locus
299 (MCCLELLAND 1966). Therefore, an almost exclusive presence of the EGFP tag in males is a
300 consequence of the transgene integration into the M-chromosome, close to the SDR. A small
301 number of EGFP-positive females and EGFP-negative males among F₁ and F₂ individuals
302 indicates rarely occurring recombination between the M factor and the transgene, and transfer
303 of the transgene onto the non-sex-specific m-chromosome during male meiosis (Figure 2). The
304 low proportion of recombinants (0.29%) in the crossing experiments described above and in the
305 progeny from mass crosses between the EGFP-tagged males and the EGFP-negative females
306 (>8,000 randomly selected F₁ individuals screened) indicates that the transgene is tightly linked

307 to the SDR. Slightly higher recombination rate (0.4%) was observed between the M locus and
308 the 'sensor' transgene (HALL *et al.* 2014).

309 We used PCR to follow the inheritance pattern of the *myo-sex* gene, which is closely
310 linked to the SDR, in the recombinant and non-recombinant C42 strain individuals. The tests
311 revealed that, regardless of possessing the transgene or not, *myo-sex* was absent in females
312 and present in males (Figure 2C). This result indicates distal (telomeric side) location of the C42
313 transgene relative to *Nix*, and is consistent with the suggestion that *myo-sex* is located closer to
314 the centromere than *Nix* (HALL *et al.* 2014).

315

316 **Developmental abnormalities revealed by crossing experiments**

317 We randomly selected two females, which acquired the EGFP marker through independent
318 recombination events, to establish mosquito lines A and B. According to the Mendelian
319 inheritance half of such females' progeny should be EGFP-positive, regardless of sex, and both
320 sexes should be present in approximately equal proportions, as is the case with broods between
321 the non-recombinant C42 strain individuals. However, the observed results deviated from the
322 expectations.

323 Line A females produced progeny with a distorted sex ratio (1♂:2♀). Whereas half of the
324 F₁ females carried the transgene, only two out of over 1,000 F₁ males did so (Table 2, Cross 1
325 and 2; see also Supplementary Figure S3), which indicates that nearly all the EGFP-positive
326 males were inviable. The same inheritance pattern was reproduced in consecutive generations,
327 wherever line A females were used for crosses (data not shown). In each generation, larval and
328 pupal mortality was low and comparable to mortality in the wild-type or C42 strain mosquitoes
329 reared under the same conditions. Thus, the EGFP-positive males failed to develop past the
330 embryonic stage. The observed lethal effect must have been caused by a factor on the maternal
331 recombinant (transgene-carrying) m-chromosome. Yet, from crosses between EGFP-positive
332 mothers and fathers lacking EGFP we recovered two males that were EGFP-positive (Table 2,

333 cross 1 and 2). We suggest that the two surviving males acquired the transgene on the *m*-
334 chromosome that had undergone a secondary crossing-over during female meiosis (cf.
335 Supplementary Figure S3), during which the male-lethal factor had been purged, and which
336 rescued male viability. A very low number of these GFP-positive males is consistent with a
337 notion of a secondary recombination. Mutations inactivating the lethal factor (assuming that the
338 lethality is caused by gene expression) could have the same effect; however, such mutations
339 would occur with a frequency orders of magnitude lower than the frequency of the EGFP-
340 positive males or other “unexpected” rare phenotypes observed during this study (cf.
341 Supplementary Figure S3).

342 Availability of viable males carrying the *m*-linked EGFP marker created an opportunity to
343 produce a line homozygous for transgene insertion. The attempts to do so were unsuccessful,
344 but they provided further insight into the SDR neighbourhood (Table 2; Supplementary Figure
345 S3). Among cross 4 progeny we expected equal sex ratios and homozygosity for transgene
346 insertion in half of the females. Instead, there was an excess of males and apparently there
347 were no homozygous females, because none of the females taken from cross 4 progeny
348 produced EGFP-only broods (Table 2; probability of drawing 12 heterozygotes in a row for cross
349 5 from an equal mixture of homo- and heterozygotes, $P = 0.0002$). Female deficiency in cross 4
350 progeny was caused by embryonic lethality, similar to the male bias described above. Further,
351 the expectation of equal proportion of homo- and heterozygotes among male progeny from
352 cross 5 was not met. None of the males sampled from that pool (for cross 6) produced EGFP-
353 only females, which indicated that all tested males were heterozygous, with the transgene linked
354 to the M-chromosome (Table 2, cross 6; probability of drawing 22 heterozygotes in a row for
355 cross 6 from an equal mixture of homo- and heterozygotes $P = 2.4 \times 10^{-7}$). Intriguingly, in the
356 majority of cross 6 families there was 2♂:1♀ sex bias caused by female embryo lethality. When
357 individuals from families with the biased sex ratios were inbred (Table 2, cross 7), or females
358 from such families were crossed with wild-type males (Table 2, cross 8), approximately half of

359 the resulting families exhibited 2♂:1♀ bias. In contrast, none of the males from the affected
360 families, when crossed with wild-type females, sired progeny deviating from parity between the
361 sexes (Table 2, cross 9).

362 In contrast to line A females, the progeny of line B females consisted of all expected
363 phenotypes, including males with the inherited *m*-linked EGFP marker (Table 2; Supplementary
364 Figure S4). Crossing such males with the wild-type females yielded an excess of males,
365 resulting from the 2♂:1♀ sex ratio caused by female embryonic lethality in 5 out of 13 analyzed
366 families (Table 2, females cross ii). We used the EGFP-positive progeny of line B females in a
367 further attempt to establish a line homozygous for the transgene insertion. It proved
368 unsuccessful, but revealed another phenotypic effect. While homozygous males developed
369 normally and survived to adulthood (Table 2, cross iii), homozygous females could not complete
370 emergence and, almost invariably, died while attempting to leave pupal exuvium (Table 2, cross
371 iv; Supplementary Figure S5). They usually had protracted larval development, with up to twice
372 the length of the fourth instar larva period, compared to heterozygotes or wild-type females
373 (data not shown).

374

375 DISCUSSION

376 Availability of an *A. aegypti* strain with an EGFP-tagged transgene inserted near the SDR
377 allowed us to easily detect crossing-over events close to the SDR boundary and to trace the
378 inheritance pattern of the fluorescent marker in relation to sex. The initial crosses involving
379 rarely occurring females possessing a recombinant EGFP-positive *m*-chromosome produced
380 broods lacking one of the anticipated phenotypes. Prompted by this surprising result, we
381 conducted further crossing experiments and identified additional phenotypic effects, with a
382 stable 1:2 sex bias driven by lethality of either males or females in all or half of the families
383 analysed.

384 Our results have clear parallels with an earlier report of naturally occurring 1:2 sex
385 biases (WOOD 1975). However, in that study the skewed sex ratios were caused by larval
386 mortality, and the male lethal locus was at a considerable distance (5-10 cM) from the sex locus
387 (WOOD 1975), indicating that different lethal factors were involved. The lethality observed in our
388 study is, in most cases, attributable to the recombinant EGFP-positive m-chromosome. To
389 explain these phenomena, we suggest that within the SDR neighbourhood there are several
390 factors, which in the affected sex are likely lost or gained through recombination, leading to
391 death (Figure 3). These may include genes (or groups of tightly linked genes) vital for sex-
392 specific development, genes carrying recessive lethal mutations, or genes that are sexually
393 antagonistic, causing highly deleterious conflict that has been resolved by tight linkage with the
394 sex locus. The hypothetical scenarios of lethality are presented in more detail below. Other,
395 more complex schemes are possible, but not discussed in this study. The scenarios presented
396 here assume that, within the sex locus and its close neighborhoods encompassing the
397 recombination region, the M- and the m-chromosomes diverged sufficiently to share no
398 functional alleles of the lethal factors considered in this study.

399

400 **Male lethality**

401 The 1♂:2♀ bias in line A female crosses (Table 2, cross 1 and 2) may be due to either a male-
402 specifically expressed or recessive lethal gene normally effectively male-specific because of its
403 close linkage to the SDR, but transferred through recombination onto the m-chromosome.
404 Female carriers of such a recombinant chromosome would be viable and fertile. However, half
405 of their male progeny could be inviable for three potential reasons. (1) If the exchanged portion
406 of the chromosome harbors a gene expressed only in male embryos, the lethality may be
407 caused by a dosage-dependent deleterious effect. In males inheriting the recombinant m-
408 chromosome, the gene would be present in a double dose (one copy on the M-chromosome
409 and one copy on the m-chromosome; Figure 3C), which would lead to its overexpression and, in

410 effect, embryonic lethality. A study on *Drosophila*, in which duplication of two short X
411 chromosome regions led to male embryonic lethality (VENKEN *et al.* 2010), lends support to this
412 scenario. (2) If a recessive embryonic lethal gene is involved, males inheriting the maternal
413 recombinant chromosome would be homozygous and inviable (Figure 3C). Loci closely linked to
414 the SDR are kept effectively heterozygous. Consequently, recessive lethal mutations arising in
415 such SDR-linked regions would be sheltered by the wild-type alleles on the m-chromosome, and
416 they may be readily fixed in a population (MULLER 1932; NEI 2013). Linkage to the SDR would
417 result in male lethal effect, even if the recessive lethal gene was non-sex-specifically expressed.
418 It is possible that expressed pseudogenes, which are a product of sequence degeneration
419 around the M factor, are involved, if two copies are sufficient to cause deleterious effects. There
420 is growing evidence for the causal link of pseudogene expression with disease (POLISENO *et al.*
421 2015). (3) Lethality could be caused by recombination-mediated loss of a gene, loss of gene
422 expression, or gene inactivation. Males inheriting a recombinant m-chromosome would have a
423 segment of the M-chromosome duplicated, but would lack a portion of the m-chromosome
424 located close to the sex locus. If the missing non-sex-specific region carries a gene essential
425 and haplosufficient for embryo development, individuals lacking the region would die (Figure
426 3D). The same lethal effect would occur if the essential haplosufficient gene was disrupted by
427 the recombination breakpoint and, in result, was not transcribed or produced incomplete
428 transcripts (Figure 3E). Alternatively, the m-linked haplosufficient gene could be silenced due to
429 position effect variegation (SCHOTTA *et al.* 2003), providing the recombination brought the gene
430 close to and under the influence of the heterochromatin from the SDR neighborhood (Figure
431 3F).

432

433 **Female lethality**

434 The female lethality observed in this study involves likely three different factors and at least two
435 alternative mechanisms. The 2♂:1♀ bias in cross 4 progeny (Table 2; Line A-derived

436 recombinants) can be explained by scenario (3) described above, with homozygous females
437 possessing two copies of an M-chromosome segment, and thus, lacking a corresponding *m*-
438 linked region carrying a haplosufficient gene essential for female embryonic development
439 (Figure 3G). Likewise, lethality can be caused by inheriting two copies of a non-functional
440 haplosufficient gene (analogous to situation depicted in Figures 3E and 3F). A lack of a
441 haplosufficient gene or inheritance of nonfunctional gene copies could also have driven late
442 lethality (during eclosion) of line B-derived homozygous females (Table 2, Line B females cross
443 iv). Alternatively, the SDR-linked female-specific recessive lethal genes recombined into the *m*-
444 chromosome could also be responsible (Figure 3H). Similarly, the SDR-linked sexually
445 antagonistic genes could produce observed developmental abnormalities, if sexual conflict was
446 not strong in females heterozygous for the antagonistic genes, but sufficiently deleterious to
447 lead to death of homozygous females (Figure 3H). Such an explanation is consistent with the
448 prediction that the region around the maleness gene may be highly detrimental if recombined
449 into females (JORDAN and CHARLESWORTH 2012). In this context we tested whether the *myo-sex*
450 gene residing close to the SDR and strongly expressed in male pupae might be responsible for
451 the female late lethality detected in our study. We found that *myo-sex* segregates with
452 maleness, rather than with the C42 transgene, consistent with its location on the opposite site of
453 the SDR relative to the transgene. As such, *myo-sex* cannot be a female-lethal factor in this
454 case. Finally, as mentioned earlier, it is possible that deleterious effects are exerted by
455 expressed pseudogenes (Figure 3H).

456 A different mechanism must have led to female deficiency in some line B-derived cross ii
457 families (Table 2). In that case, inheritance of a recessive female-lethal gene on the *m*-
458 chromosome from a wild-type heterozygous mother and on the paternal recombinant *m*-
459 chromosome could have led to inviability of the resulting homozygous females (Figure 3I). A
460 similar or a different interaction could have led to 2♂:1♀ bias in the line A-derived families from
461 crosses 6, 7 and 8 (Table 2), in which the inviable females inherited the EGFP-less, apparently

462 non-recombinant first chromosomes from both parents. These interactions may have involved
463 epigenetic phenomena, such as paramutations (HOLLICK 2010), that lead to heritable altered
464 gene expression states, whose toxic effects could result in non-Mendelian inheritance patterns.
465 Currently we do not have sufficient data to offer a plausible interpretation of these results, but
466 they indicate the existence within the sex locus neighbourhood of additional loci affecting female
467 development.

468

469 **Implausible scenarios of sex-specific lethality**

470 Sex-specific embryonic lethality in *Drosophila* is almost exclusively linked to misregulation of
471 dosage compensation machinery (CLINE and MEYER 1996). Similar mechanisms of lethality are
472 rather unlikely to be behind the phenomena described here, because dosage compensation
473 appears not to exist in *A. aegypti* (HALL *et al.* 2015). Likewise, abnormalities in chromosomal
474 transmission caused by transgene-linked inversions (cf. MCGIVERN and RAI 1974) lend no
475 plausible explanation. In our study, sex bias was observed in the progeny of the EGFP-positive
476 females (but not in the C42 strain males), and was associated with inheriting the recombinant
477 maternal (non-sex-specific) m-chromosome. Thus, even if an inversion was present on such an
478 m-chromosome, a crossover in the inversion loop during female meiosis would not yield
479 detectable sex bias, because loss of gametes would equally affect male and female progeny.
480 Moreover, published evidence argues against existence of any larger inversions encompassing
481 the SDR that might cause loss of one type of gametes, because individuals carrying a
482 recombinant sex chromosome, with a recombination breakpoint close to either side of the SDR,
483 produce progeny with all expected phenotypes, when crossed to wild-type mosquitoes (HALL *et*
484 *al.* 2014).

485

486 **Genes within the sex locus neighborhood**

487 The proposed scenarios dictate that the line A-derived viable males possessing the m-
488 chromosome-linked transgene could have originated only after secondary crossing-over events
489 during female meiosis, which eliminated the male embryonic-lethal factor from the
490 neighbourhood of the transgene. A male lacking such a lethal factor subsequently produced
491 embryonic-lethal homozygous females (Table 2, cross 4). Thus, the male-embryonic lethal gene
492 must be different from the female-embryonic lethal gene. In addition, lethality observed at two
493 distinct developmental stages suggests that at least two factors are responsible for
494 developmental abnormalities in females. We propose that all the loci involved are located on
495 chromosome 1 pair in the following linear order: *myo-sex*, *Nix*, male-embryonic lethal *E1* (likely
496 M-linked), female-embryonic lethal *E2* (m-linked, or sexually antagonistic M-linked), late
497 development female-lethal *E3* (m-linked, or sexually antagonistic M-linked), the transgene,
498 female-embryonic lethal *E4* (Supplementary Figure S6). Independent recombination events
499 within the sex locus neighbourhood, each with a single chromosomal breakpoint between
500 different loci, could have led to assortment or elimination of different lethal factors in the
501 progeny, eventually resulting in the observed alternative lethal phenotypes. This model implies a
502 persistent significant sequence similarity between the M and m haplotypes to promote
503 recombination, which may be mediated by repetitive elements.

504

505 **Evolutionary persistence of a short SDR**

506 The primary model for the evolution of sex chromosomes implicates genes with sexually
507 antagonistic properties as drivers of selection for suppressed recombination around the primary
508 sex-determining gene (FISHER 1931; CHARLESWORTH and CHARLESWORTH 1980; RICE 1987).
509 Numerous studies regarding theoretical considerations on sexual conflict and sex chromosomes
510 sharply contrast with a very limited empirical evidence for the existence of sexually antagonistic
511 loci. Among few well documented cases are genes increasing male reproductive success but
512 affecting survival, such as those encoding coloration in guppies and Lake Malawi cichlids

513 (LINDHOLM and BREDEN 2002; ROBERTS *et al.* 2009). However, virtually nothing is known about
514 sexually antagonistic effects caused by genes with sex differences in development. Our data
515 suggest that such genes are likely present in *Aedes*.

516 Our study shows that the SDR neighbourhood in *Aedes* is a fascinating test case for the
517 analysis of processes that shape boundaries of a non-recombining region stably maintained for
518 long evolutionary times. According to the models, spread of recombination suppression is fueled
519 by the availability of alleles with sexually antagonistic properties in the vicinity of the sex locus.
520 However, the process may cease, if the junctions with the SDR have high recombination rates
521 per physical distance (OTTO *et al.* 2011), or if the non-recombining region abuts a sequence of
522 sufficient recombinational length that fails to provide genetic variation for sexually antagonistic
523 traits (RICE 1987). Indeed, an extensive array of the ribosomal RNA genes adjacent to the SDR
524 in *Aedes* and *Culex* may constitute a sufficient barrier that prevents expansion of the non-
525 recombining region towards the telomere. Conversely, a cluster of lethal, likely sexually
526 antagonistic, genes positioned close to the boundary of the non-recombining region may
527 effectively guard against recombination between the neighborhoods of the sex locus and
528 prevent homogenization of the corresponding areas of the M- and m-chromosomes in a
529 population to protect the very small SDR from shrinking, which might be detrimental to the male
530 sex. The *A. aegypti* genome is not assembled in that region and no genetic markers located
531 between *Nix* and the transgene are currently available. Therefore, further considerable work is
532 needed to reveal the molecular identity of the lethal factors involved.

533

534 CONFLICT OF INTEREST

535 The authors declare no conflict of interest.

536

537 ACKNOWLEDGMENTS

538 Akio Mori provided the *A. aegypti* RED strain mosquitoes for genetic analysis. Kanchana
539 Panaram helped with scoring the phenotypes in some of the initial crosses. Luke Alpey, Gareth
540 Lycett and Dave Weetman provided valuable comments on the manuscript. We acknowledge
541 financial assistance for this work to JK from UNICEF/UNDP/World Bank/WHO Special
542 Programme for Research and Training in Tropical Diseases (TDR) grant A60345 and to ASR
543 from the National Institutes of Health grant R37 AI24716.

544 LITERATURE CITED

545

- 546 Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman, 1990 Basic local alignment
547 search tool. *J Mol Biol* 215: 403-410.
- 548 Arensburger, P., K. Megy, R. M. Waterhouse, J. Abrudan, P. Amedeo *et al.*, 2010 Sequencing
549 of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics.
550 *Science* 330: 86-88.
- 551 Bachtrog, D., 2013 Y-chromosome evolution: emerging insights into processes of Y-
552 chromosome degeneration. *Nat Rev Genet* 14: 113-124.
- 553 Bachtrog, D., M. Kirkpatrick, J. E. Mank, S. F. McDaniel, J. C. Pires *et al.*, 2011 Are all sex
554 chromosomes created equal? *Trends Genet* 27: 350-357.
- 555 Bian, G., S. W. Shin, H. M. Cheon, V. Kokoza and A. S. Raikhel, 2005 Transgenic alteration of
556 Toll immune pathway in the female mosquito *Aedes aegypti*. *Proc Natl Acad Sci U S A*
557 102: 13568-13573.
- 558 Bull, J., 1983 *Evolution of Sex Determining Mechanisms*. Benjamin Cummings, Menlo Park, CA.
- 559 Charlesworth, D., and B. Charlesworth, 1980 Sex differences in fitness and selection for centric
560 fusions between sex-chromosomes and autosomes. *Genet Res* 35: 205-214.
- 561 Cline, T. W., and B. J. Meyer, 1996 Vive la difference: males vs females in flies vs worms. *Annu*
562 *Rev Genet* 30: 637-702.
- 563 Craig, G. B., Jr., R. C. Vandehey and W. A. Hickey, 1961 Genetic variability in populations of
564 *Aedes aegypti*. *Bull World Health Organ* 24: 527-539.
- 565 Fisher, R. A., 1931 The evolution of dominance. *Biol. Rev.* 6: 345-368.
- 566 Gale, K., and J. Crampton, 1989 The ribosomal genes of the mosquito, *Aedes aegypti*. *Eur J*
567 *Biochem* 185: 311-317.
- 568 Hall, A. B., S. Basu, X. Jiang, Y. Qi, V. A. Timoshevskiy *et al.*, 2015 A male-determining factor in
569 the mosquito *Aedes aegypti*. *Science* 348: 1268-1270.
- 570 Hall, A. B., V. A. Timoshevskiy, M. V. Sharakhova, X. Jiang, S. Basu *et al.*, 2014 Insights into
571 the preservation of the homomorphic sex-determining chromosome of *Aedes aegypti*
572 from the discovery of a male-biased gene tightly linked to the M-locus. *Genome Biol Evol*
573 6: 179-191.
- 574 Hickey, W. A., and G. B. Craig, Jr., 1966 Genetic distortion of sex ratio in a mosquito, *Aedes*
575 *aegypti*. *Genetics* 53: 1177-1196.
- 576 Hollick, J. B., 2010 Paramutation and development. *Annu Rev Cell Dev Biol* 26: 557-579.
- 577 Horn, C., and E. A. Wimmer, 2000 A versatile vector set for animal transgenesis. *Dev Genes*
578 *Evol* 210: 630-637.
- 579 Jordan, C. Y., and D. Charlesworth, 2012 The potential for sexually antagonistic polymorphism
580 in different genome regions. *Evolution* 66: 505-516.
- 581 Kokoza, V., A. Ahmed, E. A. Wimmer and A. S. Raikhel, 2001 Efficient transformation of the
582 yellow fever mosquito *Aedes aegypti* using the *piggyBac* transposable element vector
583 pBac[3xP3-EGFP afm]. *Insect Biochem Mol Biol* 31: 1137-1143.
- 584 Kokoza, V., A. Ahmed, S. Woon Shin, N. Okafor, Z. Zou *et al.*, 2010 Blocking of *Plasmodium*
585 transmission by cooperative action of Cecropin A and Defensin A in transgenic *Aedes*
586 *aegypti* mosquitoes. *Proc Natl Acad Sci U S A* 107: 8111-8116.
- 587 Krzywinska, E., and J. Krzywinski, 2009 Analysis of expression in the *Anopheles gambiae*
588 developing testes reveals rapidly evolving lineage-specific genes in mosquitoes. *BMC*
589 *Genomics* 10: 300.
- 590 Kumar, A., and K. S. Rai, 1990 Chromosomal localization and copy number of 18S + 28S
591 ribosomal RNA genes in evolutionarily diverse mosquitoes (Diptera, Culicidae).
592 *Hereditas* 113: 277-289.

593 Kumar, A., and K. S. Rai, 1993 Molecular organization and evolution of mosquito genomes.
594 Comp Biochem Physiol B 106: 495-504.

595 Li, X., N. Lobo, C. A. Bauser and M. J. Fraser, Jr., 2001 The minimum internal and external
596 sequence requirements for transposition of the eukaryotic transformation vector
597 piggyBac. Mol Genet Genomics 266: 190-198.

598 Lindholm, A., and F. Breden, 2002 Sex chromosomes and sexual selection in poeciliid fishes.
599 Am Nat 160 Suppl 6: S214-224.

600 McClelland, G. A., 1966 Sex-linkage at two loci affecting eye pigment in the mosquito *Aedes*
601 *aegypti* (Diptera: Culicidae). Can J Genet Cytol 8: 192-198.

602 McGivern, J. J., and K. S. Rai, 1974 Sex-ratio distortion and directed alternate segregation of
603 interchange complexes in a mosquito. J Hered 65: 71-77.

604 Mori, A., D. W. Severson and B. M. Christensen, 1999 Comparative linkage maps for the
605 mosquitoes (*Culex pipiens* and *Aedes aegypti*) based on common RFLP loci. J Hered
606 90: 160-164.

607 Muller, H. J., 1932 Further studies on the nature and causes of gene mutations., pp. 213–255.
608 in *Proc. 6th Int. Congr. Genet.*, edited by D. F. Jones, Ithaca, NY.

609 Nei, M., 2013 *Mutation-driven evolution*. Oxford University Press, Oxford.

610 Nene, V., J. R. Wortman, D. Lawson, B. Haas, C. Kodira *et al.*, 2007 Genome sequence of
611 *Aedes aegypti*, a major arbovirus vector. Science 316: 1718-1723.

612 Otto, S. P., J. R. Pannell, C. L. Peichel, T. L. Ashman, D. Charlesworth *et al.*, 2011 About PAR:
613 the distinct evolutionary dynamics of the pseudoautosomal region. Trends Genet 27:
614 358-367.

615 Palavesam, A., C. Esnault and D. A. O'Brochta, 2013 Post-integration silencing of piggyBac
616 transposable elements in *Aedes aegypti*. PLoS One 8: e68454.

617 Poliseno, L., A. Marranci and P. P. Pandolfi, 2015 Pseudogenes in Human Cancer. Front Med
618 (Lausanne) 2: 68.

619 Reidenbach, K. R., S. Cook, M. A. Bertone, R. E. Harbach, B. M. Wiegmann *et al.*, 2009
620 Phylogenetic analysis and temporal diversification of mosquitoes (Diptera: Culicidae)
621 based on nuclear genes and morphology. BMC Evol Biol 9: 298.

622 Rice, W. R., 1987 The accumulation of sexually antagonistic genes as a selective agent
623 promoting the evolution of reduced recombination between primitive sex chromosomes.
624 Evolution 41: 911-914.

625 Roberts, R. B., J. R. Ser and T. D. Kocher, 2009 Sexual conflict resolved by invasion of a novel
626 sex determiner in Lake Malawi cichlid fishes. Science 326: 998-1001.

627 Sambrook, J., E. F. Fritsch and T. Maniatis, 1998 *Molecular Cloning: A Laboratory Manual*. Cold
628 Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

629 Schotta, G., A. Ebert, R. Dorn and G. Reuter, 2003 Position-effect variegation and the genetic
630 dissection of chromatin regulation in *Drosophila*. Semin Cell Dev Biol 14: 67-75.

631 Severson, D. W., 1997 RFLP analysis of insect genomes, pp. pp. 309–320 in *The Molecular*
632 *Biology of Insect Disease Vectors*, edited by J. M. Crampton, B. J. Beaty and C. Louis.
633 Chapman&Hall, London.

634 Severson, D. W., A. Mori, Y. Zhang and B. M. Christensen, 1993 Linkage map for *Aedes*
635 *aegypti* using restriction fragment length polymorphisms. J Hered 84: 241-247.

636 Song, J., F. Dong, J. W. Lilly, R. M. Stupar and J. Jiang, 2001 Instability of bacterial artificial
637 chromosome (BAC) clones containing tandemly repeated DNA sequences. Genome 44:
638 463-469.

639 Timoshevskiy, V. A., N. A. Kinney, B. S. deBruyn, C. Mao, Z. Tu *et al.*, 2014 Genomic
640 composition and evolution of *Aedes aegypti* chromosomes revealed by the analysis of
641 physically mapped supercontigs. BMC Biol 12: 27.

642 Timoshevskiy, V. A., D. W. Severson, B. S. DeBruyn, W. C. Black, I. V. Sharakhov *et al.*, 2013
643 An integrated linkage, chromosome, and genome map for the yellow fever mosquito
644 *Aedes aegypti*. PLoS Negl Trop Dis 7: e2052.
645 Venken, K. J., E. Popodi, S. L. Holtzman, K. L. Schulze, S. Park *et al.*, 2010 A molecularly
646 defined duplication set for the X chromosome of *Drosophila melanogaster*. Genetics
647 186: 1111-1125.
648 Veyrunes, F., P. D. Waters, P. Miethke, W. Rens, D. McMillan *et al.*, 2008 Bird-like sex
649 chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome
650 Res 18: 965-973.
651 Vicoso, B., J. J. Emerson, Y. Zektser, S. Mahajan and D. Bachtrog, 2013 Comparative sex
652 chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global
653 dosage compensation. PLoS Biol 11: e1001643.
654 Wood, R. J., 1975 Lethal genes on the sex chromosomes concealed in a population of the
655 mosquito *Aedes aegypti* L. Genetica 46: 49-66.
656

657 Table 1 Results of crossing experiments to establish linkage group of the transgene
 658 insertion in the C42 strain.

659

| Cross | Male | | | | Female | | | |
|--|---------|-----------------|----------------|----------------|----------------|-----|-----------------|-----|
| | EGFP(+) | | EGFP(-) | | EGFP(+) | | EGFP(-) | |
| | black | red | black | red | black | red | black | red |
| RED ♀ X C42 ♂ | 652 | - | 3 ^a | - | 1 ^a | - | 695 | - |
| RED ♀ X F ₁ (RED ♀ X C42 ♂) ♂ | 878 | 20 ^b | - | 3 ^a | 2 ^a | - | 25 ^b | 787 |

660

661 Figures represent numbers of progeny individuals with a given phenotype. Mosquitoes were
 662 scored at the pupal stage for sex, EGFP expression (+, EGFP-positive; -, EGFP-negative) and
 663 eye color. The RED strain individuals are homozygous for a recessive red eye color mutation.
 664 Black denotes a wild-type eye phenotype.

665 ^a Individuals carrying a recombinant chromosome 1 with the breakpoint between the sex locus
 666 and the transgene.

667 ^b Individuals carrying a recombinant chromosome 1 with the breakpoint between the sex locus
 668 and the *re* locus.

669

670

671 Table 2 Summary of crossing experiments using recombinant chromosome 1-carrying female
 672 lines and their derivatives (see also Supplementary Figure S3 and S4).
 673

| Cross | Genotype ^a | Families with 1:2 sex bias ^b | Male | | Female | |
|------------------------------------|---|--|-------------------------|---------|-------------------------|-------------------------|
| | | | EGFP(+) | EGFP(-) | EGFP(+) | EGFP(-) |
| <u>Line A females</u> | | | | | | |
| 1 | C42A EGFP ♀ X wt ♂ | 22/22 F | 1 ^{(3) c} | 458 | 452 | 394 |
| 2 | C42A EGFP ♀ X C42 (-) ♂ | 26/26 F | 1 | 623 | 636 | 644 |
| <u>Line A-derived recombinants</u> | | | | | | |
| 3 | C42A EGFP ♀ X Cross 2 F ₁ EGFP ♂ | 9/9 F | 2 ⁽⁴⁾ | 196 | 417 ⁽⁴⁾ | 2 |
| 4 | Cross 3 F ₁ EGFP ♀ X Cross 3 F ₁ EGFP ♂ | 3/3 M | 54 | 38 | 31 ⁽⁵⁾ | - |
| 5 | Cross 4 F ₁ EGFP ♀ X C42 EGFP ♂ | 1/12 F | 499 ⁽⁶⁾ | 1 | 282 | 239 ⁽⁶⁾ |
| 6 | Cross 5 F ₁ (-) ♀ X Cross 5 F ₁ EGFP ♂ | 20/22 M | 761 ^{(7, 9) d} | 1 | 1 | 398 ^{(7, 8) d} |
| 7 | Cross 6 F ₁ (-) ♀ X Cross 6 F ₁ EGFP ♂ | 21/40 M | 1468 | 1 | 3 | 981 |
| 8 | Cross 6 F ₁ (-) ♀ X wt ♂ | 15/35 M | - | 1492 | - | 1055 |
| 9 | wt ♀ X Cross 6 F ₁ EGFP ♂ | 0/18 | 569 | 4 | 1 | 579 |
| <u>Line B females</u> | | | | | | |
| i | C42B EGFP ♀ X C42 (-) ♂ | 0/3 | 53 ^(ii, iv) | 43 | 46 ^(ii, iii) | 55 |
| ii | wt ♀ X Cross i F ₁ EGFP ♂ | 5/13 M | 1 | 422 | 306 | 3 |
| iii | Cross i F ₁ EGFP ♀ X C42 EGFP ♂ | 0/10 | 585 ^e | - | 271 | 269 |
| iv | Cross i F ₁ EGFP ♀ X Cross i F ₁ EGFP ♂ | 0/9 | 159 | 147 | 274 ^f | - |

674
 675 Figures represent cumulative numbers of progeny from a given cross, counted at the pupal
 676 stage. The observed 1:2 sex biases reveal sex-specific developmental abnormalities mediated
 677 by different factors linked to the sex locus.

678 ^a The EGFP-positive and the EGFP-negative individuals used in crosses are denoted as EGFP
 679 and (-), respectively. The C42 (-) males are the C42 strain-derivatives, which inherited a
 680 paternal *M* allele-bearing chromosome 1 that lost EGFP through recombination (cf. Figure 2).
 681 The wild-type (wt) individuals were from the Rockefeller strain.

682 ^b Number of families with the sex bias out of the total number of families studied; a letter
 683 following the numbers denotes excess of either females (F) or males (M) in the affected
 684 families.

685 ^c Superscript numbers in parentheses indicate cross, for which individuals from a given
686 phenotype were taken.

687 ^d Only individuals from families with sex bias were taken for crosses.

688 ^e Of 303 male pupae examined, 151 were scored as homozygotes and 152 as heterozygotes for
689 transgene insertion; 8% and 3% of those, respectively, died during eclosion.

690 ^f Of 186 female pupae examined, 97 were scored as homozygotes and 89 as heterozygotes for
691 transgene insertion; 89% and 4% of those, respectively, died during eclosion. Small difference
692 in fluorescence intensity can make distinction of homo- and heterozygotes ambiguous;
693 therefore, females from the former group that survived to adulthood are likely heterozygotes
694 incorrectly scored as homozygotes.

695

696 **Titles and legends to figures**

697

698 **Figure 1** The *piggyBac* transgene integration into the *A. aegypti* C42 strain genome. (A) A map
699 of the construct flanked by the identified genomic DNA (not drawn to scale). Approximate
700 positions of inverse PCR primers used to isolate the junctions between the transgene and
701 flanking genomic DNA (thick zigzag lines) are marked directly above the map. Primer names
702 are abbreviated for clarity; see Table S1 for complete names and sequences of the primers.
703 Target sites of the restriction enzymes used to generate templates for inverse PCR are marked
704 as vertical bars below the map. Primers used to verify integrity of the junctions are shown at the
705 top of the figure. Target sites of the endonucleases used for Southern blot analysis are marked
706 by long vertical arrows below the map; for clarity target sites for different enzymes are
707 represented by different arrow styles. Solid horizontal line represents a DNA fragment used as a
708 probe in Southern blot analysis. (B) Southern blot analysis of the C42 males carrying the EGFP
709 tag. Names of the restriction enzymes used for DNA digestion and sizes of the hybridizing
710 fragments (in kb) are shown above the lanes. The position of the HindIII digested Lambda DNA
711 fragments used as a high molecular weight marker is indicated on the left. (C) A PCR confirming
712 integrity of the junctions in the EGFP-positive individuals. See Materials and Methods for primer
713 combinations used. Quality of the template DNA was evaluated using the EGFP-specific
714 primers (control).

715

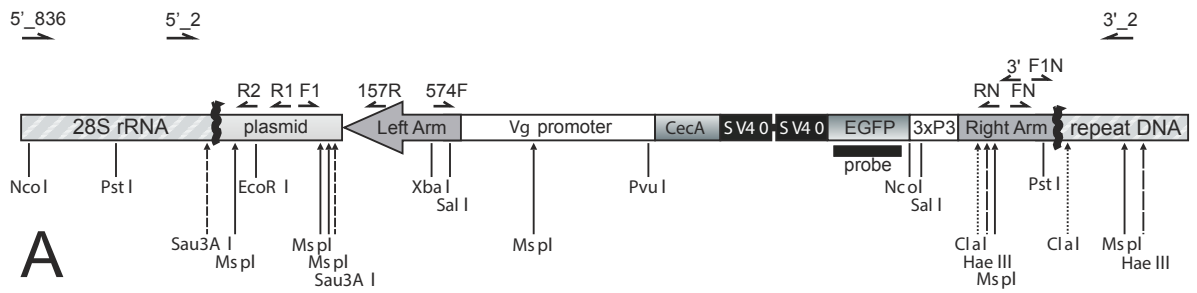
716 **Figure 2** Loss of linkage between the maleness locus and the EGFP tag during male meiosis.
717 (A) A diagram of a crossover between two non-sister chromatids during meiotic prophase,
718 giving rise to recombinant chromosomes (only pericentromeric fragments of chromosomes 1 are
719 presented). The sex locus (black line or square denoted either *m* or *M*) and the transgene (white
720 square denoted *G*) are shown. (B) A PCR confirming integrity of the junctions in the EGFP-
721 positive individuals and indicating loss of the transposon in males and gain thereof in females in

722 a generation following the recombination. The phenotypes of individuals analyzed are given at
723 the top. For each phenotype lanes are marked as follows: 5', a product spanning the 5' junction;
724 G, a fragment of the *EGFP* gene; 3', a product spanning the 3' junction; +, a fragment of the
725 *Aams* gene (positive control of DNA quality). A smeary ladder-like pattern of the PCR product
726 spanning the 3' junction results from binding of the near_pBac3'_2 primer to multiple target sites
727 within a tandemly repeated flanking DNA. Combinations of chromosome 1 pairs corresponding
728 to each phenotype are depicted below the gel image. (C) Test for presence of *myo-sex* in the
729 EGFP-positive and EGFP-negative individuals, denoted as G+ or G-. M, male; F, female; (-),
730 negative control.

731

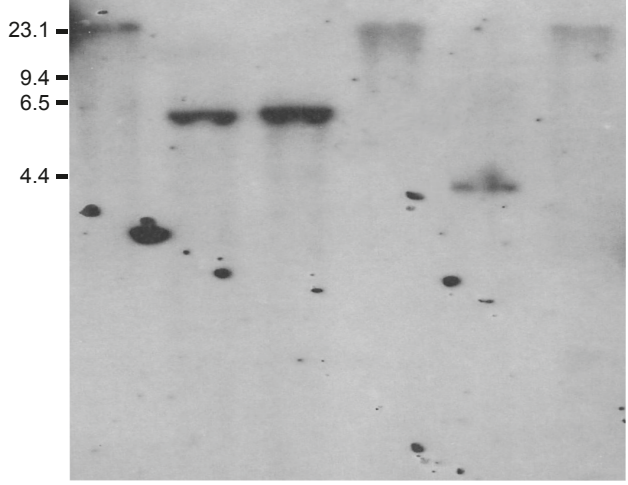
732 **Figure 3** Hypothetical mechanisms of sex-specific lethality observed in this study. (A) Crossing-
733 over during male meiosis results in decoupling of the transgene from the SDR. Following
734 assortment of gametes, the recombinant chromosomes are transmitted to the F₁ generation, in
735 which females inherit EGFP along with a portion of the M-chromosome (shown on the right). (B)
736 Females carrying a recombinant m-chromosome produce progeny with four potential karyotypes
737 (pictured), including a lethal karyotype (crossed; for clarity karyotypes sired by wild-type males
738 are shown). Individuals with the lethal karyotype carry on the m-chromosome a duplicated
739 portion of the M-chromosome region (marked by a curly bracket), which is tightly linked to the
740 SDR and normally present exclusively in a single copy in males. (C-F) Potential mechanisms of
741 male lethality. (C) If the duplicated region harbors a male-specifically expressed embryonic
742 gene *E1*, the male lethality (Table 2, Line A females cross 1 and 2) may be caused by a
743 dosage-dependent deleterious effect (overexpression from two copies). Alternatively, the male
744 lethality may result from a recessive lethal gene *E1*, or from expressed pseudogenes. (D) The
745 EGFP-positive males inheriting the transgene on the m-chromosome lack a non-sex specific
746 chromosomal region located close to the m locus. If the missing region carries a haplosufficient
747 gene essential for embryo development (missing gene *E1*), individuals with such a

748 chromosomal arrangement would be inviable. (E) If the essential haplosufficient gene *E1* is
749 disrupted by the recombination breakpoint, it would not be transcribed or it would produce
750 incomplete transcripts. (F) Positioning of the m-linked haplosufficient gene *E1* under the
751 influence (depicted as a wavy arrow) of the heterochromatin from the SDR neighborhood could
752 lead to silencing due to position effect variegation. (G-I) Potential mechanisms of female
753 lethality. (G) Lack of non-sex specific region and a concomitant lack of haplosufficient genes
754 *E2/3* residing in that region may drive lethality of homozygous females at embryo (*E2*) or late
755 (*E3*) stage (cf. Table 2, Line A-derived recombinants cross 4 and Line B females cross iv). (H)
756 The SDR-linked region recombined into the m-chromosome could contain female-specific
757 recessive lethal genes, sexually antagonistic genes, or deleterious expressed pseudogenes
758 *E2/3*. (I) A recessive lethal m-linked gene *E4* located close to the sex locus may drive female-
759 specific embryonic lethality not related to recombination (cf. Table 2, Line B females cross ii).
760 Presence or absence of hypothetical genes, that may cause lethal effects, is denoted by black
761 solid circles and dashed line circles, respectively.

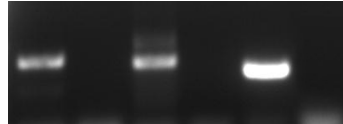


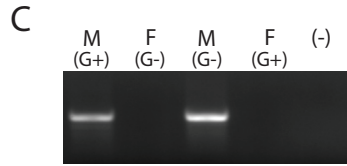
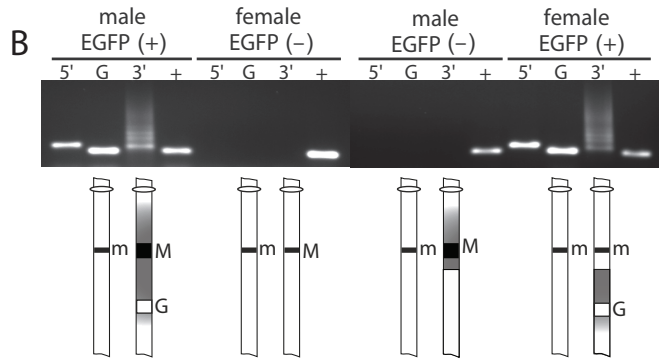
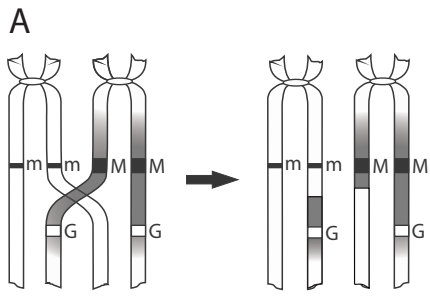
A

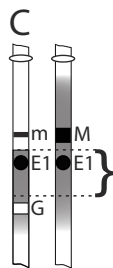
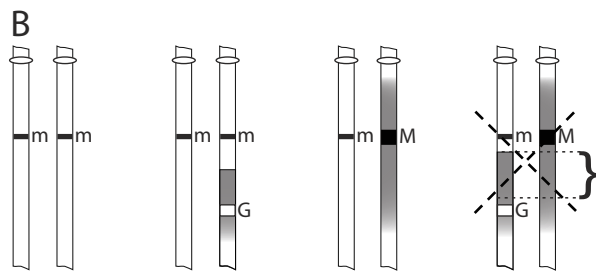
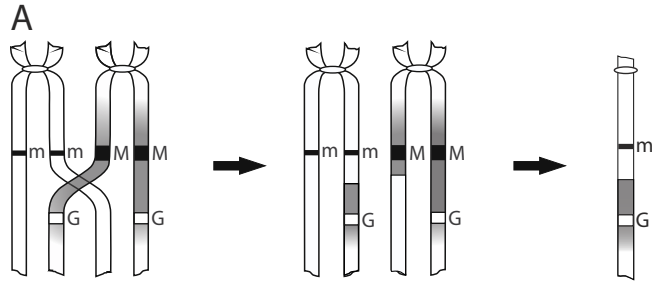
B *EcoRI* ~30 *NcoI* 6.1 *PstI* 6.3 *PvuI* ~30 *SalI* 4.0 *XbaI* ~30



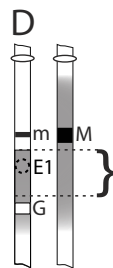
C 5' flank 3' flank control
M (-) M (-) M (-)



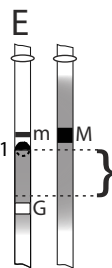




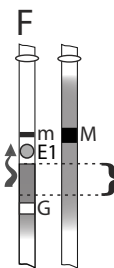
overexpression
or
recessive lethal
or
expressed pseudogene



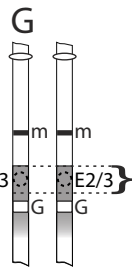
absence of essential
haplosufficient
gene



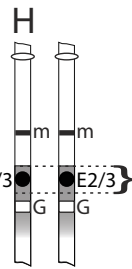
loss of essential
haplosufficient
gene expression



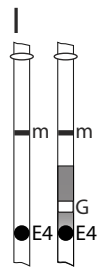
silencing of essential
haplosufficient
gene through
position effect
variegation



absence of essential
haplosufficient
gene



recessive lethal
or
sexually antagonistic
gene
or
expressed pseudogene



recessive lethal