

1 **Set2 and H3K36 regulate the *Drosophila* male X chromosome in a context-specific**
2 **manner, independent from MSL complex spreading**

3

4

5 Harmony R. Salzler¹, Vasudha Vandadi¹, and A. Gregory Matera^{1,2,3,4*}

6

7 *Address correspondence to: matera@unc.edu

8

9 ¹Integrative Program for Biological and Genome Sciences, University of North Carolina,
10 Chapel Hill, NC, USA.

11 ²Department of Genetics, University of North Carolina, Chapel Hill, NC, USA.

12 ³Department of Biology, University of North Carolina, Chapel Hill, NC, USA

13 ⁴RNA Discovery and Lineberger Comprehensive Cancer Centers, University of North
14 Carolina, Chapel Hill, NC, USA.

15

16

17 **Keywords:** Histone lysine methyltransferase, HMT, KMT; SET domain; chromatin;
18 SETD2; histone variant; H3.3; dosage compensation; post-translational modification, PTM

19

20

21

22 **Abstract**

23 Dosage compensation in *Drosophila* involves upregulating male X-genes two-fold. This
24 process is carried out by the MSL (male-specific lethal) complex, which binds high-affinity
25 sites and spreads to surrounding genes. Current models of MSL spreading focus on
26 interactions of MSL3 (male-specific lethal 3) with histone marks; in particular, Set2-
27 dependent H3 lysine-36 trimethylation (H3K36me3). However, Set2 might affect DC via
28 another target, or there could be redundancy between canonical H3.2 and variant H3.3
29 histones. Further, it is difficult to parse male-specific effects from those that are simply X-
30 specific. To discriminate among these possibilities, we employed genomic approaches in
31 *H3K36* (residue) and *Set2* (writer) mutants. The results confirm a role for Set2 in X-gene
32 regulation, but show that expression trends in males are often mirrored in females. Instead
33 of global male-specific reduction of X-genes in *Set2/H3K36* mutants, the effects were
34 heterogeneous. We identified cohorts of genes whose expression was significantly altered
35 following loss of H3K36 or Set2, but the changes were in opposite directions, suggesting
36 that H3K36me states have reciprocal functions. In contrast to *H4K16R* controls, analysis of
37 combined *H3.2K36R/H3.3K36R* mutants neither showed consistent reduction in X-gene
38 expression, nor any correlation with MSL3 binding. Examination of other developmental
39 stages/tissues revealed additional layers of context-dependence. Our studies implicate
40 BEAF-32 and other insulator proteins in Set2/H3K36-dependent regulation. Overall, the
41 data are inconsistent with the prevailing model wherein H3K36me3 directly recruits the
42 MSL complex. We propose that Set2 and H3K36 support DC indirectly, via processes that
43 are utilized by MSL but common to both sexes.

44

45

46 Introduction

47 The evolution of heterogametic sexes necessitates that the number of X chromosome
48 transcripts from XY males and XX females be equalized to prevent maladaptive disparities
49 in gene dosage. In mammals, this dosage compensation (DC) system involves stochastic
50 inactivation of one female X chromosome [1]. In contrast, *Drosophila melanogaster* relies on
51 a roughly 2-fold upregulation of transcripts generated from the male X. Importantly, many
52 elements of the *Drosophila* DC system are conserved in mammals [2], and relevant to
53 human health and disease research [3-5].

54 The most extensively studied mediator of DC in *Drosophila* is the Male-Specific
55 Lethal (MSL) complex, which carries out histone H4 lysine 16 acetylation (H4K16ac),
56 primarily on the male X [6, 7]. One estimate suggests that the MSL complex accounts for
57 ~40-50% of the upregulation of the male X [8]. Genetic mutations in MSL complex members
58 demonstrate that it is essential for male survival [9-11]. Current evidence supports
59 involvement of the MSL complex in regulating RNA polymerase II elongation [12-14] as
60 well as in genome organization [14-18]. Importantly, recent work also demonstrates that
61 the H4K16 residue itself is essential in male flies, and that the H4K16 acetylation function
62 of the MSL complex is crucial [19, 20].

63 The core MSL complex is comprised of five proteins (MSL1, MSL2, MSL3, MLE, and
64 MOF) and two lncRNAs (*roX1* and *roX2*) [14, 21]. Four of the five MSL proteins are also
65 present in females, excepting MSL2 [22, 23]. The MOF acetyltransferase, which catalyzes
66 acetylation of H4K16ac, also acts on housekeeping genes throughout the genome in the
67 context of the non-specific lethal (NSL) complex [24]. The distributions of H4K16ac
68 resulting from these two complexes are distinct, as MSL acetylates over gene bodies,

69 whereas NSL preferentially targets promoters [25, 26]. Other MSL-interacting proteins
70 have been identified, [27], many of which have substantiated roles in DC [27-30].

71 Current models of MSL function posit that the complex is initially targeted to the
72 male X via binding of MSL2•MSL1 dimers to high-affinity binding sites (HASs), followed
73 by subsequent spreading to nearby genes [31](for reviews see [14, 21]). The CLAMP protein
74 is an important cofactor for MSL2•MSL1 binding [32, 33], although CLAMP-independent
75 binding to a small subset of so-called PionX (pioneering on the X) sites is required for initial
76 recognition of the male X [34]. Following initial targeting, MSL activity spreads to
77 surrounding active genes by way of the MSL3 chromodomain [35, 36].

78 To date, our understanding of how MSL3 facilitates spreading to nearby active genes
79 remains incomplete and controversial. Early evidence pointed to the importance of histone
80 H3 lysine-36 trimethylation (H3K36me3) and its cognate lysine methyltransferase, Set2, in
81 propagating the MSL complex across the male X. First, *Set2* null male larvae exhibit a 2-10
82 fold reduction in MSL complex recruitment to a subset of X-genes [37]. Second, recombinant
83 MSL3 displays an affinity for H3K36me3 modified nucleosomes [37]. Despite these
84 findings, MSL recruitment defects observed in *Set2* mutants were inconsistent regarding
85 H4K16ac and/or mRNA levels over the genes examined [37]. Furthermore, a plasmid model
86 of DC also called into question the importance of H3K36me3 [38].

87 More recently, RNA-seq analysis of *Set2* mutant male larvae substantiated a small,
88 but significant decrease in X-gene expression, but the same study also found that *H3.2^{K36R}*
89 and *H3.3B;H3.3A* null mutants failed to display this effect [39]. Given that many histone
90 methyltransferases are known to target non-histone substrates [40-42], including the
91 mammalian ortholog of Set2 (SETD2) [43-45], it is plausible that the effect of *Set2* loss on

92 male X-expression is mediated by a target other than H3K36. However, other plausible
93 interpretations of these data remain.

94 The absence of females in previous studies also makes it difficult to discern whether
95 global X chromosome effects in male cells are due to “maleness” or “X-ness” in the sense
96 that the X itself has unique features not specific to sex that could impact gene regulation
97 [46-50]. Furthermore, the issue of functional redundancy between H3.2 and H3.3 K36
98 residues [51] was not considered [39]. Finally, with respect to DC, the potential for
99 heterogeneous regulation of X-genes has been underexplored. In particular, work
100 investigating “non-canonical” DC mechanisms provide important hints that mechanisms
101 for balancing sex chromosome gene dosage may not be entirely mediated by the MSL
102 complex [8, 18].

103 In this study, we utilize histone genetics and transcriptome profiling to clarify the
104 relationship between Set2, H3K36me3, MSL3 recruitment, and X chromosome gene
105 regulation. We confirm previous reports that Set2 impairs gene expression on the X
106 chromosome. However, our inclusion of females combined with nuanced bioinformatic
107 analyses reveal that the effects of Set2, H3.2K36, and H3.3K36 on X chromosome gene
108 expression are surprisingly heterogeneous. Importantly, our analysis of an
109 *H3.3^{K36R}/H3.2^{K36R}* combined mutant addresses the possibility of functional redundancy
110 between histone variants, and we find no evidence of involvement of H3K36 on promoting
111 expression of dosage-compensated genes. Interestingly, we frequently observe opposite
112 effects on gene expression between *Set2* and *H3^{K36R}* mutants at multiple developmental
113 stages suggestive of a regulatory switch between methyl states of H3K36. Lastly, we find
114 that X-genes with decreased expression in *Set2* and *H3^{K36R}* mutants in larval brain are
115 enriched in components of the BEAF-32 insulator complex compared to unaffected genes.

116 Based on these findings, we conclude that neither *Set2* nor *H3K36* are required for MSL3
117 recruitment, as their effects are gene-specific, context-dependent, and do not reliably
118 correlate with the presence MSL3 binding or H4K16ac. Rather, we argue that the evidence
119 is more compatible with *Set2* mediated H3K36 trimethylation impacting other processes
120 utilized in DC, but not specific to DC (such as elongation control or 3D genome
121 organization).

122

123

124 **Results**

125

126 ***H3.2^{K36R}* and *H3.3^{K36R}* mutations do not specifically impair male viability**

127 Male-specific lethality is a defining feature of mutations that affect DC in *Drosophila*
128 (reviewed in [21, 52]). Remarkably, this specificity extends all the way down to the histone
129 residues themselves, as an *H4^{K16R}* mutation causes developmental delay and death in male
130 progeny whereas their female siblings are completely viable [19]. This male-specific
131 lethality can be bypassed by expression of an acetylation mimicking *H4^{K16Q}* mutation [20].
132 Together, these results demonstrate that H4K16ac is the critical PTM of the DC machinery
133 in *Drosophila*. Moreover, they show that H4K16 is not required for basal genome function,
134 as female gene expression and viability were unaffected.

135 If H3K36me3 plays an important role in the localization or spreading of the MSL
136 complex, one might expect to observe decreased male viability in mutants that inhibit
137 H3K36 methylation. To test this idea, we assayed the fraction of adult males in H3.2 and
138 H3.3 K36R mutants, along with H4 K16R and HWT (histone wild type) controls. For
139 complete genotypes and genetic schemes for generating these animals see Figures 1A, S1

140 and Table S1. Note that *Set2* null and 12xH3.2^{K36R} animals fail to eclose as adults, but
141 wandering L3 males from these lines are readily obtained [37, 53]. To ascertain whether
142 H3.2 K36 and H4 K16 residues interact genetically, we carried out complementation
143 analysis between multi-gene families [51]. That is, we combined two 12x histone constructs
144 *in trans*, and assayed pupation and eclosion frequencies of the resulting progeny. A
145 significant change in viability by comparison to control crosses would suggest that the two
146 residues cooperate in common pathways. Previously, we found that H3.2 K36R interacted
147 strongly with K27R but was fully complemented by a K9R mutation [51].

148 As shown in Fig. 1B, addition of an *HWT* transgene fully rescued the larval and
149 pupal viability defects seen in *K36R*-only animals [51]. However, there was no significant
150 change in the number of males that eclose from a *K36R/HWT* cross compared to
151 *HWT/HWT* controls (Fig. 1C). If anything, there was a modest *increase* in *K36R/HWT*
152 adult males. Consistent with its known role in DC, we did observe a slight but insignificant
153 decrease in the fraction of males emerging from a *K16R/HWT* cross compared to the control
154 (Fig. 1C). However, the opposing sex skew of the *K36R/HWT* and *K16R/HWT* adults
155 resulted in a statistically significant difference (Fig. 1C). We also observed that modifying
156 the *K36R/HWT* genotype to *K36R/K16R* resulted in a significant decrease in males, but
157 the converse was not true (Fig. 1C). Modification of *K16R/HWT* to *K16R/K36R* resulted in
158 no change (Fig. 1C). Together, these two observations imply that the male-diminishing
159 effect of *K16R* predominates over the male-promoting effect of *K36R*. Thus H3.2K36R
160 histones appear to be slightly more toxic to females, whereas H4K16R histones specifically
161 affect males.

162 Importantly, we note the significant absence of adult males in *K16R/K16R* crosses,
163 despite the presence of wildtype copies of *His4r* in this background (Fig. 1C). Although

164 *His4r* is a replication-independent histone gene, it expresses an identical H4 protein.
165 Previously, we found that animals bearing a single 12x *K16R* transgene (crossed in
166 maternally) in a *His4r* positive background resulted in 8.5% eclosed males [19]. Taken
167 together, these findings support the notion that the proportion of zygotically expressed H4,
168 compared to the amount of wild-type maternal histones and *His4r*, is a critical determinant
169 of male viability.

170 In contrast with the results for H3.2, we found that *H3.3^{K36R}* mutants complete
171 development and eclose at a frequency of ~80%, which is nearly identical to that of *H3.3^{Ctrl}*
172 animals [51] (for full genotypes see Fig S2). We therefore assessed the ratio of males and
173 females in adults of these genotypes. We found that *H3.3^{K36R}* males comprise ~50%, of
174 eclosed adults, which is slightly but not significantly greater than that of the *H3.3^{Ctrl}* (Fig.
175 S3). These data suggest that an *H3.3^{K36R}* mutation does not substantially weaken dosage
176 compensation.

177
178 ***H3.3^{K36R}* interacts genetically with *H4^{K16R}***
179 Synthetic lethal (or synthetic sick) interactions are those wherein the combination of two
180 different mutations produces death or other strong phenotypes, whereas single mutations
181 do not. Synthetic interactions can thus implicate two genes as participating in a common
182 pathway [54, 55]. Given the importance of H4K16ac to *Drosophila* DC, we wondered
183 whether genetic evidence for involvement of H3.3K36 in DC might emerge in the sensitized
184 background of an *H4^{K16R}* mutation.

185 We hypothesized that if H3.3K36 were involved in DC, the male lethal phenotype of
186 the *H4^{K16R}* mutant would be enhanced. We therefore assayed overall viability and
187 male:female ratios in genotypes combining *H3.3^{Ctrl}* and *H3.3^{K36R}* mutations with *H4^{K16R}* (Fig

188 1A) (For full genotypes, see Fig. 4). In these experiments, *His4r* was wild type, as deletion
189 of this locus rendered the *H4^{K16R}* male lethal phenotype too severe to detect synthetic effects
190 (32% adult males vs 0%; see [19]). As expected, overall viability levels for Oregon R (*OreR*),
191 *H4^{HWT}*, and *H3.3^{Ctrl}H4^{HWT}* control genotypes were similar for both pupation and eclosion
192 (Fig. 1D). The addition of *H3.3B^{K36R}* to generate *H3.3^{K36R}H4^{HWT}* animals had no significant
193 impact on viability, though recent work shows that this mutation does reduce adult lifespan
194 (Fig. 1D, [56]). In contrast, *H3.3^{Ctrl}H4^{K16R}* mutants exhibited a significant reduction in
195 viability (~45% eclosion). This value is comparable to the eclosion frequency reported for
196 *H4^{K16R}* animals bearing wild type *H3.3* genes (50%) (Fig. 1D, [19]). Interestingly, when
197 *H3.3^{K36R}* and *H4^{K16R}* mutations are combined, adult survival is severely impaired (~20%; see
198 Fig. 1D), strongly suggesting that H3.3K36 and H4K16 regulate common pathways.
199 However, the degree of synthetic lethality also suggests that both males and females are
200 affected.

201 Given that H4K16ac is also deposited in the context of autosomal promoters, we
202 examined whether there was a more severe viability defect in males, suggestive of an
203 impairment to DC. We calculated the proportions of males and females from the eclosed
204 viable adults. As expected, *OreR* and *H4^{HWT}* produced roughly equal numbers of males and
205 females, but the *H3.3^{Ctrl}H4^{HWT}* control skewed significantly female (Fig. 1E). We note that
206 this imbalance was unexpectedly ‘rescued’ by mutation of *H3.3B^{K36R}* (*H3.3^{K36R}H4^{HWT}*; Fig.
207 1E), suggesting that loss of H3.3K36 can promote male survival in the context of H3.3
208 insufficiency. Strikingly, the *H3.3^{Ctrl}H4^{K16R}* genotype exhibited dramatic impairment of
209 male survival, despite the presence of a wild-type *His4r* gene. Compared to previous
210 reports, ablation of *H3.3A* reduced male survival 10-fold in the context of an *H4^{K16R}*
211 mutation (3.4%, Fig. 1E compared to 32%, [19]). Interestingly, combining *H3.3^{K36R}* and

212 $H4^{K16R}$ mutations ($H3.3^{K36R}H4^{K16R}$) completely eliminated eclosion of viable males. This
213 finding is consistent with the possibility that H3.3K36 performs a role in DC, however,
214 given that females were also affected to a lesser extent, the possibility that combining these
215 mutations confers a global reduction in viability that disproportionately affects weakened
216 males cannot be excluded.

217

218 **Transcriptomic analysis of *Set2* and $H3^{K36}$ mutants in the larval brain**

219 Although the genetic interaction between $H3.3^{K36R}$ and $H4^{K16R}$ was intriguing, we wanted to
220 assay the effects of K36 residue and writer mutations on male and female transcriptomes.
221 A previous study had analyzed brains of male $Set2^l$ (a null allele), $H3.3^{WT}H3.2^{K36R}$ and
222 $H3.3\Delta$ ($H3.3B^{null};H3.3A^{null}$) wandering 3rd instar (WL3) larvae [39]. These investigators
223 identified a role for Set2 in supporting expression of X chromosome transcripts in males,
224 however the exclusion of females from that study makes it unclear if this effect is truly
225 male-specific or simply X-specific. Moreover, the complete absence of H3.3 protein removes
226 an important nucleosomal subunit from many different subcompartments of the genome,
227 presumably replacing it with wildtype H3.2.

228 To extend the analysis to females and to better parse the relative involvement of
229 *Set2*, H3.2K36, and H3.3K36 in the regulation of gene expression, we performed poly-A
230 selected RNA-seq followed by DESeq2 differential expression analyses in WL3 brains.
231 Altogether, there were six replicates (3 male and 3 female) of three different mutant
232 genotypes plus three corresponding controls: $Set2^l$ and *yw*; $H3.3^{WT}H3.2^{K36R}$ and
233 $H3.3^{WT}H3.2^{HWT}$; $H3.3^{K36R}H3.2^{HWT}$ and $H3.3^{Ctrl}H3.2^{HWT}$ (see Fig. S5 and Table 1 for detailed
234 descriptions). Note that we analyzed the $H3.3B^{K36R}$ mutation on the $H3.2^{HWT}$ histone
235 replacement background to enable direct comparison with the $H3.3^{WT}H3.2^{K36R}$ animals. We

236 also sequenced samples to high read depth (62-95 million paired-end reads per replicate)
237 and avoided cutoffs based on a log₂ fold-change (LFC) thresholds in downstream analyses
238 because previous work has shown that mutation and knockdown of MSL complex members
239 yield subtle LFC values X chromosome-wide [31, 57].

240 Principal component analysis (PCA) revealed tight groupings of replicates by
241 genotype, as well as by sex (Fig. S6A). For our initial DESeq2 runs, we combined replicates
242 for both sexes into a single genotype class to simplify general trends in expression patterns
243 between the mutants. MA plots highlighting all differentially expressed genes (DEGs)
244 (adjusted *P* value < 0.05) revealed a notably greater number of DEGs in the *Set2^l* mutant
245 (7,042) than either the *H3.3^{WT}H3.2^{K36R}* (4,519) or the *H3.3^{K36R}H3.2^{HWT}* (1,835) mutant alone,
246 or their sum (6,344) (Fig. 2A). When adjusting this sum to account for genes that are DEGs
247 in both H3K36 mutant genotypes (5,508 for one or both H3K36R mutants; Fig S6B), these
248 data not only suggest significant functional compensation between H3.2K36 and H3.3K36,
249 but also the possibility of Set2 functions that are not related to H3K36. This pattern was
250 maintained when an LFC cutoff of > |1| was employed (Fig. S6C). We also note that,
251 within the subset of DEGs identified in all three mutant genotypes (618 genes), the largest
252 group of genes was upregulated in all three mutants (43%, see Fig. S6B). Additionally, a
253 substantial fraction (25%) was upregulated in both H3K36R mutant genotypes, but
254 downregulated in the *Set2^l* mutant, suggesting a regulatory relationship between H3K36
255 trimethylation and other modification states (Fig. S6B). Importantly, these data hint at
256 other possible regulatory scenarios besides H3K36-independent functions of Set2 or
257 redundancy between H3.2 and H3.3 residues.

258

259 **Individual Set2 or H3K36 mutations exert weak and inconsistent effects on global
260 X chromosome gene expression**

261 To understand the extent to which H3.3K36, H3.2K36, and Set2 might play a role in DC,
262 we performed additional DESeq2 comparisons, this time separated by sex (Fig.S7). Overall
263 patterns of gene expression were similar to the combined analysis when separated in this
264 manner (Fig.S7). To gain insight into whether, expression of X chromosome genes is
265 inhibited in *H3.3^{K36R}H3.2^{HWT}* mutants, we plotted the LFC of each mutant genotype relative
266 to its control, binned by chromosome arm for both males and females (Fig. 2B,C). In line
267 with previous work [39], we observed a significant decrease in chrX gene expression in male
268 *Set2^l* mutants. Importantly, we did not see this effect in females indicating that this X
269 chromosome-wide decrease is male-specific (Fig. 2C). We also observed a very slight, but
270 statistically significant decrease (adjusted $P < 0.01$) in the *H3.3^{K36R}H3.2^{HWT}* mutant males
271 (Fig. 2B). No change was observed in *H3.3^{K36R}H3.2^{HWT}* females, or in either sex in the
272 *H3.3^{WT}H3.2^{K36R}* genotype (Fig. 2B,C). These results suggest that there must be either
273 functional compensation between H3.3K36 and H3.2K36 with respect to male X
274 chromosome gene expression, or that Set2 regulates male X gene expression via some other
275 target. Remarkably, we also observed strong sexual dimorphism in the effect of all three
276 mutant genotypes with respect to the 4th chromosome, implying that sex differences in
277 chromosome-wide gene expression may not always be due to dosage compensation (Fig.
278 2B,C).

279 One feature of reduced expression of MSL complex members is a change in the
280 severity of male X gene expression impairment that varies by distance from high-affinity
281 MSL binding sites (HASs; see [31, 57]). Impairment of MSL2 binding to HAS loci results in
282 the greatest degree of gene expression loss overlapping the site itself, whereas impairment

283 of MSL3 exhibits the opposite pattern with the greatest decrease farthest from HAS sites
284 [31, 57]. These and other findings suggest that MSL2 is required for initiation of MSL
285 mediated DC and that MSL3 is involved in spreading of the complex to surrounding genes
286 (reviewed in [14, 58]). We were curious if the small, but significant decrease in X-gene
287 expression in $H3.3^{K36R}H3.2^{HWT}$ males would exhibit an HAS distance trend, consistent with
288 a role in DC. We also wanted to examine whether the previously observed relationship
289 between HAS site distance and gene expression in *Set2^l* mutants [39] was male-specific.

290 To probe these questions, we performed HAS distance analysis in both male and
291 female *Set2^l*, $H3.3^{WT}H3.2^{K36R}$, and $H3.3^{K36R}H3.2^{HWT}$ mutants. As shown previously, we
292 observed the greatest decrease in chrX gene expression nearest to HASs in *Set2^l* males,
293 suggestive of an initiation defect rather than a spreading defect (Fig. 2D [31, 39, 57]). We
294 also detected a similar, but smaller, effect in female brains (Fig. 2E). Analysis of
295 $H3.3^{K36R}H3.2^{HWT}$ mutants demonstrates gene expression trend related to HAS distance,
296 suggesting that the small difference in male X expression may not be due to DC. Inversely,
297 we observed an overall trend in the $H3.3^{WT}H3.2^{K36R}$ males and females resembling that of
298 *Set2^l* mutants, though weaker and less consistently. On the whole, these observations call
299 into question whether Set2 is likely to be involved in MSL complex spreading, as the
300 observed effects are neither male-specific, nor do they resemble a situation of impaired
301 MSL3 function. Furthermore, we found no evidence that either H3.2 or H3.3 K36R
302 mutation impacts DC at this developmental stage.

303

304 **H3.3K36 exhibits differential effects on X-gene expression during development**
305 Genetic redundancy between H3.2K36 and H3.3K36 complicates a determination of the
306 requirement for H3K36me3 in MSL complex spreading. However, one would expect that

307 compensation between H3 variants might be partially bypassed in tissues or developmental
308 stages where one variant predominates. In the adult brain, cells are largely senescent and
309 H3.3 incorporation increases with age [59, 60]. We therefore, took advantage of *H3.3^{K36R}*
310 mutant transcriptomic data obtained in adult male and female heads of both “young”
311 (newly eclosed) and “old” (~23 days post-eclosion) flies [56]. Indeed, transcriptomic
312 dysregulation on the whole increases in *H3.3^{K36R}* mutants with age in brain/head tissue
313 (Fig. 2A, [56]). Of note, *H3.3^{K36R}* mutant and *H3.3^{Ctrl}* animals were on a genetic background
314 with a wild-type RD histone locus in these analyses from adult heads [56].

315 Chromosome arm plots of LFC values by age and sex show a larger decrease in
316 median LFC for chrX genes relative to the large autosomes for both young and old flies of
317 both sexes (Fig. 3A,B). The magnitude of decrease increases with age, concurrent with
318 increased H3.3 incorporation (Fig. 3A,B). The presence of this decrease in both sexes
319 suggests this effect is due to “X-ness” rather than to DC. If this were true, one prediction
320 would be that despite decreased global X expression, there would be no relationship
321 between LFC and HAS distance. In fact, we observe no relationship in young males and old
322 females, and a significant *upregulation* of chromosome X genes by HAS distance in young
323 females (Fig. 3C, D). In old males, the overall trend is significant, but does not exhibit a
324 consistent change at each increment as would be expected if *H3.3K36* mutation were
325 impeding MSL complex spreading (Fig. 3D)

326 Finally, if H3.3K36me3 promotes DC in aged male flies, we would expect to observe
327 the greatest decreases in X-gene expression on genes with the highest levels of H4K16ac.
328 To assess the relationship between gene expression change and H4K16ac, we binned chrX
329 genes by mean H4K16ac signal in adult heads and plotted LFC in these bins (Fig. 3E).
330 Unexpectedly, in young male *H3.3^{K36R}* fly heads, we observed a compelling, male-specific

331 trend whereby gene expression *increases* with increasing H4K16ac (Fig. 3E). This is
332 precisely the opposite of what one would expect if H3.3K36me3 enables MSL3 spreading.
333 Instead, this pattern is more consistent with H3.3K36 *inhibiting* DC in some way. Also
334 unexpectedly, this relationship changes in the ageing male flies where the genes in the top
335 six deciles of H4K16ac exhibit decreased expression (Fig. 3E). This effect is mirrored (but to
336 a lesser extent) in females (Fig. 3E). These data argue against a simple role for
337 H3.3K36me3 in mediating MSL complex spreading, and instead hint that the effect of
338 H3.3K36 on X-gene expression may be mediated by other processes. Furthermore, these
339 data imply that effects of H3.3K36 on chrX gene expression are influenced by
340 developmental stage and age.

341
342 **The effect of Set2 and H3K36 mutations on X genes depends on chromatin context**
343 The effects of *Set2^l*, *H3.3^{WT}H3.2^{K36R}*, and *H3.3^{K36R}H3.2^{HWT}* mutations on global X
344 chromosome expression neither track consistently by sex, nor do they exhibit predicted
345 trends in gene expression by proximity to HASs. These findings suggest that such effects
346 are unlikely to be caused by a defect in MSL spreading. Furthermore, the largest effect in
347 *Set2^l* mutant males is considerably weaker than that observed following depletion of MSL
348 complex proteins, and stands in marked contrast to effects in *H4^{K16R}* mutants [19, 31, 57].
349 Given that all chromosomes harbor genes within different chromatin environments, subject
350 to different modes of regulation and activity [61, 62], we wondered whether our
351 observations could be explained by heterogeneous responses to Set2/K36 mutation within
352 different chromatin compartments.

353 To investigate this hypothesis, we utilized the genome-wide chromatin
354 characterization model defined by Kharchenko and colleagues [62]. This study applied a

355 machine learning approach to ChIP-seq data to define 9 basic chromatin states in two cell
356 culture models. We used their BG3 model (derived from male WL3 larval brain) for this
357 analysis. The 9 chromatin states include 5 “active” states (1-5) and 4 “repressive” states (6-
358 9). Though most genes span multiple states, we were able to identify a “predominant”
359 chromatin state for most genes, defined as the state covering > 50% of gene body length
360 (Fig. 3A). When genes were classified in this way, the composition of the male X was clearly
361 different from the autosomes, with three states comprising the bulk of genes (Fig. 3A).
362 State 5 genes, marked by H4K16 acetylated chromatin, encompass nearly half of the genes
363 on the male X. State 1, marked by H3K4me3 and H3K9ac and common at active promoters
364 accounts for about ~25%. Lastly, repressive State 8, marked by moderate levels of H3K9 di-
365 and trimethylation, covers ~12% of genes.

366 To examine whether Set2, H3.2K36 and H3.3K36 regulate chromosome X genes
367 heterogeneously within different chromatin states, we next plotted WL3 brain LFC values
368 of chrX genes for each mutant and sex binned by predominant state (Fig. 3B). Of note,
369 because BG3 cells are male, the chromatin features of these “State 5” genes in females are
370 unknown, but unlikely to be characterized by genic H4K16ac since this is a hallmark of
371 male DC. Remarkably, we observe different patterns of effects in the three mutant
372 genotypes depending on chromatin state (Fig. 3B). For State 5 genes, we observe a
373 significant median decrease specifically in *Set2^l* males, and no change in *Set2^l* females or
374 the *H3K36R* mutants. However, we note that a substantial fraction (>25%) of State 5 genes
375 are actually upregulated in *Set2^l* mutant males. In contrast, State 1 genes exhibit
376 significantly reduced expression in both sexes for all three mutant genotypes. This
377 difference reveals that State 5 and State1 chrX genes are differentially sensitive to H3K36
378 mutation. Even so, the median decrease in expression of State 1 genes in *Set2^l* males is

379 substantially greater than for the other genotypes (~6 fold > than *Set2^l* females; ~2.5
380 >*H3.3^{WT}H3.2^{K36R}* males). The disproportionate effect in both active states in *Set2^l* males
381 demonstrates that Set2 enhances expression of active genes on the male X in a distinctive
382 manner. Whether this outsized effect is due to an alternative function of Set2 or
383 redundancy between H3.2K36 and H3.3K36 at these genes remains unclear.

384 In contrast, expression of genes in repressive State 8 are substantially *increased* in
385 *Set2^l* and *H3.3^{WT}H3.2^{K36R}* mutants of both sexes, and slightly in *H3.3^{K36R}H3.2^{HWT}* males.
386 This adds to mounting evidence implicating H3K36 in repressing inactive or lowly
387 expressed genes [56, 63], and implies that that Set2 may support gene repression in some
388 contexts as well. Taken together, these data hint that the effects of Set2, H3.2K36, and
389 H3.3K36 on chrX gene expression are context-dependent.

390

391 **Set2 and H3K36 variants exhibit variable patterns of X chromosome gene
392 regulation**

393 Thus far, our analyses hint that chrX genes respond in a pleotropic manner to mutation of
394 *Set2*, *H3.2K36*, and *H3.3K36*, suggesting that regulation by these players is context-
395 dependent, and potentially multi-faceted. We wanted to better understand the interplay of
396 these mutations on specific genes and genomic contexts, and ascertain whether any of these
397 contexts were associated with sexually dimorphic effects. To address these questions, we
398 first identified groups of genes likely to be similarly regulated. We reasoned that genes with
399 common regulatory mechanisms would exhibit similar patterns of expression changes with
400 respect to genotype and sex. To assess global patterns of regulation across differentially
401 expressed genes on the X, we constructed a *k*-means clustered heatmap of the combined
402 DEGs for all mutants. We used the *z*-score difference of DESeq2 normalized counts

403 (individual replicates – mean of controls of combined sexes) for each gene to enable
404 comparison of genes with vastly different expression levels (Fig. 5A). From this heatmap,
405 we were able to extract gene names for further analysis of cluster features. For each
406 cluster, we calculated the base mean gene expression (Fig. 5B), LFC between mutants and
407 same-sex controls (Fig. 5C), relative levels of H3K36 methylation states (Fig. 5D) and DC
408 proteins (Fig. 5E), and relative enrichment of proteins and marks associated with the
409 Kharchenko chromatin states (Fig. S8). For analyses of cluster features, chrX genes
410 unchanged in any of the Set2/H3K36 mutants (nonDEGs) were included for comparison. Of
411 interest, this *k*-means clustering approach reveals that many X-genes exhibit mild sexual
412 dimorphism in expression in wild type males and females (Fig. 5A), as male and female
413 replicates are consistently on opposite sides of the genotype mean (L3-c1, c2, c3, c4, c9) in
414 the *yw* control (Fig. 5A).

415 With respect to our genotypes of interest, we identified nine distinct patterns of
416 regulation amongst all genotypes and sexes, three of which (clusters L3-c1, c2, c3; 604/2017
417 of total chrX genes) align with what would be expected if H3K36me3 enabled spreading of
418 the MSL complex (Fig. 5A,C). For these clusters, we observed male-specific expression
419 decreases in the *Set2^l* mutant, and to a lesser extent in either the *H3.3^{WT}H3.2^{K36R}* or the
420 *H3.3^{K36R}H3.2^{WT}* mutants (Fig. 5C). These clusters were also amongst the highest in
421 relative enrichment of H3K36me3 and MSL complex proteins (Fig. 5D, E). Notably, we did
422 not observe any gene clusters with expression changes in the *Set2^l* mutants, but where
423 *H3.3^{WT}H3.2^{K36R}* and *H3.3^{K36R}H3.2^{WT}* resembled controls, suggesting that the role of *Set2^l* in
424 promoting expression of chrX genes in males is likely to occur by way of H3K36 in this
425 tissue/stage, rather than by some other target or function of Set2. L3-c1,c2, c3 are
426 compatible with the idea of redundancy between variants, as the magnitude of change in

427 the *Set2^l* mutant is greater than either *H3^{K36R}* mutant even while changing in the same
428 direction (Fig. 5C). These observations are consistent with the possibility that Set2 via
429 H3K36me3 may promote gene expression of some dosage-compensated genes.

430 Two other clusters also exhibited sexually dimorphic expression changes, but
431 different from what would be expected if H3K36me3 were facilitating canonical DC. Cluster
432 L3-c4 shows decreased expression in *Set2^l* males, but *increased* expression in
433 *H3.3^{WT}H3.2^{K36R}* females, whereas L3-c6 shows *increased* expression in *Set2^l* females and
434 decreased expression in *H3.3^{WT}H3.2^{K36R}* males (Fig. 5A,C). L3-c6 is among the most
435 enriched in H3K36me3 in gene bodies and L3-c4 is relatively less so. Increased expression
436 in female mutants resembles what would be predicted in response to a defect in “non-
437 canonical dosage compensation” whereby lowly expressed genes in heterochromatin
438 depleted of MSL complex in males, are inhibited in females by way of homolog pairing [8].
439 However, neither cluster is depleted in MSL complex proteins (Fig. 5E) or enriched in
440 repressive histone marks or chromatin proteins (Fig. S8). Furthermore, L3-c4 contains
441 genes with the highest base mean (Fig. 5B). These observations suggest that L3-c4 and L3-
442 c6 are unlikely to employ non-canonical DC as defined previously.

443 Clusters L3-c7, c8, c9 are primarily defined by upregulation in one or more mutant
444 genotype. L3-c8 and L3-c9 are relatively enriched in H3K36me1 and depleted in H3K36me3
445 (Fig.5D). These genes were lowly expressed on the whole and enriched in heterochromatic
446 marks (Fig. 5B, S8). Even so, gene expression was significantly increased in L3-c9 in the
447 *Set2^l* mutant (Fig.5C). This is consistent with the possibility of indirect effects, or these
448 genes may correspond to genes where Set2 depletion results in increased H3K36me1 on the
449 chromosome arms [47]. Lastly, clusters L3-c5 and L3-c7 are driven primarily by H3.3
450 mutation. These genes also have intermediate levels of DC proteins. Overall, these data

451 imply a large degree of heterogeneity in how H3.2K36, H3.3K36, and Set2 impact X
452 chromosome gene expression, which is inconsistent with a role in chromosome-wide dosage
453 compensation.

454

455 **Insulator proteins associate with X chromosome DEGs in Set2/H3K36 mutants**

456 We next wanted to gain insight into what might be driving the diverse patterns of gene
457 expression changes observed in the *Set2* and *H3^{K36R}* mutants. To this end, we performed
458 motif enrichment analysis using the SEA (Simple Enrichment Analysis) tool [64] on the
459 WL3 brain mRNA-seq heatmap clusters (Fig 5A; Fig 6A) [65]. Promoter and gene body
460 regions for genes in each cluster were compared to these regions in nonDEGs. We focused
461 on the most enriched motifs, those exhibiting a *q*-value < 0.05 and enrichment over control
462 sequences > 2 (Fig. 6A).

463 Interestingly, BEAF-32 and Dref motifs were enriched at promoters across multiple
464 clusters, and exhibiting diverse expression patterns between mutants (Fig. 6A). BEAF-32 is
465 a protein linked to 3D genome organization, insulator function, and gene regulation [66-69].
466 Dref is a transcription factor involved in insulator function, chromatin organization, gene
467 expression, and telomere maintenance [66, 70, 71]. Interestingly, BEAF-32 and Dref bind
468 similar, often overlapping, DNA motifs [72]. Both functional redundancy [66, 73] and
469 inverse binding profiles have been reported for these factors in different contexts [72]. The
470 most significantly enriched clusters for Dref motifs (L3-c1, L3-c2, L3-c3, and L1-c4), also
471 have the highest median gene expression and exhibit a male-specific decrease in gene
472 expression in the *Set2^l* mutants (Fig. 5A,B). Three of these clusters (c1, c2, and c3) are also
473 the most significantly enriched in BEAF-32 motifs (Fig. 6A). L3-c8 was also enriched in
474 BEAF-32 motifs, though these genes were upregulated in *Set2/H3^{K36R}* mutants (Fig. 5A,B).

475 Next, we assessed whether motif enrichment corresponded to increased insulator
476 protein binding at the promoters of these genes. We constructed heatmaps of relative
477 insulator protein binding for each L3 heatmap cluster for factors with available
478 modENCODE ChIP data (as in Fig.5D,E). We included proteins known to work in
479 conjunction with BEAF-32 (CP190 and Chromator) along with others that operate in
480 different insulator complexes (SuHw and GAF) [69]. We observed substantial relative
481 enrichment of BEAF-32, CP190, and Chromator in L3-c1, c2, c3, and c4 (Fig. 6B). Of note,
482 L3-c8 was relatively depleted in binding of these proteins, despite enrichment of BEAF-32
483 motifs (Fig. 6B). We observed peaks of BEAF-32 and CP190 at many promoters and some
484 3' ends of genes, but these peaks did not always overlap with each other (Fig.6C). For
485 comparison, we saw no enrichment of SuHw on any cluster or the NonDEGs (Fig. 6B).

486 We also constructed metaplots of BEAF-32 and CP190 to assess the distribution of
487 signal across genes with similar levels of binding (Fig. 6D). Consistent with previous
488 reports, BEAF-32 and CP190 peak near the TSS, with a much smaller enrichment after the
489 TES ([74]; Fig. 6D). This effect was strongest in L3-clusters 1-4, and weakest in the
490 nonDEGs (Fig. 6D). In contrast, a metaplot of SuHw showed relative depletion in L3-
491 clusters 1-4 (Fig. 6D).

492 The male X chromosome of BEAF-32 mutants exhibit unusual morphology in
493 polytene spreads, despite normal recruitment of MOF [75]. Tissues and cells with impaired
494 levels of BEAF-32 also have widespread transcriptomic changes [66, 76]. We wondered
495 whether cells with a reduction in BEAF-32 might exhibit a decrease in chrX gene
496 expression relative to autosomes, as was observed in *Set2^l* mutant males ([39]; Fig. 2B). To
497 address this question, we reanalyzed RNA-seq data from a previous study of BG3 cells
498 RNAi depleted for insulator complex transcripts [66]. We calculated LFC values for

499 knockdown (KD) conditions of BEAF-32, BEAF-32 + Dref, and CP190 + Chromator and
500 plotted these values by chromosome arm (Fig. 6E).

501 Like *Set2^l* mutant males, median gene expression for autosomal genes was elevated
502 for all three insulator KD conditions (Fig. 2B, 6E). Expression of chrX genes was also
503 elevated in the insulator KD conditions, but for the BEAF-32 and BEAF-32 + Dref
504 conditions, this increase in expression was significantly less than what was observed in
505 autosomes (Fig. 6E). In contrast, there was no significant difference in the CP190 +
506 Chromator condition between autosomes and chrX, despite ~90% and ~70% reductions in
507 CP190 and Chromator proteins, respectively (Fig. 6E, [66]. These data imply that BEAF-32
508 promotes gene repression to a lesser degree on the male X chromosome than on autosomes.

509 Given the heterogeneous, context-dependent effects on chrX gene expression when
510 components of the Set2/H3K36 axis are mutated, we wanted to determine if reduction of
511 insulator components demonstrated similarly heterogeneous changes. We hypothesized
512 that if Set2/K36 and BEAF-32 dependent mechanisms of gene regulation were operating on
513 the same genes in a collaborative manner, one would observe similar gene expression
514 trends in BEAF-32 knockdown cells when binned according to Set2/H3K36 expression
515 clusters. When this analysis was performed, we observed a remarkable concordance
516 between the gene expression trends in the *Set2^l* mutant males and the insulator protein
517 knockdowns for nearly all L3 clusters (Fig. 5C, 6F). The exceptions were L3-c5 and L3-c7
518 which were primarily driven by changes in the *H3.3K36R*/*H3.2^{HWT}* mutant. In summary, these
519 data demonstrate that BEAF-32 binds the promoters of Set2 responsive chrX genes in male
520 cells, and that mutation of both factors have similar effects on expression of dosage-
521 compensated genes. This is consistent with the possibility that Set2 and H3K36 may

522 enhance expression of many male X genes by impacting insulator function rather than by
523 way of MSL complex spreading.

524

525 **H3K36me3 does not play an essential role in MSL3 spreading**

526 Our experiments thus far suggest that H3K36me3 is unlikely to be uniquely important for
527 MSL complex spreading. For chrX genes, mutation of *Set2* in males causes small decreases
528 in downregulated genes, and upregulates many others. Moreover, many of the same
529 changes can be observed in females to a lesser extent. In some gene groups, the effects of
530 *Set2* and *H3^{K36}* mutation do not align. These effects are consistent with the possibility that
531 the Set2/H3K36 axis is affecting gene expression by one or more other means, including by
532 impacting insulator function. However, recent work suggests that MSL3 might also bind
533 H3K36me2, which could explain the weak and inconsistent effect on chrX gene expression
534 in the *Set2^l* mutants [47]. Furthermore, we have not yet fully investigated the prospect of
535 functional redundancy between H3.2K36 and H3.3K36.

536 To address these alternatives, we performed total RNA-seq and DESeq2 analysis at
537 the L1 stage in *Set2^l* and combined *H3.3^{K36R}H3.2^{H3K36R}* mutants where all zygotic H3K36
538 has been mutated, alongside control genotypes. The *H3.3^{K36R}H3.2^{K36R}* genotype addresses
539 both genetic redundancy between variants and the possibility that MSL3 might bind to
540 H3K36me2, simultaneously. We used a mixed sex population because sexing them at this
541 stage in the context of a transgenic system already using YFP selection was not yet
542 possible. Because we used mixed sex larvae, we also included the *H3.3^{Ctrl}H4^{K16R}* mutant
543 genotype to verify that we could detect a signature of male DC in a mixed sex population.
544 We examined this developmental stage because the *H3.3^{K36R}H3.2^{K36R}* mutants are L1 lethal
545 [51].

546 Genome-wide MA plots of *Set2^l* and *H3.3^{K36R}H3.2^{K36R}* mutants illustrate that large
547 numbers of genes are differentially expressed in both mutants (6,533 and 5,799
548 respectively), comparable to that observed in *Set2^l* mutant WL3 brains, indicating that
549 maternal contribution of wild type proteins is unlikely to be masking an effect on gene
550 regulation (Fig. 7A, 2A). In contrast, a modest number (645) of genes reached statistical
551 significance in the *H3.3^{Ctrl}H4^{K16R}* mutants. These overall trends were preserved when a
552 cutoff of LFC > |1| was employed for these DEGs (Fig. S9A). Despite the relatively small
553 number of DEGs in the *H3.3^{Ctrl}H4^{K16R}* animals, when we plotted LFC values by chromosome
554 arm, there was a highly significant ($p < 10^{-15}$) decrease in global chrX gene expression in
555 these animals, demonstrating the ability to detect a DC defect in a mixed population (Fig.
556 7B). In contrast, despite much greater changes to their respective transcriptomes, we
557 observed no change in the *Set2^l* mutants and a highly significant *increase* in the
558 *H3.3^{K36R}H3.2^{K36R}* genotype (Fig. 7B).

559 HAS distance analyses were concordant with these results. In the *H3.3^{Ctrl}H4^{K16R}*
560 mutants, we observed clear and statistically significant incremental change in the
561 magnitude of transcript reduction varying by distance from the HAS site (Fig. 7C).
562 Conversely, we found no such correlation in the *Set2^l* and *H3.3^{K36R}H3.2^{K36R}* mutants (Fig.
563 7C).

564 If H3K36 methylation were required for MSL complex spreading, one prediction
565 would be that the greatest loss of expression would be on genes with the most MSL
566 complex. To test this, we plotted median LFC for decile bins corresponding to mean gene
567 body MSL3 signal (Fig. 7D). In the *H3.3^{Ctrl}H4^{K16R}* mutant controls, we observed a nearly
568 perfect incremental relationship between bin medians whereby the greatest decrease in
569 gene expression occurs at the highest MSL3 levels (Fig. 7D). This trend was clearly visible

570 in mixed sex samples and with relatively less transcriptome dysregulation overall. In
571 contrast, the *Set2^l* mutants tended to *increase* at genes with the highest MSL3 (Fig. 7D). In
572 the *H3.3^{K36R}H3.2^{K36R}* genotype, there was little or no change in the top two deciles of MSL3
573 occupancy, with the most substantial median decrease in gene expression occurring in the
574 fourth highest decile (Fig. 7D). Importantly, neither the *Set2^l* or *H3.3^{K36R}H3.2^{K36R}* genotype
575 showed any clear relationship with MSL3 occupancy. Nor did those two genotypes resemble
576 one another in this aspect. Instead, they trended opposite to each other in all but one bin
577 (Fig. 7D). These opposite trends also held when LFC values were binned by base mean gene
578 expression (Fig. S9B).

579 Next, we wanted to look directly at the patterns of gene expression among *Set2^l*,
580 *H3.3^{K36R}H3.2^{K36R}*, and *H3.3^{Ctrl}H4K16R* mutants for genes on the X chromosome. We
581 constructed a *k*-means clustered heatmap of *z*-score differences for the combined set of chrX
582 DEGs, as in Fig. 5A (Fig. 7E). Strikingly, we observed that most genes exhibit an opposite
583 expression trend between the *Set2^l* and *H3.3^{K36R}H3.2^{K36R}* mutants, providing further
584 evidence of a regulatory “switch” between methylation states (Fig. 7E, F). We also observed
585 that the cluster with the strongest decrease in expression in the *H3.3^{Ctrl}H4K16R* mutants (L1-
586 c2), the highest relative H3K36me3 (Fig. 7G), and greatest relative occupancy of DC related
587 proteins (Fig. 7H), showed a trend toward *upregulation* in both the *Set2^l* and
588 *H3.3^{K36R}H3.2^{K36R}* mutants, which argues against a role for H3K36me3 in promoting
589 H4K16ac (Fig. 7E). Furthermore, the three clusters with the highest relative enrichment of
590 H3K36me3 (L1-c2, c3, c4), show *upregulation* in the *Set2^l* mutant suggesting that Set2 is
591 acting to dampen expression at these genes (Fig. 7F,G).

592

593 **Set2 and H3K36 exhibit context-specific expression discordance**

594 We also noted that *k*-means clustered heatmaps looked very different at L1 and L3 stages
595 (Fig. 5A, Fig. 7E). In the L3 heatmap, *Set2^l* and *H3^{K36R}* mutations resulted in only 3 of the 9
596 clusters (L3-c4, c6, and c7, comprising ~27% of L1 DEGs) exhibiting discordant expression
597 changes (Fig. 5A,C). In contrast, for nearly all gene clusters in the L1 heatmap, *Set2^l* and
598 the combined *H3.3^{K36R}H3.2^{K36R}* mutant resulted in opposite trends (~81% of L1 DEGs),
599 excepting L1-c1 and L1-c2 (Fig. 7E,F). In the case of the *Set2^l* mutant, we also see
600 discordance between developmental stage/tissue type within the very same genotype.
601 Analyses of the 3 most common male X Kharchenko states (Fig. S9C; States 1, 5, and 9 in
602 S2 cells) reveals contradictory trends in State 1 for this genotype (Fig. 4B, S9D). This
603 reveals an additional layer of context-dependence in X chromosome regulation related to
604 developmental stage or tissue type. Intriguingly, the relative levels of the three *Drosophila*
605 H3K36 methyltransferases can also differ between WL3 brain and whole L1 larvae,
606 consistent with the possibility that differential methylation profiles at particular loci could
607 mediate these changes (See Discussion) (Fig. S10). In summary, these data provide
608 compelling evidence for context- and stage-dependent regulation of the X chromosome by
609 Set2/H3K36. Moreover, the data do not support a requirement for a specific H3K36
610 methylation state in MSL complex spreading, even when all zygotic copies of H3 cannot be
611 methylated at lysine-36.

612

613

614 **Discussion**

615

616 **Trimethylation of H3K36 is not essential for spreading of the MSL complex**

617 This study provides strong evidence against the prevailing dogma that H3K36me3
618 mediates spreading of the MSL complex. Although many gene clusters enriched in MSL
619 complex members in males are downregulated in *Set2* mutant males, most of these genes
620 exhibit the same general trends in females (Fig. 5A,C, E). Furthermore, we have identified
621 genes marked by MSL and highly decorated with H3K36me3 that are unaffected in *Set2*
622 males, but trend upwards in *Set2* females (Fig. 5, L3-c6). *H3.3A^{null}H4^{K16R}*We also note that
623 HAS analyses of *Set2* mutants resemble the pattern observed in depletion of MSL2
624 (involved in initiation at HASs) rather than MSL3 (involved in MSL complex spreading
625 ([31, 57]; Fig. 2D,E). *H3.3^{Ctrl}H4^{K16R}* mutants, even at an early stage and in a mixed sex
626 population, exhibit a nearly ubiquitous downward trend in chrX gene expression (Fig. 7). In
627 contrast, mutations of Set2 and H3K36 elicit heterogeneous effects across the X
628 chromosome at multiple developmental stages (Figs 4, 5, and 7).

629 Yet, clearly for a large proportion of genes exhibiting enrichment of H4K16ac and
630 MSL complex, Set2 exerts an outsized effect in males (Fig. 5A,C,E). We propose a model
631 whereby Set2 (via H3K36) likely supports expression of genes by other mechanisms such as
632 nucleosome turnover [77], elongation control [78-80], recruitment of HDACs [81, 82], or as
633 suggested in this study, functional relationships with insulator proteins (Fig. 6; [67, 83]).
634 In males, one or more of these mechanisms may synergize with the MSL complex, which is
635 believed to utilize both elongation control and 3D genome organization in propagating its
636 function [12, 49, 57, 84].

637 If H3K36me3 is not essential for MSL complex spreading, what are some
638 alternatives? One possibility is methylation of histone H4 lysine 20 (H4K20). Like
639 H3K36me3 and H4K16ac, H4K20 monomethylation localizes preferentially to gene bodies
640 [85-87]. *In vitro* studies demonstrate that H4K20me1 and H4K20me2 peptides have an up-

641 to 50fold higher affinity for the MSL3 chromodomain compared to H3K36me3 [88-90]. A
642 Y31A mutation in the MSL3 chromodomain that weakens *in vitro* binding of H4K20
643 methylated peptides, also reduces survival of males when introduced *in vivo* [90]. The K9-
644 S10 portion of the H3 tail has also been connected to regulation of male X genes. H3K9me2
645 on X-specific 1.688^X satellite sequences has been shown to support proper expression of
646 surrounding genes [91], and ectopic expression of siRNA from these repeats can partially
647 rescue *roX1roX2* mutant males [92].

648 Importantly, these possibilities are not mutually exclusive. MSL complex might
649 make use of multiple chromatin features for targeting, including H3K36me3, H4K20me,
650 and H3K9me2. This could occur either redundantly between marks, or with specificity on a
651 gene-by-gene basis depending on which marks predominate. The second possibility might
652 be evidenced by preferential regulation of different subsets of male X genes in H3K36,
653 H4K20, H9K9 mutants. There is precedent for redundancies in the DC system regarding
654 both *roX1* and *roX2*, as well as replication-dependent *H4K16* and replication-independent
655 *His4r* [19, 20]. Further studies addressing the impact of these other histone tail residues on
656 DC, either alone or in concert, would be informative.

657

658 **Relationships between H3K36, insulator proteins, and dosage compensation**

659 Given that we found enrichment of BEAF-32 and CP190 in the promoters of Set2
660 responsive X-genes (Fig. 6A,B), and similar effects on many gene clusters when Set2 and
661 BEAF-32 are impaired (Figs. 5C & 6F), we believe that 3D genome structure and insulator
662 function are especially promising areas of potential synergy between H3K36 and DC. The
663 male and female X chromosomes have surprisingly similar large-scale organization [49, 93],
664 but with more mid- to long-range interactions on the male X [94]. Intriguingly, Clamp, a
665 protein essential for *Drosophila* DC [32, 33, 95, 96] promotes the interaction of HASs in 3D

666 space [97]. Furthermore, Clamp and MSL complex binding are enriched at BEAF-32/CP190
667 domain boundaries that are weakened in males [94]. Like H3K36me3, Clamp binds
668 genome-wide where it can impact gene expression independently of the MSL complex, as
669 well as synergize with the MSL complex during DC [96, 98-100]. Thus, Clamp sets a
670 precedent for the model that we espouse.

671 Interestingly, Clamp is known to interact with two separate insulator
672 complexes: the late boundary complex [101] and the *gypsy* insulator [102]. Furthermore,
673 depletion of Clamp results in reduction of CP190 at some sites [102]. Clamp has also been
674 shown to interact with two separate insulator complexes: the late boundary complex[101]
675 [102] [102]Clamp also [33, 99]interacts with several histone proteins, including H3.2 and
676 H3.3 [103], and can bind nucleosomal DNA to increase chromatin accessibility [98]. Thus, it
677 is tempting to speculate that H3K36 and Clamp may cooperate in some manner.

678 BEAF-32 peaks occur most often near the TSS, while H3K36me3 is enriched at the
679 3' ends of genes, thus any model of interplay between these factors must account for their
680 different spatial positions. One possibility is an interaction between BEAF-32 and
681 H3K36me3 chromatin. Indeed, one 4C study identifying the most prevalent chromatin
682 states for BEAF-32 interactions showed that BEAF-32 had the strongest interaction with
683 active chromatin harboring H3K36me3, rather than active chromatin depleted of
684 H3K36me3, consistent with the possibility of a functionally important interaction [61, 68].
685 One study reports that weakening of domain boundaries containing BEAF-32 parallels
686 binding of the MSL complex on the male X [94]. In conjunction with our data, this suggests
687 the intriguing possibility that H3K36me3 might assist in weakening these boundaries
688 somehow. Future 4C or Hi-C studies, as well as chromatin binding studies of BEAF-32 and
689 other insulator proteins in *Set2* and *H3K36R* mutants would be of great interest in evaluating
690 this hypothesis.

691

692 **Context-dependence of X-gene expression at different developmental stages**

693 One surprising conclusion of our study is the strong effect of developmental stage/tissue

694 type on X chromosome gene expression heterogeneity. We enumerate two distinct effects.

695 First, we find that the degree of agreement between Set2 and H3K36 mutants differs

696 widely between the L1 and WL3 brain datasets, with much greater discordance in the L1

697 samples (Fig. 5A,C & Fig. 7E, F). Secondly, we find that individual genotypes can trend

698 differently in the same chromatin states between these datasets. The best example of this is

699 in the *Set2¹* mutant genotype in State 1 (Fig.4B, Fig.S9D).

700 What could be causing these variations? One exciting possibility is that differential

701 expression of H3K36 methyltransferases (KMTs) at different stages or in different tissues

702 could be driving these differences. In our RNA-seq data, we see distinct relative levels of

703 H3K36 KMTs between L1, WL3 brain, and adult head (Fig. S10). At L1, NSD and Ash1 are

704 ~40% and ~15% more highly expressed than Set2 (Fig. S10). In contrast, NSD is ~15% more

705 highly expressed than Set2 in WL3 brain, while Ash1 expression falls below that of Set2.

706 In adult heads, NSD expression is less than 50% of that of Set2 and Ash1, which are

707 roughly equal (Fig. S10). Some of these differences may be specific to nervous system tissue,

708 as another study examined levels of these KMTs and found different trends in whole WL3

709 larvae and whole aged adults [39].

710 One model driven primarily by experiments in female Kc cells posits a direct

711 interaction between BEAF-32 and NSD which preconditions H3K36me2 for Dref/Set2

712 driven trimethylation [67, 83]. Bulk modifications by H4K16ac by Western blot elicited the

713 conclusion that decrease of H3K36me3 alone leads to decreased H4K16ac, while decrease of

714 both H3K36 di- and trimethylation led to *increased* H4K16ac [83]. Since *H3^{K36R}* mutation

715 eliminates all methylation states while *Set2* mutation eliminates only trimethylation, this

716 is consistent with the idea of a regulatory switch between methylation states, and could
717 account for some of the discordance we observed, while also explaining how these
718 differences could be exacerbated by varying levels of H3K36 KMTs. It is also intriguing to
719 speculate that given this connection with insulators, differential KMT levels might also
720 exert differential effects on insulator function.

721 Though interesting, this “preconditioning model” has recently been challenged by a
722 genome-level study in S2 cells of the three *Drosophila* H3K36 methyltransferases (KMTs),
723 their binding patterns, and the subsequent effects on H3K36 methylation and the
724 transcriptome when these writer enzymes are subjected to RNAi knockdown [47]. This
725 study suggests that Set2 does not require H3K36me2 to trimethylated H3K36, and that
726 most genes are primarily methylated by one particular KMT on a gene-by-gene basis [47].
727 Even so, reduction of one KMT can also affect activity of other KMTs in a “see-saw effect”
728 [47]. The authors also report that NSD can perform trimethylation on some genes. One
729 possible implication of this study is that differential levels of KMTs would be expected to
730 exert genome-wide, locus-specific, and context-dependent effects that could conceivably vary
731 by tissue and/or developmental stage. A comprehensive investigation of H3K36 readers and
732 writers in different cell types, tissues, and stages would shed additional light on the basis
733 for these context-dependent effects.

734 Although we believe we can make many strong conclusions, it is important to point
735 out potential limitations of this study. First, these results are limited to specific
736 developmental timepoints/tissues. While we would expect findings related directly to MSL
737 complex function to be broadly applicable, other sources of heterogeneity are likely to vary
738 in other tissues and stages, as we have found to be the case in this study. The use of mixed
739 sex larvae at L1, while suggestive, necessitates cautious interpretation. ChIP-Seq data
740 were obtained from cell culture models. Additionally, we have not directly measured MSL3

741 binding, but have inferred it by examining gene expression. In future studies, we would like
742 to generate antibodies to test this directly.

743

744 **Conclusions**

745 In summary, the work here does not support the widely held view [21, 34, 104-106] that
746 H3K36me3 is essential for *Drosophila* MSL complex spreading. Our transcriptomic study of
747 X-gene regulation in *Set2*, *H3.2^{K36}*, *H3.3^{K36}* and combined *H3^{K36}* mutants of both sexes is
748 inconsistent with this idea. Instead, the data point to mechanisms whereby Set2 and
749 H3K36 support X chromosome gene expression via processes common to both sexes, that
750 synergize with the MSL complex in males. These findings lead to a more accurate
751 understanding of the relationship between H3K36 writers and residues and its effects on
752 the activity of MSL complex. As these same regulatory paradigms and processes are
753 conserved in mammals, these findings will be important for our understanding of human
754 health and disease.

755

756

757 **Methods**

758 ***Drosophila* lines and husbandry**

759 To obtain experimental progeny, parental flies were housed in cages sealed with grape juice
760 agar plates smeared with supplemental yeast paste. Plates were changed daily. L1 larvae
761 were obtained directly from the grape juice plates. Older animals were picked at the L2
762 stage, 50 per vial, and raised on cornmeal-molasses food. All experimental animals were
763 raised at 25°C. Details concerning construction of BAC transgenes generated previously
764 containing the 12xH3.2 and 12xH4K16R histone gene arrays can be found in [19, 53, 107].

765 *HisΔ* indicates *Df(HisC^{ED1429})*; flies containing the *HisΔ*, *twGal4*,
766 and *HisΔ*, *UAS:2xYFP* chromosomes [108] were received from A. Herzig.
767 The *H3.3A^{2x1}* (*H3.3A^{null}*) [109], *Set2¹* allele and rescue transgene [84], *Df(2*
768 *L)Bsc110* deficiency, and the beta-tubulin GFP protein trap stock used for recombination
769 with the rescue transgene were obtained from Bloomington Stock Center (nos. 68240,
770 77917, 8835, and 50867). The *H3.3B^{K36R}* CRISPR allele was generated previously [51]. Gene
771 names, annotations, genome sequence, references, and other valuable information useful to
772 this study were acquired from FlyBase [110].

773 **Generation of mutant genotypes**

774 For detailed genetic schemes, see Figs. S1, S2, S4, & S5). *HisΔ* animals were obtained by
775 selection for yellow fluorescent protein (YFP). Other H3.3 genotypes were selected for
776 absence of a *CyO*, *twGFP* balancer chromosome. *Set2* mutants were detected by absence of
777 GFP from both a maternal FM7i, act>GFP balancer and a paternal chromosome carrying a
778 *Set2* rescue transgene linked to a transgene expressing GFP tagged B-tubulin.

779 **Pupal and adult viability and sex ratio assays**

780 For each genotype, fifty L2 larvae were picked from grape juice agar plates and transferred
781 to vials containing molasses-cornmeal food. Full plates were picked to prevent bias due to
782 different developmental timing between males and females. Pupae and eclosed adults were
783 counted until 13 and 18 days after egg laying, respectively. Pupal and adult eclosion
784 percentages were calculated per-vial by dividing the number of pupal cases or eclosed
785 adults per 50 input larvae and multiplying by 100. Each vial constituted one biological
786 replicate for statistical purposes. Between 400 and 500 total animals (8-10 replicate vials)
787 were analyzed per genotype. For male and female ratios, number of males and females were

788 determined from eclosed adults from the above viability assays. Statistical significance for
789 % eclosion was obtained with Brown-Forsythe and Welch ANOVA tests, followed by
790 Dunnett's T3 multiple comparisons test. Statistical significance for sex ratio was
791 obtained with Fisher's Exact Test, followed by the Benjamini-Hochberg False Discovery
792 Rate (FDR) correction for multiple comparisons ($Q=0.05$). Graphpad Prism was used for
793 calculations.

794

795 **RNA Seq library preparation and sequencing**

796 For the wandering L3 brain experiment, 25 brains were dissected per replicate and
797 homogenized in 1ml Trizol solution. RNA was obtained from the Trizol aqueous phase using
798 the Zymo RNA Clean and Concentrator-5 kit (Genesee Scientific #11-352) plus DNase I
799 treatment, according to manufacturer's instructions. PolyA-selected libraries were prepared
800 using the KAPA stranded mRNA kit (Roche # 07962207001) and sequenced using the
801 NOVASEq-S1 paired-end 100 platform. For the L1 experiment, 25-30 larvae were picked,
802 rinsed with PBS, homogenized in 1mL Trizol, and isolated above. Total RNA Seq libraries
803 were prepared with Nugen Ovation Universal Drosophila kit and sequenced with
804 NOVASEq-S4 paired-end 100 platform.

805

806 **Bioinformatic analyses**

807 For both sequencing experiments, reads were trimmed for adaptor sequence/low-quality
808 sequence using BBduk (bbmap). FastQC was used for quality control [111], and reads were
809 aligned to genome build DM6 using the STAR aligner [112]. Aligned reads were counted
810 with featureCounts [113] and differential expression analyses were completed with DESeq2
811 [114]. Of note, for the L1 data, one genotype ($H3.3^{K36R}H3.2^{HWT}$) from the same sequencing
812 run was included in construction of the DESeq model, but not included in any downstream

813 analysis. k -means clustered heatmaps of z -score differences from RNA Seq data were
814 produced as follows. The combined set of chromosome X DEGs for all mutant genotypes
815 were used for each heatmap. z -scores for each gene were obtained from DESeq2 normalized
816 counts for each replicate. For each gene, z -score differences were obtained by: $z^{replicate} -$
817 $z^{mean_ctrl_reps_both_sexes}$. For each z -score difference, the mean of the most appropriate control
818 genotype was used. Scree plots were used to determine the value of k . The
819 ComplexHeatmap package was used to plot z -score differences [115]. Gene lists for each
820 cluster were exported for downstream analyses of cluster features. Boxplots were made
821 using ggplot2 from the Tidyverse package [116]. Heatmaps displaying median LFC values
822 per bins of MSL or H4K16ac were made using GraphPad Prism for Mac, GraphPad
823 Software, www.graphpad.com. Heatmaps displaying median z -scores of ChIP Seq data per
824 RNA Seq cluster were produced as follows. For modENCODE data files, DM3 aligned
825 bedGraph files were converted to bigwig files using Crossmap [117]. For H3K36me2 ChIP,
826 data was downloaded from SRA, and sequences were trimmed, quality checked, and aligned
827 as above. BAM files from ChIP files were normalized to input files and output to bigwig
828 format using deepTools [118]. For Clamp, MSL2, MSL3, Jasper, and Jil-1 RNAi data
829 generated by previously, DM6 aligned bigwigs were downloaded directly from the GEO
830 repository [28, 32]. BEDTools was used to calculate mean ChIP signal over promoter
831 regions (500bp upstream of the TSS) and gene bodies for each gene [119]. z -scores for mean
832 promoter and gene body ChIP signal were obtained relative to all chrX genes. For each
833 heatmap of median ChIP Seq signal values (Figs. 5D, 5E, 6B, 7G, 7H, SXX) for RNA Seq
834 gene clusters generated in Fig. 5A & Fig. 7E, a median z -score for each cluster for each
835 ChIP dataset was calculated and plotted using the pheatmap package [120]. z -score
836 normalization enabled relative comparisons between different histone modification or
837 chromatin binding protein datasets obtained using different antibodies and conditions.

838 Motif analysis was performed by the SEA (Simple Enrichment Analysis) tool using a
839 predefined set of motifs [65]. Metaplots were generated from modENCODE ChIP data for
840 genes in each RNA heatmap cluster using deepTools [118]. Browser tracks for genomic data
841 were visualized on the Integrated Genomics Viewer (IGV) [121].

842 Statistical analyses for RNA-seq data is as follows. Significant DEGs were
843 determined by DESeq2 with and adjusted *p*-value <0.05. For chromosome arm plots, LFC
844 values of X-chromosome genes were compared to the combined set of large autosome (2L,
845 2R, 3L, and 3R) genes, and *p*-values computed using the Kruskall-Wallis ANOVA, followed
846 by Dunn's multiple comparisons tests. For predominant chromatin state analyses based on
847 [62], Statistical significance of the difference between medians was obtained using the
848 Wilcoxon signed rank test and the Benjamini-Hochberg False Discovery Rate (FDR)
849 multiple comparisons correction.

850
851

852 Acknowledgements

853 We thank the UNC High Throughput Sequencing Facility (HTSF) for their assistance with
854 generating the datasets used here, and members of the Matera laboratory for helpful
855 discussions and critical reading of the manuscript. This work was supported by the
856 National Institutes of Health (NIGMS) grant R35-GM136435 (to A.G.M.).

857
858

859 Figure Legends

860 **Fig. 1 H3.3K36R interacts genetically with H4K16R.** (For all genotypes in B&C,
861 the *H3.3A* gene (chr. 2L) and *H3.3B* gene (chr. X) are WT, and the endogenous

replication-dependent histone gene cluster, *HisC* (chr. 2L) is Δ . The transgenic insertion site *VK33* (chr. 3L, band 65B2) was used for all $12x$ histone transgenes. For each genotype, 2 copies of $12x$ transgenes (*HWT*, *H3.2K36R*, or *H4K16R*) are present in trans. All C&D genotypes, except Oregon R (OR), are *HisC Δ . Status of *H3.3A*, *H3.3B*, and $12x$ transgenes are indicated in the table below the graph in (D). *H3.3B* is either WT or K36R; *H3.3A* WT or null. *HisC* is either intact (WT) or Δ (null). $12xH4$ transgenes contain 12 copies of the histone repeat unit, each containing all five replication-dependent histone genes. Transgenes used in this study carry the following alleles of *H4*: *HWT* or *K16R*. Cartoon of genetic loci used in panel D. Panels A and C were created using BioRender.com. **A)** Cartoon of genetic loci used in panels B-D. For complete genotypes, see Figs. S1, S2, S4, & S5. **(B)** Developmental viability assay. For each genotype, % pupation and % eclosion of 8-10 biological replicates (50 larvae/replicate vial) were calculated, and means and SD of these percentages were plotted. Statistical significance for % eclosion was calculated with GraphPad Prism software using Brown-Forsythe and Welch ANOVA tests, followed by the Dunnett's T3 multiple comparisons test. *** $P < 0.001$. **** $P < 0.0001$. ns, not significant. **(C)** Proportion of male and female eclosed animals were calculated. Statistical significance for sex ratio was calculated with GraphPad Prism software using Fisher's Exact Test, followed by the Benjamini-Hochberg False Discovery Rate (FDR) correction for multiple comparisons ($Q=0.05$). ** $P < 0.01$. **** $P < 0.0001$. ns, not significant. **(D)** Viability assay, as in B. **(E)** Sex ratio of adults, as in C.*

Fig. 2 Transcriptomic analyses of dosage compensation in third instar *Set2* and *H3K36R* larval brains. (A) M/A plots comparing gene expression in WL3 brain from combined male and female replicates of mutants relative to control. Mutants represented

887 from left to right with control genotype in parentheses: $H3.3^{WT}H3.2^{K36R}$ ($H3.3^{WT}H3.2^{HWT}$),
888 $H3.3^{K36R}H3.2^{HWT}$ ($H3.3^{Ctrl}H3.2^{HWT}$), and $Set2^l$ (yw). Magenta and blue dots represent
889 differentially expressed genes (DEGs) that were significantly (adjusted p-value, p-adj <
890 0.05) up- or down-regulated, respectively. The number of DEGs in each direction is shown
891 in the upper and lower corners. **(B)** For DESeq2 analyses separated by sex, all genes with
892 a defined P value (not NA), Log₂ Fold-change values of mutant genotypes in A, relative to
893 controls were plotted for male replicates and binned by chromosome arm. Median Log₂
894 Fold-change values of X-chromosome genes were compared to the combined set of large
895 autosome (2L, 2R, 3L, and 3R) genes, and p-values computed using the Kruskall-Wallis
896 ANOVA, followed by Dunn's multiple comparisons tests. ** $P < 0.01$, **** $P < 10^{-15}$. ns, not
897 significant. **(C)** Same as C, but for female replicates. **(D)** HAS site analysis of mutant males
898 relative to controls. Log₂ Fold-change values of $Set2^l$ mutant males were plotted, binned by
899 distance from chrX HAS sites defined previously [31]. *** $P < 0.001$. **** $P < 0.0001$. ns,
900 not significant. **(E)** Same as D, but for female replicates.

901
902 **Fig. 3 Transcriptomic analyses of dosage compensation in $H3.3^{K36R}$ adults. (A)** For
903 DESeq2 analyses in adult heads separated by sex, and all genes with a defined P value (not
904 NA), Log₂ Fold-change values of $H3.3^{K36R}$ mutants, relative to $H3.3A^{null}$ controls for young
905 (~1 day post-eclosion) and old (~23 day post-eclosion) were plotted for male replicates and
906 binned by chromosome arm. Median Log₂ Fold-change values of X-chromosome genes were
907 compared to the combined set of large autosome (2L, 2R, 3L, and 3R) genes, and p-values
908 computed using the Kruskall-Wallis ANOVA, followed by Dunn's multiple comparisons
909 tests. * $P < 0.05$, **** $P < 10^{-15}$. **(B)** Same as A, but for females. **(C)** HAS site analysis of
910 mutant males relative to controls. Log₂ Fold-change values of $H3.3^{K36R}$ mutant males were
911 plotted, binned by distance from chrX HAS sites defined previously [31]. *** $P < 0.001$.

912 **** $P < 0.0001$. ns, not significant. **(D)** Same as D, but for female replicates. **(E)** For ChrX
913 genes, median Log2 Fold-change values for each group were binned by mean H4K16ac
914 ChIP-seq signal in gene bodies from male adult heads and plotted on a heatmap .
915

916 **Fig. 4 Chromosome X genes with Different Predominant Chromatin States**

917 **Respond Differently to *Set2* and *H3K36R* mutation. (A)** Pie charts depicting
918 predominant chromatin states (defined in [62]) of six *Drosophila* chromosomes in BG3 cells.
919 BEDtools was used to assign genes to a predominant chromatin state. Genes were binned to
920 a given state if > 50% of the gene was marked by that state. Genes where no state color was
921 > 50% of gene length were designated as “Mixed”. Representative histone marks in each
922 state depicted in the legend. A full characterization of each state is described in the source
923 publication [62]. **(B)** Log₂ Fold-change values of mutant genotypes described in Fig. S5 for
924 ChrX genes were plotted separately for genes in the three predominant states on the male
925 X: states 1 (n=499), 5 (n=798), and 8 (n=239). Statistical significance of difference between
926 medians was assessed using the Wilcoxon signed rank test, followed by the Benjamini-
927 Hochberg False Discovery Rate (FDR) correction for multiple comparisons. ** $P < 0.001$.
928 *** $P < 0.001$. **** $P < 0.0001$. ns, not significant.

929

930 **Fig. 5 *k*-means clustering of gene expression reveals heterogeneous regulation of**
931 **Chromosome X genes by *Set2/H3K36R*. (A)** DESeq2 normalized count values for the
932 combined set of Chromosome X DEGs from WL3 brain from all mutant and control
933 genotypes were *z*-score normalized by gene to put all expression values on the same scale
934 (mean=0, SD=1). Differences of individual replicate values for each gene were calculated
935 relative to the mean *z*-score of control replicates. A *k*-means clustered heatmap (*k*=9) was
936 generated from these values. Below the heatmap, genotype and sex of each replicate is

937 indicated. To the left, cluster numbers for subsequent analyses are indicated. N values for
938 each cluster are as follows: c1 (272), c2 (57), c3 (275), c4 (180), c5 (141), c6 (120), c7 (132), c8
939 (219), c9 (205), not differentially expressed (nonDEG, 416). **(B)** Base Mean of normalized
940 DESeq2 counts for DEGs (binned by Cluster Number in panel A) and Non-DEGs (genes not
941 differentially expressed in any mutant genotype). **(C)** Chromosome X genes were grouped
942 by gene expression cluster (Panel A), and a heatmaps of median Log₂ Fold-change values for
943 mutant genotypes, separated by sex, was constructed. **(D)** Mean levels of H3K36me1, me2,
944 and me3 in BG3 cells were calculated for genes from all chromosomes for both promoter
945 and gene body regions. For each methyl state, *z*-scores were computed for all genes. For X-
946 genes, median *z*-score was computed for Panel A heatmap clusters and non-DEGs, and a
947 heatmap of these values was constructed to highlight relative levels of H3K36 modification
948 states between clusters. **(E)** Relative abundance of DC modifications and proteins within
949 gene bodies were calculated and plotted as in (D), using datasets generated in S2 cells.
950

951 **Fig. 6 Motif analysis reveals that Set2/K36 and BEAF-32 regulate common gene**
952 **sets similarly.** **(A)** SEA (Simple Enrichment Analysis) was performed on promoters and
953 gene bodies of gene groups from Fig. 5A using motifs from the FLYREG.v2 database (Fig
954 5A; Fig 6A [64, 65]). Motifs with *q*-value > 0.05 and enrichment value over nonDEG control
955 sequences > 2 were displayed. Motifs not meeting either threshold were designated as ns
956 and colored black. **(B)** A heatmap of insulator protein binding was generated as in Fig. 5D
957 and 5E, with the addition of hierarchical clustering. **(C)** Browser shot of insulator protein
958 binding relative to H3K36me3, H4K16ac, L3 cluster annotation, and WL3 brain male RNA-
959 seq data. RPGC normalized RNA-seq data is colored for each mutant with control data
960 shown in gray on the same track line. **(D)** Scaled gene metaplots of insulator proteins with
961 genes grouped by similar enrichment of BEAF-32 and CP190. SuHw is included for

962 comparison. **(E)** Chromosome arm plots and accompanying statistical analyses were
963 produced as in Fig. 2B from RNA-seq data from insulator transcript knockdowns in BG3
964 cells generated by [66]. **F)** Median LFC values for BEAF-32, BEAF-32 + DREF, and CP190
965 + Chromator were determined for each L3 gene cluster from Fig. 5A and plotted as a
966 heatmap.

967

968 **Fig. 7 Transcriptomic analyses of dosage compensation in first instar *Set2* and**
969 **combined *H3K36R* larvae.** **(A)** M/A plots comparing gene expression changes from mixed
970 sex, whole L1 animals. Mutants represented from left to right with control genotype in
971 parentheses: $H3.3^{Ctrl}H4^{K16R}$ ($H3.3^{Ctrl}H3.2^{HWT}$), $Set2^l$ (yw), and $H3.3^{K36R}H3.2^{K36R}$
972 ($H3.3^{Ctrl}H3.2^{HWT}$). Magenta and blue dots represent differentially expressed genes (DEGs)
973 that were significantly (adjusted p-value, p-adj < 0.05) up- or down-regulated, respectively.
974 The number of DEGs in each direction is shown in the upper and lower corners. **(B)** For all
975 genes with a defined P value (not NA), Log₂ Fold-change values of mutant genotypes in A,
976 relative to controls were plotted and binned by chromosome arm as in Fig.2B,C. **(C)** HAS
977 site analysis of mixed sex animals relative to controls was analysed as in Fig. 2D,E. **(D)**
978 Mean MSL3 levels across gene bodies in S2 cells (see Methods). ChrX genes were binned by
979 MSL3 decile, and heatmaps of median Log₂ Fold-change values for mutant genotypes were
980 constructed. **(E)** A k-means clustered heatmap ($k=6$) was generated as in Fig.5A (see
981 Methods). Genotypes are indicated below heatmap; clusters to the left. N values for each
982 cluster are as follows: c1 (163), c2 (125), c3 (182), c4 (385), c5 (333), c6 (353), not
983 differentially expressed (nonDEG, 456). **(F)** ChrX genes were grouped by gene expression
984 cluster (Panel E), and a heatmap of median Log₂ Fold-change values for mutant genotypes,
985 was constructed. **(G)** Heatmaps of median H3K36me1, me2, and me3 per cluster were

986 constucted as in Fig. 5E, except using ChIP data from S2 cell. **(H)** Relative abundance of
987 DC modifications and proteins within gene bodies were calculated and plotted as in (G).

988

989

990 **References**

- 991 1. Schwämmle T, Schulz EG: **Regulatory principles and mechanisms governing the onset of**
992 **random X-chromosome inactivation.** *Curr Opin Genet Dev* 2023, **81**:102063.
- 993 2. Keller CI, Akhtar A: **The MSL complex: juggling RNA-protein interactions for dosage**
994 **compensation and beyond.** *Curr Opin Genet Dev* 2015, **31**:1-11.
- 995 3. Basilicata MF, Bruel AL, Semplicio G, Valsecchi CIK, Aktaş T, Duffourd Y, Rumpf T, Morton J,
996 Bache I, Szymanski WG, et al: **De novo mutations in MSL3 cause an X-linked syndrome**
997 **marked by impaired histone H4 lysine 16 acetylation.** *Nat Genet* 2018, **50**:1442-1451.
- 998 4. Monserrat J, Morales Torres C, Richardson L, Wilson TS, Patel H, Domart MC, Horswell S,
999 Song OR, Jiang M, Crawford M, et al: **Disruption of the MSL complex inhibits tumour**
1000 **maintenance by exacerbating chromosomal instability.** *Nat Cell Biol* 2021, **23**:401-412.
- 1001 5. Sun Y, Wiese M, Hmadi R, Karayol R, Seyfferth J, Martinez Greene JA, Erdogan NU, Deboutte
1002 W, Arrigoni L, Holz H, et al: **MSL2 ensures biallelic gene expression in mammals.** *Nature*
1003 2023, **624**:173-181.
- 1004 6. Akhtar A, Becker PB: **Activation of transcription through histone H4 acetylation by MOF,**
1005 **an acetyltransferase essential for dosage compensation in Drosophila.** *Mol Cell* 2000,
1006 **5**:367-375.
- 1007 7. Conrad T, Cavalli FM, Holz H, Hallaci E, Kind J, Ilik I, Vaquerizas JM, Luscombe NM, Akhtar A:
1008 **The MOF chromobarrel domain controls genome-wide H4K16 acetylation and**
1009 **spreading of the MSL complex.** *Dev Cell* 2012, **22**:610-624.
- 1010 8. Lee H, Oliver B: **Non-canonical Drosophila X chromosome dosage compensation and**
1011 **repressive topologically associated domains.** *Epigenetics Chromatin* 2018, **11**:62.
- 1012 9. Belote JM, Lucchesi JC: **Male-specific lethal mutations of Drosophila melanogaster.**
1013 *Genetics* 1980, **96**:165-186.
- 1014 10. Meller VH, Wu KH, Roman G, Kuroda MI, Davis RL: **roX1 RNA paints the X chromosome of**
1015 **male Drosophila and is regulated by the dosage compensation system.** *Cell* 1997,
1016 **88**:445-457.
- 1017 11. Tanaka A, Fukunaga A, Oishi K: **Studies on the sex-specific lethals of Drosophila**
1018 **melanogaster. II. Further studies on a male-specific lethal gene, maleless.** *Genetics*
1019 1976, **84**:257-266.
- 1020 12. Ferrari F, Plachetka A, Alekseyenko AA, Jung YL, Ozsolak F, Kharchenko PV, Park PJ, Kuroda
1021 MI: **"Jump start and gain" model for dosage compensation in Drosophila based on**
1022 **direct sequencing of nascent transcripts.** *Cell Rep* 2013, **5**:629-636.
- 1023 13. Larschan E, Bishop EP, Kharchenko PV, Core LJ, Lis JT, Park PJ, Kuroda MI: **X chromosome**
1024 **dosage compensation via enhanced transcriptional elongation in Drosophila.** *Nature*
1025 2011, **471**:115-118.

- 1026 14. Samata M, Akhtar A: **Dosage Compensation of the X Chromosome: A Complex Epigenetic**
1027 **Assignment Involving Chromatin Regulators and Long Noncoding RNAs.** *Annu Rev*
1028 *Biochem* 2018, **87**:323-350.
- 1029 15. Bai X, Larschan E, Kwon SY, Badenhorst P, Kuroda MI: **Regional control of chromatin**
1030 **organization by noncoding roX RNAs and the NURF remodeling complex in Drosophila**
1031 **melanogaster.** *Genetics* 2007, **176**:1491-1499.
- 1032 16. Grimaud C, Becker PB: **The dosage compensation complex shapes the conformation of the**
1033 **X chromosome in Drosophila.** *Genes Dev* 2009, **23**:2490-2495.
- 1034 17. Samata M, Alexiadis A, Richard G, Georgiev P, Nuebler J, Kulkarni T, Renschler G, Basilicata
1035 MF, Zenk FL, Shvedunova M, et al: **Intergenerationally Maintained Histone H4 Lysine 16**
1036 **Acetylation Is Instructive for Future Gene Activation.** *Cell* 2020.
- 1037 18. Sheveliov YY, Ulianov SV, Gelfand MS, Belyakin SN, Razin SV: **Dosage Compensation in**
1038 **Drosophila: Its Canonical and Non-Canonical Mechanisms.** *Int J Mol Sci* 2022, **23**.
- 1039 19. Armstrong RL, Penke TJR, Strahl BD, Matera AG, McKay DJ, MacAlpine DM, Duronio RJ:
1040 **Chromatin conformation and transcriptional activity are permissive regulators of DNA**
1041 **replication initiation in Drosophila.** *Genome Res* 2018, **28**:1688-1700.
- 1042 20. Copur O, Gorchakov A, Finkl K, Kuroda MI, Muller J: **Sex-specific phenotypes of histone H4**
1043 **point mutants establish dosage compensation as the critical function of H4K16**
1044 **acetylation in Drosophila.** *Proc Natl Acad Sci U S A* 2018.
- 1045 21. Kuroda MI, Hilfiker A, Lucchesi JC: **Dosage Compensation in Drosophila—a Model for the**
1046 **Coordinate Regulation of Transcription.** *Genetics* 2016, **204**:435-450.
- 1047 22. Bashaw GJ, Baker BS: **The msl-2 dosage compensation gene of Drosophila encodes a**
1048 **putative DNA-binding protein whose expression is sex specifically regulated by Sex-**
1049 **lethal.** *Development* 1995, **121**:3245-3258.
- 1050 23. Kelley RL, Solovyeva I, Lyman LM, Richman R, Solovyev V, Kuroda MI: **Expression of msl-2**
1051 **causes assembly of dosage compensation regulators on the X chromosomes and**
1052 **female lethality in Drosophila.** *Cell* 1995, **81**:867-877.
- 1053 24. Sheikh BN, Guhathakurta S, Akhtar A: **The non-specific lethal (NSL) complex at the**
1054 **crossroads of transcriptional control and cellular homeostasis.** *EMBO Rep* 2019,
1055 **20**:e47630.
- 1056 25. Lam KC, Muhlford F, Vaquerizas JM, Raja SJ, Holz H, Luscombe NM, Manke T, Akhtar A:
1057 **The NSL complex regulates housekeeping genes in Drosophila.** *PLoS Genet* 2012,
1058 **8**:e1002736.
- 1059 26. Raja SJ, Charapitsa I, Conrad T, Vaquerizas JM, Gebhardt P, Holz H, Kadlec J, Fraterman S,
1060 Luscombe NM, Akhtar A: **The nonspecific lethal complex is a transcriptional regulator in**
1061 **Drosophila.** *Mol Cell* 2010, **38**:827-841.
- 1062 27. Wang CI, Alekseyenko AA, LeRoy G, Elia AE, Gorchakov AA, Britton LM, Elledge SJ,
1063 Kharchenko PV, Garcia BA, Kuroda MI: **Chromatin proteins captured by ChIP-mass**
1064 **spectrometry are linked to dosage compensation in Drosophila.** *Nat Struct Mol Biol*
1065 **2013, 20**:202-209.
- 1066 28. Albig C, Wang C, Dann GP, Wojcik F, Schauer T, Krause S, Maenner S, Cai W, Li Y, Girton J, et
1067 al: **JASPer controls interphase histone H3S10 phosphorylation by chromosomal kinase**
1068 **JIL-1 in Drosophila.** *Nat Commun* 2019, **10**:5343.

- 1069 29. Cugusi S, Ramos E, Ling H, Yokoyama R, Luk KM, Lucchesi JC: **Topoisomerase II plays a role**
1070 **in dosage compensation in Drosophila.** *Transcription* 2013, **4**:238-250.
- 1071 30. Lerach S, Zhang W, Deng H, Bao X, Girton J, Johansen J, Johansen KM: **JIL-1 kinase, a**
1072 **member of the male-specific lethal (MSL) complex, is necessary for proper dosage**
1073 **compensation of eye pigmentation in Drosophila.** *Genesis* 2005, **43**:213-215.
- 1074 31. Straub T, Grimaud C, Gilfillan GD, Mitterweger A, Becker PB: **The Chromosomal High-**
1075 **Affinity Binding Sites for the Drosophila Dosage Compensation Complex.** *PLoS Genetics*
1076 2008, **4**:e1000302.
- 1077 32. Albig C, Tikhonova E, Krause S, Maksimenko O, Regnard C, Becker PB: **Factor cooperation**
1078 **for chromosome discrimination in Drosophila.** *Nucleic Acids Res* 2019, **47**:1706-1724.
- 1079 33. Soruco MM, Chery J, Bishop EP, Siggers T, Tolstorukov MY, Leydon AR, Sugden AU, Goebel K,
1080 Feng J, Xia P, et al: **The CLAMP protein links the MSL complex to the X chromosome**
1081 **during Drosophila dosage compensation.** *Genes Dev* 2013, **27**:1551-1556.
- 1082 34. Villa R, Schauer T, Smialowski P, Straub T, Becker PB: **PionX sites mark the X chromosome**
1083 **for dosage compensation.** *Nature* 2016, **537**:244-248.
- 1084 35. Alekseyenko AA, Larschan E, Lai WR, Park PJ, Kuroda MI: **High-resolution ChIP-chip analysis**
1085 **reveals that the Drosophila MSL complex selectively identifies active genes on the**
1086 **male X chromosome.** *Genes Dev* 2006, **20**:848-857.
- 1087 36. Sural TH, Peng S, Li B, Workman JL, Park PJ, Kuroda MI: **The MSL3 chromodomain directs a**
1088 **key targeting step for dosage compensation of the Drosophila melanogaster X**
1089 **chromosome.** *Nat Struct Mol Biol* 2008, **15**:1318-1325.
- 1090 37. Larschan E, Alekseyenko AA, Gortchakov AA, Peng S, Li B, Yang P, Workman JL, Park PJ,
1091 Kuroda MI: **MSL complex is attracted to genes marked by H3K36 trimethylation using a**
1092 **sequence-independent mechanism.** *Mol Cell* 2007, **28**:121-133.
- 1093 38. Yokoyama R, Pannuti A, Ling H, Smith ER, Lucchesi JC: **A plasmid model system shows that**
1094 **Drosophila dosage compensation depends on the global acetylation of histone H4 at**
1095 **lysine 16 and is not affected by depletion of common transcription elongation**
1096 **chromatin marks.** *Mol Cell Biol* 2007, **27**:7865-7870.
- 1097 39. Lindehell H, Glotov A, Dorafshan E, Schwartz YB, Larsson J: **The role of H3K36 methylation**
1098 **and associated methyltransferases in chromosome-specific gene regulation.** *Sci Adv*
1099 2021, **7**:eab4390.
- 1100 40. Cornett EM, Ferry L, Defossez PA, Rothbart SB: **Lysine Methylation Regulators**
1101 **Moonlighting outside the Epigenome.** *Mol Cell* 2019, **75**:1092-1101.
- 1102 41. Huang J, Berger SL: **The emerging field of dynamic lysine methylation of non-histone**
1103 **proteins.** *Curr Opin Genet Dev* 2008, **18**:152-158.
- 1104 42. Morgan MAJ, Shilatifard A: **Reevaluating the roles of histone-modifying enzymes and their**
1105 **associated chromatin modifications in transcriptional regulation.** *Nat Genet* 2020,
1106 **52**:1271-1281.
- 1107 43. Bhattacharya S, Levy MJ, Zhang N, Li H, Florens L, Washburn MP, Workman JL: **The**
1108 **methyltransferase SETD2 couples transcription and splicing by engaging mRNA**
1109 **processing factors through its SHI domain.** *Nat Commun* 2021, **12**:1443.
- 1110 44. Khan A, Metts JM, Collins LC, Mills CA, Li K, Brademeyer AL, Bowman BM, Major MB, Aubé
1111 J, Herring LE, et al: **SETD2 maintains nuclear lamina stability to safeguard the genome.**
1112 *bioRxiv* 2023.

- 1113 45. Molenaar TM, van Leeuwen F: **SETD2: from chromatin modifier to multipronged regulator**
1114 **of the genome and beyond.** *Cell Mol Life Sci* 2022, **79**:346.
- 1115 46. DiBartolomeis SM, Tartof KD, Jackson FR: **A superfamily of Drosophila satellite related (SR)**
1116 **DNA repeats restricted to the X chromosome euchromatin.** *Nucleic Acids Res* 1992,
1117 **20**:1113-1116.
- 1118 47. Jayakrishnan M, Havlová M, Veverka V, Regnard C, Becker PB: **Genomic context-dependent**
1119 **histone H3K36 methylation by three Drosophila methyltransferases and**
1120 **implications for dedicated chromatin readers.** *bioRxiv* 2024:2024.2002.2006.577191.
- 1121 48. Kaufman TC: **A Short History and Description of Drosophila melanogaster Classical**
1122 **Genetics: Chromosome Aberrations, Forward Genetic Screens, and the Nature of**
1123 **Mutations.** *Genetics* 2017, **206**:665-689.
- 1124 49. Ramirez F, Lingg T, Toscano S, Lam KC, Georgiev P, Chung HR, Lajoie BR, de Wit E, Zhan Y, de
1125 Laat W, et al: **High-Affinity Sites Form an Interaction Network to Facilitate Spreading of**
1126 **the MSL Complex across the X Chromosome in Drosophila.** *Mol Cell* 2015, **60**:146-162.
- 1127 50. Waring GL, Pollack JC: **Cloning and characterization of a dispersed, multicopy, X**
1128 **chromosome sequence in Drosophila melanogaster.** *Proc Natl Acad Sci U S A* 1987,
1129 **84**:2843-2847.
- 1130 51. Salzler HR, Vandadi V, McMichael BD, Brown JC, Boerma SA, Leatham-Jensen MP, Adams
1131 KM, Meers MP, Simon JM, Duronio RJ, et al: **Distinct roles for canonical and variant**
1132 **histone H3 lysine-36 in Polycomb silencing.** *Sci Adv* 2023, **9**:eadf2451.
- 1133 52. Lucchesi JC, Kuroda MI: **Dosage compensation in Drosophila.** *Cold Spring Harb Perspect*
1134 **Biol** 2015, **7**.
- 1135 53. Meers MP, Henriques T, Lavender CA, McKay DJ, Strahl BD, Duronio RJ, Adelman K, Matera
1136 AG: **Histone gene replacement reveals a post-transcriptional role for H3K36 in**
1137 **maintaining metazoan transcriptome fidelity.** *Elife* 2017, **6**.
- 1138 54. Costanzo M, VanderSluis B, Koch EN, Baryshnikova A, Pons C, Tan G, Wang W, Usaj M,
1139 Hanchard J, Lee SD, et al: **A global genetic interaction network maps a wiring diagram**
1140 **of cellular function.** *Science* 2016, **353**.
- 1141 55. Housden BE, Valvezan AJ, Kelley C, Sopko R, Hu Y, Roesel C, Lin S, Buckner M, Tao R, Yilmazel
1142 B, et al: **Identification of potential drug targets for tuberous sclerosis complex by**
1143 **synthetic screens combining CRISPR-based knockouts with RNAi.** *Sci Signal* 2015, **8**:rs9.
- 1144 56. Brown JC, McMichael BD, Vandadi V, Mukherjee A, Salzler HR, Matera AG: **Lysine-36 of**
1145 **Drosophila histone H3.3 supports adult longevity.** *G3 (Bethesda)* 2024, **14**.
- 1146 57. Valsecchi CIK, Basilicata MF, Georgiev P, Gaub A, Seyfferth J, Kulkarni T, Panhale A, Semplicio
1147 G, Manjunath V, Holz H, et al: **RNA nucleation by MSL2 induces selective X chromosome**
1148 **compartmentalization.** *Nature* 2021, **589**:137-142.
- 1149 58. Straub T, Becker PB: **Transcription modulation chromosome-wide: universal features and**
1150 **principles of dosage compensation in worms and flies.** *Curr Opin Genet Dev* 2011,
1151 **21**:147-153.
- 1152 59. Konev AY, Tribus M, Park SY, Podhraski V, Lim CY, Emelyanov AV, Vershilova E, Pirrotta V,
1153 Kadonaga JT, Lusser A, Fyodorov DV: **CHD1 motor protein is required for deposition of**
1154 **histone variant H3.3 into chromatin in vivo.** *Science* 2007, **317**:1087-1090.

- 1155 60. Schoberleitner I, Bauer I, Huang A, Andreyeva EN, Sebald J, Pascher K, Rieder D, Brunner M,
1156 Podhraski V, Oemer G, et al: **CHD1 controls H3.3 incorporation in adult brain chromatin**
1157 **to maintain metabolic homeostasis and normal lifespan.** *Cell Rep* 2021, **37**:109769.
- 1158 61. Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de
1159 Castro IJ, Kerkhoven RM, Bussemaker HJ, van Steensel B: **Systematic protein location**
1160 **mapping reveals five principal chromatin types in Drosophila cells.** *Cell* 2010, **143**:212-
1161 224.
- 1162 62. Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ,
1163 Larschan E, Gorchakov AA, Gu T, et al: **Comprehensive analysis of the chromatin**
1164 **landscape in Drosophila melanogaster.** *Nature* 2011, **471**:480-485.
- 1165 63. Chaouch A, Berlandi J, Chen CCL, Frey F, Badini S, Harutyunyan AS, Chen X, Krug B, Hébert S,
1166 Jeibmann A, et al: **Histone H3.3 K27M and K36M mutations de-repress transposable**
1167 **elements through perturbation of antagonistic chromatin marks.** *Molecular Cell* 2021,
1168 **81**:4876-4890.e4877.
- 1169 64. Bailey TL, Grant CE: **SEA: Simple Enrichment Analysis of motifs.** *bioRxiv*
1170 2021:2021.2008.2023.457422.
- 1171 65. Bergman CM, Carlson JW, Celtniker SE: **Drosophila DNase I footprint database: a systematic**
1172 **genome annotation of transcription factor binding sites in the fruitfly, Drosophila**
1173 **melanogaster.** *Bioinformatics* 2005, **21**:1747-1749.
- 1174 66. Chathoth KT, Mikheeva LA, Crevel G, Wolfe JC, Hunter I, Beckett-Doyle S, Cotterill S, Dai H,
1175 Harrison A, Zabet NR: **The role of insulators and transcription in 3D chromatin**
1176 **organization of flies.** *Genome Res* 2022, **32**:682-698.
- 1177 67. Lhoumaud P, Hennion M, Gamot A, Cuddapah S, Queille S, Liang J, Micas G, Morillon P,
1178 Urbach S, Bouchez O, et al: **Insulators recruit histone methyltransferase dMes4 to**
1179 **regulate chromatin of flanking genes.** *Embo j* 2014, **33**:1599-1613.
- 1180 68. Shrestha S, Oh DH, McKown JK, Dassanayake M, Hart CM: **4C-seq characterization of**
1181 **Drosophila BEAF binding regions provides evidence for highly variable long-distance**
1182 **interactions between active chromatin.** *PLoS One* 2018, **13**:e0203843.
- 1183 69. Yang J, Corces VG: **Chromatin insulators: a role in nuclear organization and gene**
1184 **expression.** *Adv Cancer Res* 2011, **110**:43-76.
- 1185 70. Tue NT, Yoshioka Y, Mizoguchi M, Yoshida H, Zurita M, Yamaguchi M: **DREF plays multiple**
1186 **roles during Drosophila development.** *Biochim Biophys Acta Gene Regul Mech* 2017,
1187 **1860**:705-712.
- 1188 71. Valadez-Graham V, Yoshioka Y, Velazquez O, Kawamori A, Vázquez M, Neumann A,
1189 Yamaguchi M, Zurita M: **XNP/dATRX interacts with DREF in the chromatin to regulate**
1190 **gene expression.** *Nucleic Acids Research* 2012, **40**:1460-1474.
- 1191 72. Gurudatta BV, Yang J, Van Bortle K, Donlin-Asp PG, Corces VG: **Dynamic changes in the**
1192 **genomic localization of DNA replication-related element binding factor during the cell**
1193 **cycle.** *Cell Cycle* 2013, **12**:1605-1615.
- 1194 73. Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, Habermann B, Akhtar A,
1195 Manke T: **High-resolution TADs reveal DNA sequences underlying genome organization**
1196 **in flies.** *Nat Commun* 2018, **9**:189.
- 1197 74. Bushey AM, Ramos E, Corces VG: **Three subclasses of a Drosophila insulator show distinct**
1198 **and cell type-specific genomic distributions.** *Genes Dev* 2009, **23**:1338-1350.

- 1199 75. Roy S, Gilbert MK, Hart CM: **Characterization of BEAF mutations isolated by homologous**
1200 **recombination in Drosophila.** *Genetics* 2007, **176**:801-813.
- 1201 76. Gurudatta BV, Ramos E, Corces VG: **The BEAF insulator regulates genes involved in cell**
1202 **polarity and neoplastic growth.** *Dev Biol* 2012, **369**:124-132.
- 1203 77. Venkatesh S, Smolle M, Li H, Gogol MM, Saint M, Kumar S, Natarajan K, Workman JL: **Set2**
1204 **methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes.**
1205 *Nature* 2012, **489**:452-455.
- 1206 78. Ginsburg DS, Anlembom TE, Wang J, Patel SR, Li B, Hinnebusch AG: **NuA4 links methylation**
1207 **of histone H3 lysines 4 and 36 to acetylation of histones H4 and H3.** *J Biol Chem* 2014,
1208 **289**:32656-32670.
- 1209 79. Koutná E, Lux V, Kouba T, Škerlová J, Nováček J, Srb P, Hexnerová R, Šváchová H, Kukačka Z,
1210 **Novák P, et al: Multivalency of nucleosome recognition by LEDGF.** *Nucleic Acids Res*
1211 2023, **51**:10011-10025.
- 1212 80. Venkatesh S, Workman JL: **Set2 mediated H3 lysine 36 methylation: regulation of**
1213 **transcription elongation and implications in organismal development.** *Wiley Interdiscip*
1214 *Rev Dev Biol* 2013, **2**:685-700.
- 1215 81. Lee KY, Ranger M, Meneghini MD: **Combinatorial Genetic Control of Rpd3S Through**
1216 **Histone H3K4 and H3K36 Methylation in Budding Yeast.** *G3 (Bethesda)* 2018, **8**:3411-
1217 3420.
- 1218 82. Li B, Gogol M, Carey M, Lee D, Seidel C, Workman JL: **Combined action of PHD and chromo**
1219 **domains directs the Rpd3S HDAC to transcribed chromatin.** *Science* 2007, **316**:1050-
1220 1054.
- 1221 83. Bell O, Wirbelauer C, Hild M, Scharf AN, Schwaiger M, MacAlpine DM, Zilberman F, van
1222 Leeuwen F, Bell SP, Imhof A, et al: **Localized H3K36 methylation states define histone**
1223 **H4K16 acetylation during transcriptional elongation in Drosophila.** *EMBO J* 2007,
1224 **26**:4974-4984.
- 1225 84. Larschan E, Bishop EP, Kharchenko PV, Core LJ, Lis JT, Park PJ, Kuroda MI: **X chromosome**
1226 **dosage compensation via enhanced transcriptional elongation in Drosophila.** *Nature*
1227 2011, **471**:115-118.
- 1228 85. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K: **High-**
1229 **resolution profiling of histone methylations in the human genome.** *Cell* 2007, **129**:823-
1230 837.
- 1231 86. Beck DB, Oda H, Shen SS, Reinberg D: **PR-Set7 and H4K20me1: at the crossroads of**
1232 **genome integrity, cell cycle, chromosome condensation, and transcription.** *Genes Dev*
1233 2012, **26**:325-337.
- 1234 87. Crain AT, Butler MB, Hill CA, Huynh M, McGinty RK, Duronio RJ: **Drosophila melanogaster**
1235 **Set8 and L(3)mbt function in gene expression independently of histone H4 lysine 20**
1236 **methylation.** *bioRxiv* 2024.
- 1237 88. Buscaino A, Legube G, Akhtar A: **X-chromosome targeting and dosage compensation are**
1238 **mediated by distinct domains in MSL-3.** *EMBO Rep* 2006, **7**:531-538.
- 1239 89. Kim D, Blus BJ, Chandra V, Huang P, Rastinejad F, Khorasanizadeh S: **Corecognition of DNA**
1240 **and a methylated histone tail by the MSL3 chromodomain.** *Nat Struct Mol Biol* 2010,
1241 **17**:1027-1029.

- 1242 90. Moore SA, Ferhatoglu Y, Jia Y, Al-Jiab RA, Scott MJ: **Structural and biochemical studies on**
1243 **the chromo-barrel domain of male specific lethal 3 (MSL3) reveal a binding preference**
1244 **for mono- or dimethyllysine 20 on histone H4.** *J Biol Chem* 2010, **285**:40879-40890.
- 1245 91. Deshpande N, Meller VH: **Chromatin That Guides Dosage Compensation Is Modulated by**
1246 **the siRNA Pathway in Drosophila melanogaster.** *Genetics* 2018, **209**:1085-1097.
- 1247 92. Menon DU, Coarfa C, Xiao W, Gunaratne PH, Meller VH: **siRNAs from an X-linked satellite**
1248 **repeat promote X-chromosome recognition in Drosophila melanogaster.** *Proc Natl Acad*
1249 *Sci U S A* 2014, **111**:16460-16465.
- 1250 93. Schauer T, Ghavi-Helm Y, Sexton T, Albig C, Regnard C, Cavalli G, Furlong EE, Becker PB:
1251 **Chromosome topology guides the Drosophila Dosage Compensation Complex for**
1252 **target gene activation.** *EMBO Rep* 2017, **18**:1854-1868.
- 1253 94. Pal K, Forcato M, Jost D, Sexton T, Vaillant C, Salviato E, Mazza EMC, Lugli E, Cavalli G,
1254 Ferrari F: **Global chromatin conformation differences in the Drosophila dosage**
1255 **compensated chromosome X.** *Nat Commun* 2019, **10**:5355.
- 1256 95. Soruco MM, Larschan E: **A new player in X identification: the CLAMP protein is a key**
1257 **factor in Drosophila dosage compensation.** *Chromosome Res* 2014, **22**:505-515.
- 1258 96. Urban JA, Doherty CA, Jordan WT, 3rd, Bliss JE, Feng J, Soruco MM, Rieder LE, Tsiarli MA,
1259 Larschan EN: **The essential Drosophila CLAMP protein differentially regulates non-**
1260 **coding roX RNAs in male and females.** *Chromosome Res* 2017, **25**:101-113.
- 1261 97. Jordan W, 3rd, Larschan E: **The zinc finger protein CLAMP promotes long-range chromatin**
1262 **interactions that mediate dosage compensation of the Drosophila male X-**
1263 **chromosome.** *Epigenetics Chromatin* 2021, **14**:29.
- 1264 98. Duan J, Rieder L, Colonna MM, Huang A, McKenney M, Watters S, Deshpande G, Jordan
1265 W, Fawzi N, Larschan E: **CLAMP and Zelda function together to promote Drosophila**
1266 **zygotic genome activation.** *Elife* 2021, **10**.
- 1267 99. Kuzu G, Kaye EG, Chery J, Siggers T, Yang L, Dobson JR, Boor S, Bliss J, Liu W, Jogl G, et al:
1268 **Expansion of GA Dinucleotide Repeats Increases the Density of CLAMP Binding Sites**
1269 **on the X-Chromosome to Promote Drosophila Dosage Compensation.** *PLoS Genet* 2016,
1270 **12**:e1006120.
- 1271 100. Rieder LE, Koreski KP, Boltz KA, Kuzu G, Urban JA, Bowman SK, Zeidman A, Jordan WT, 3rd,
1272 Tolstorukov MY, Marzluff WF, et al: **Histone locus regulation by the Drosophila dosage**
1273 **compensation adaptor protein CLAMP.** *Genes Dev* 2017, **31**:1494-1508.
- 1274 101. Kaye EG, Kurbidaeva A, Wolle D, Aoki T, Schedl P, Larschan E: **Drosophila Dosage**
1275 **Compensation Loci Associate with a Boundary-Forming Insulator Complex.** *Mol Cell*
1276 *Biol* 2017, **37**.
- 1277 102. Bag I, Dale RK, Palmer C, Lei EP: **The zinc-finger protein CLAMP promotes gypsy chromatin**
1278 **insulator function in Drosophila.** *J Cell Sci* 2019, **132**.
- 1279 103. Urban JA, Urban JM, Kuzu G, Larschan EN: **The Drosophila CLAMP protein associates with**
1280 **diverse proteins on chromatin.** *PLoS One* 2017, **12**:e0189772.
- 1281 104. Makki R, Meller VH: **When Down Is Up: Heterochromatin, Nuclear Organization and X**
1282 **Upregulation.** *Cells* 2021, **10**.
- 1283 105. McCarthy A, Sarkar K, Martin ET, Upadhyay M, Jang S, Williams ND, Forni PE, Buszczak M,
1284 Rangan P: **Msl3 promotes germline stem cell differentiation in female Drosophila.**
1285 *Development* 2022, **149**.

- 1286 106. Prayitno K, Schauer T, Regnard C, Becker PB: **Progressive dosage compensation during**
1287 **Drosophila embryogenesis is reflected by gene arrangement.** *EMBO Rep* 2019,
1288 **20**:e48138.
- 1289 107. McKay DJ, Klusza S, Penke TJ, Meers MP, Curry KP, McDaniel SL, Malek PY, Cooper SW,
1290 Tatomer DC, Lieb JD, et al: **Interrogating the function of metazoan histones using**
1291 **engineered gene clusters.** *Dev Cell* 2015, **32**:373-386.
- 1292 108. Gunesdogan U, Jackle H, Herzig A: **A genetic system to assess in vivo the functions of**
1293 **histones and histone modifications in higher eukaryotes.** *EMBO Rep* 2010, **11**:772-776.
- 1294 109. Sakai A, Schwartz BE, Goldstein S, Ahmad K: **Transcriptional and developmental functions**
1295 **of the H3.3 histone variant in Drosophila.** *Curr Biol* 2009, **19**:1816-1820.
- 1296 110. Gramates LS, Agapite J, Attrill H, Calvi BR, Crosby MA, Dos Santos G, Goodman JL, Goutte-
1297 Gattat D, Jenkins VK, Kaufman T, et al: **FlyBase: a guided tour of highlighted features.**
1298 *Genetics* 2022, **220**.
- 1299 111. Andrews S: **FastQC: A Quality Control Tool for High Throughput Sequence Data.**
1300 Cambridge, United Kingdom: Babraham Bioinformatics, Babraham Institute; 2010.
- 1301 112. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras
1302 TR: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics* 2013, **29**:15-21.
- 1303 113. Liao Y, Smyth GK, Shi W: **featureCounts: an efficient general purpose program for**
1304 **assigning sequence reads to genomic features.** *Bioinformatics* 2014, **30**:923-930.
- 1305 114. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for**
1306 **RNA-seq data with DESeq2.** *Genome Biology* 2014, **15**.
- 1307 115. Gu Z, Eils R, Schlesner M: **Complex heatmaps reveal patterns and correlations in**
1308 **multidimensional genomic data.** *Bioinformatics* 2016, **32**:2847-2849.
- 1309 116. Wickham H, Averick M, Bryan J, Chang W, McGowan LDA, François R, Grolemund G, Hayes
1310 A, Henry L, Hester J: **Welcome to the Tidyverse.** *Journal of open source software* 2019,
1311 **4**:1686.
- 1312 117. Zhao H, Sun Z, Wang J, Huang H, Kocher JP, Wang L: **CrossMap: a versatile tool for**
1313 **coordinate conversion between genome assemblies.** *Bioinformatics* 2014, **30**:1006-
1314 **1007.**
- 1315 118. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke
1316 T: **deepTools2: a next generation web server for deep-sequencing data analysis.** *Nucleic*
1317 *Acids Res* 2016, **44**:W160-165.
- 1318 119. Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features.**
1319 *Bioinformatics* 2010, **26**:841-842.
- 1320 120. Kolde R: **Pheatmap: pretty heatmaps. R package version 1.0. 12. 2019.** 2019.
- 1321 121. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP:
1322 **Integrative genomics viewer.** *Nat Biotechnol* 2011, **29**:24-26.
- 1323

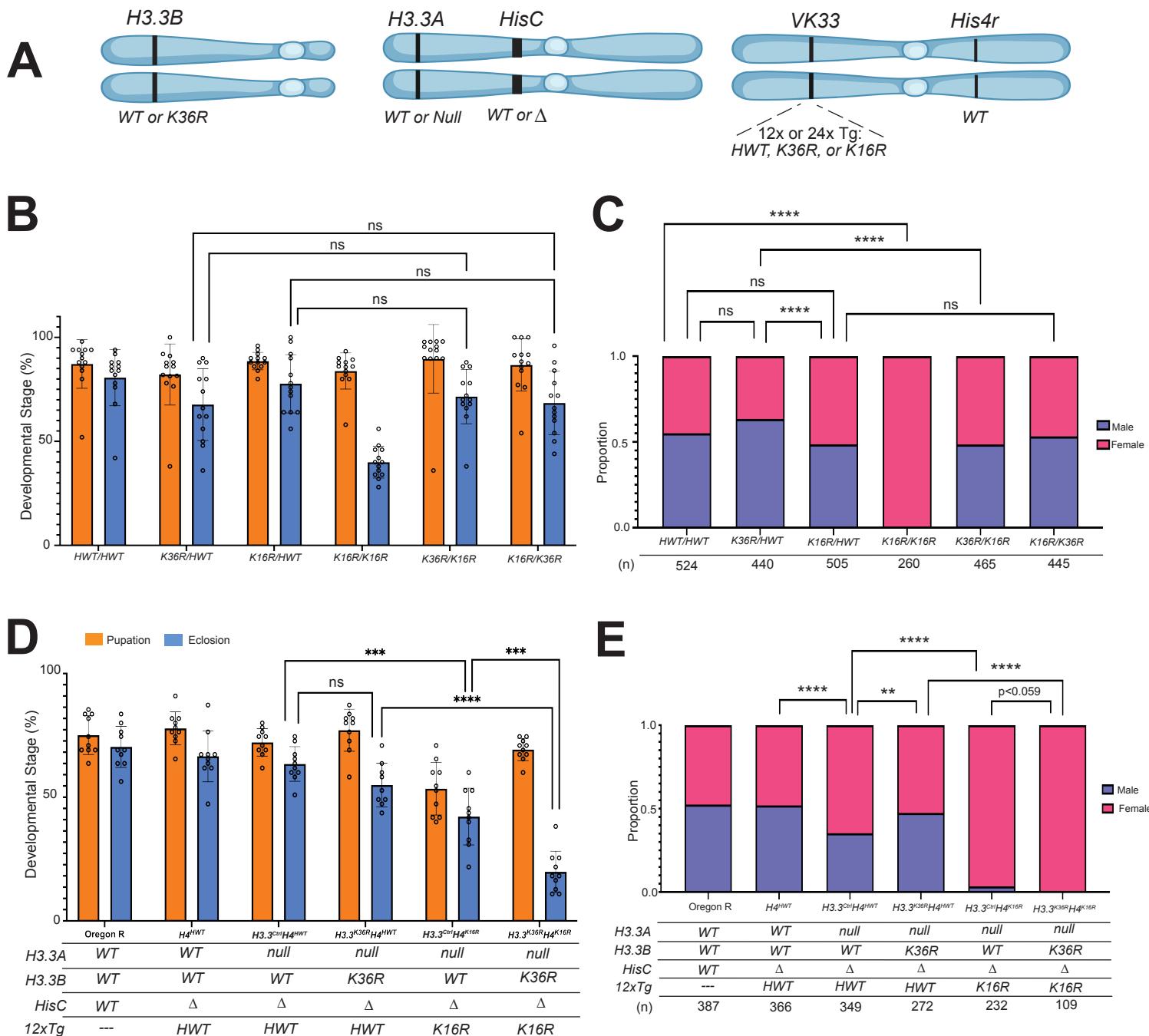


Figure 1

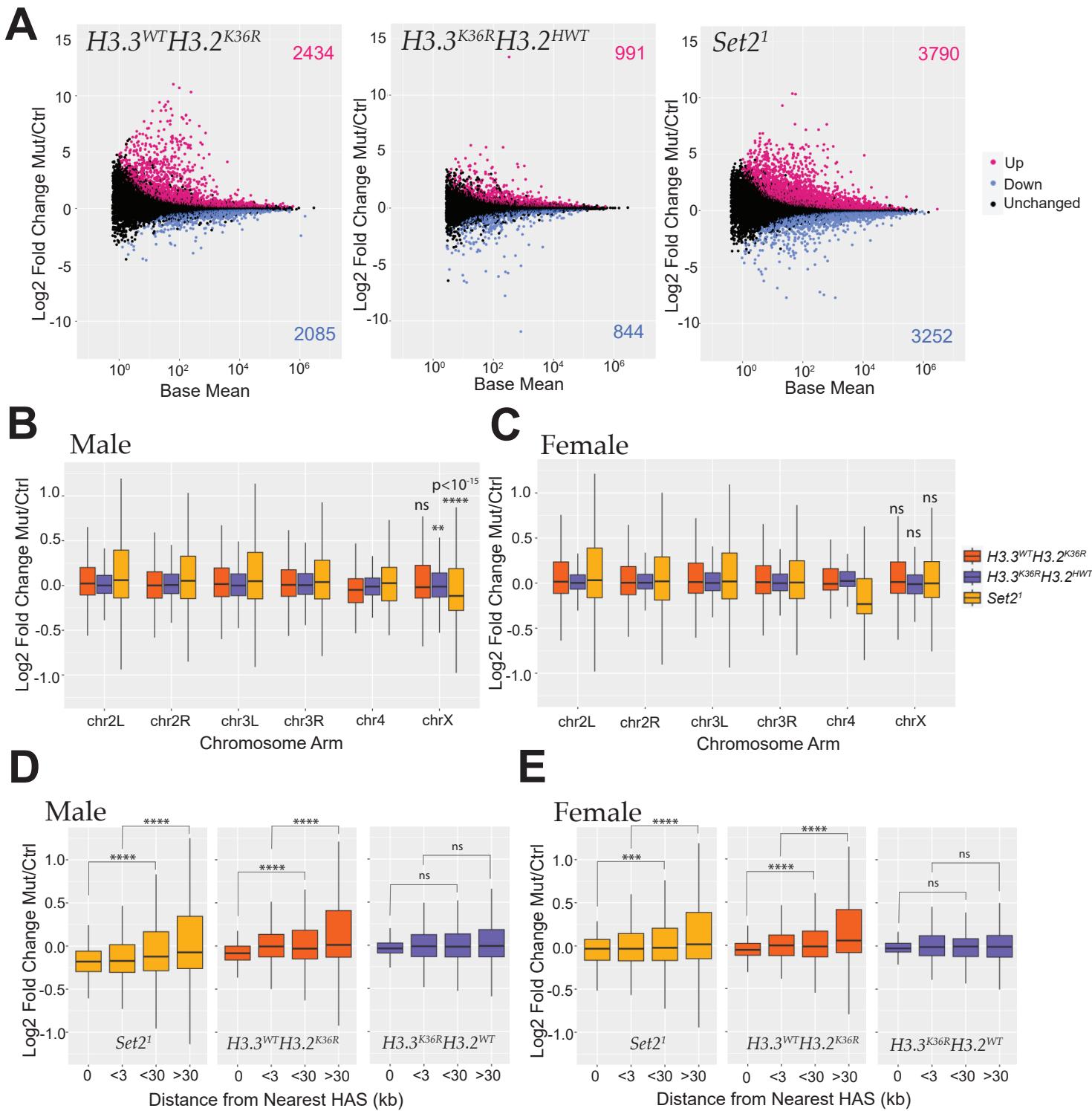


Figure2

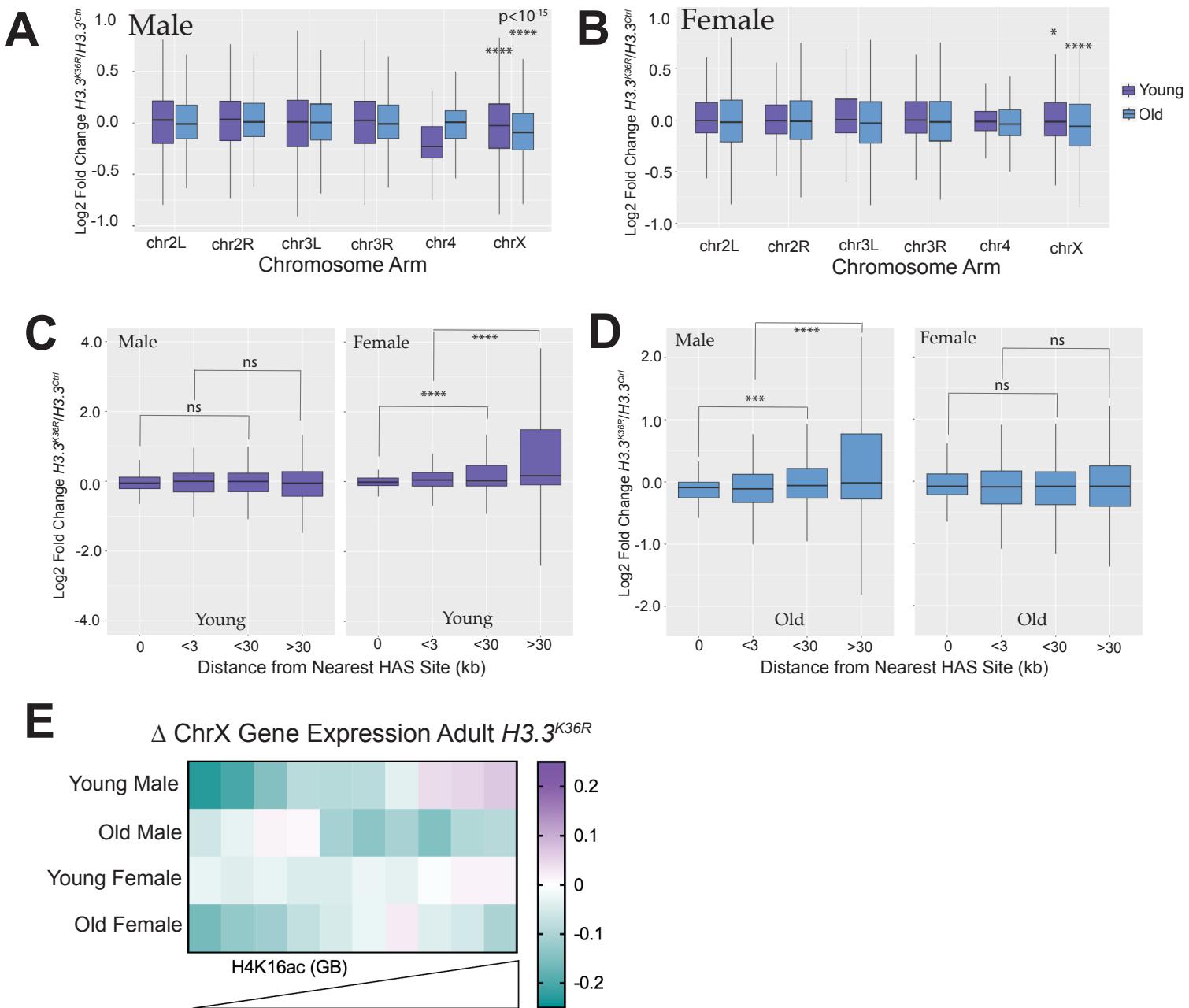
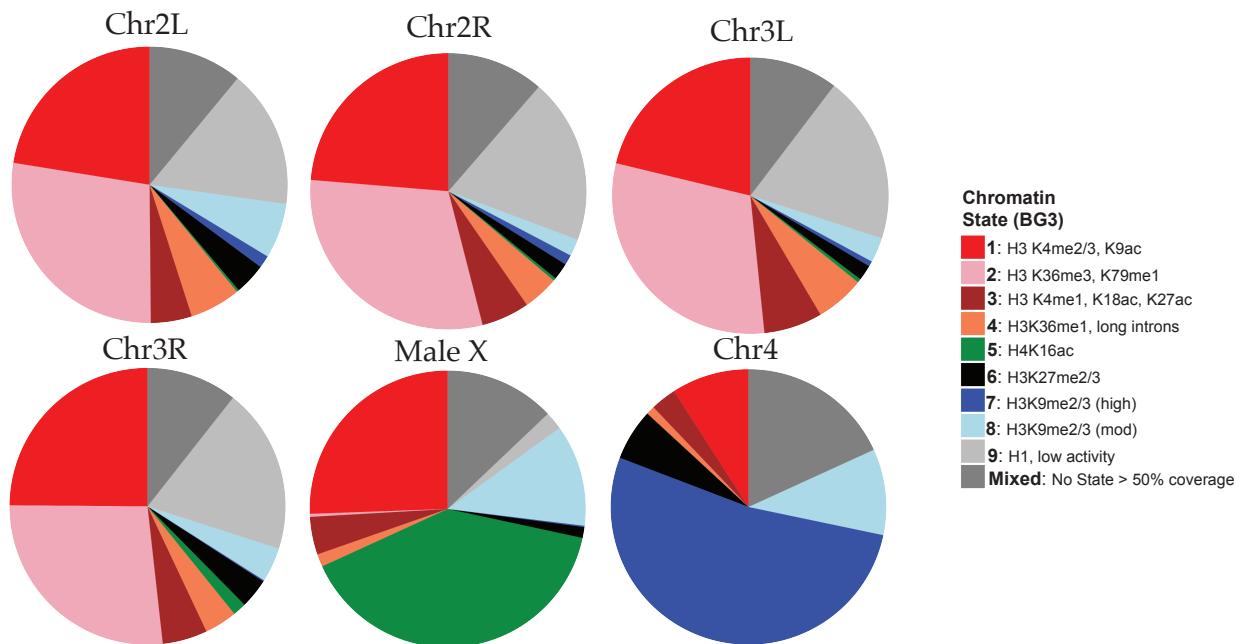


Figure 3

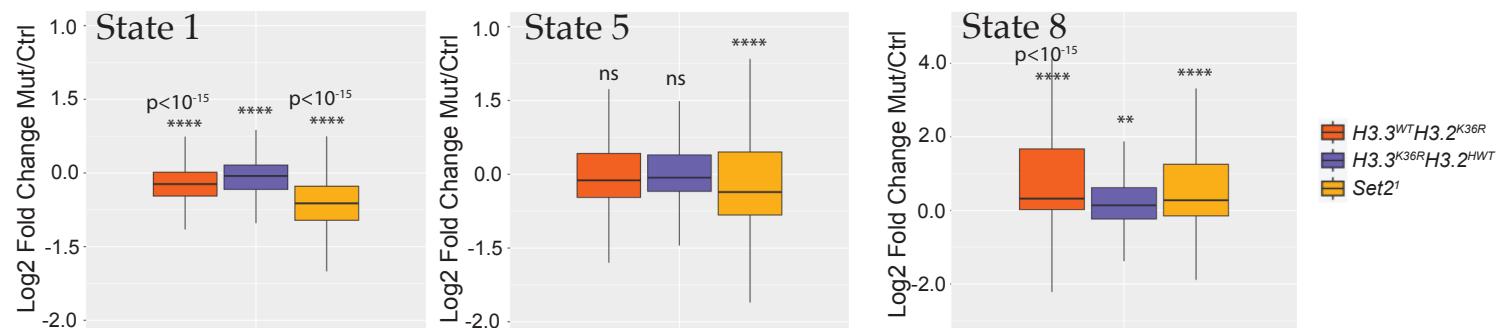
A



B

Chromosome X Genes

Male



Female

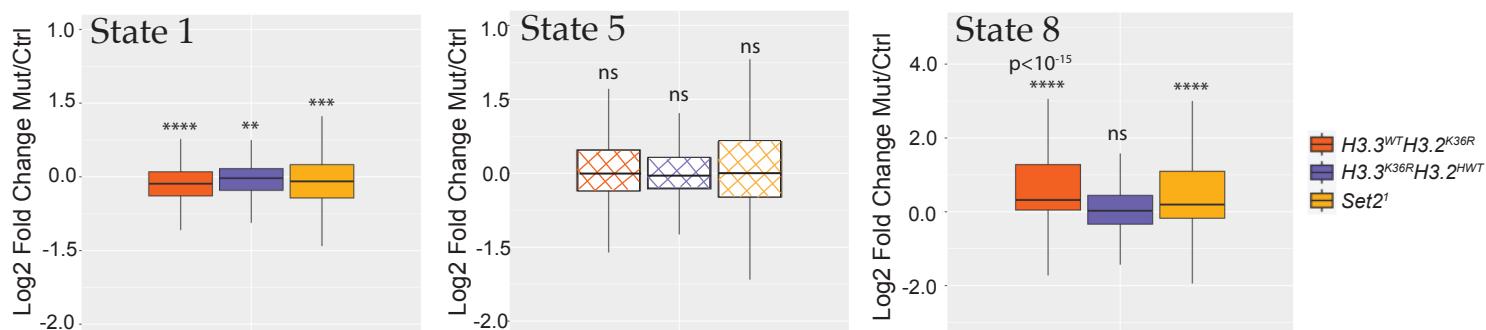
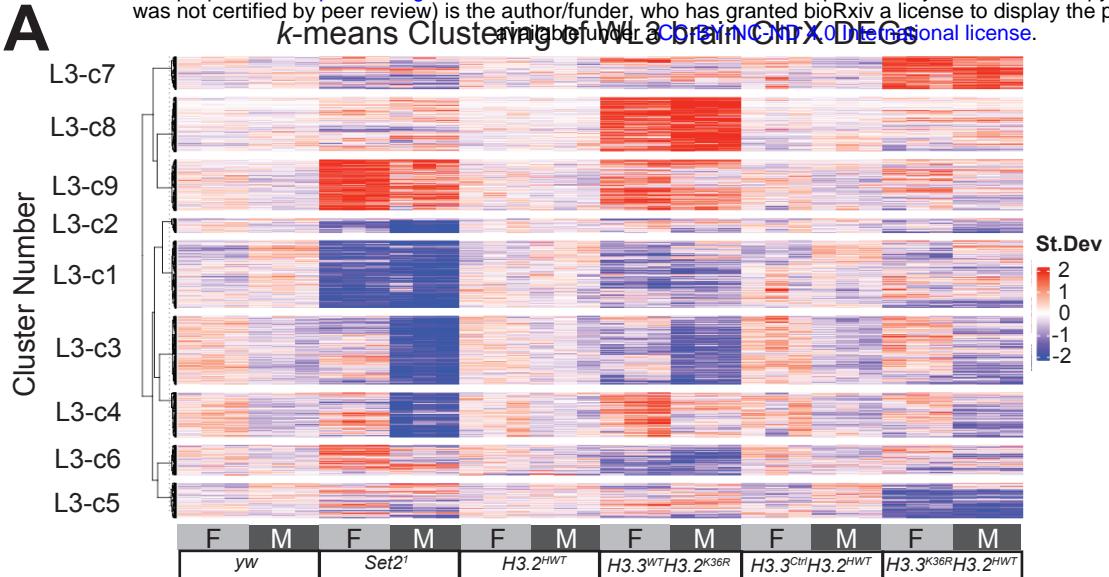
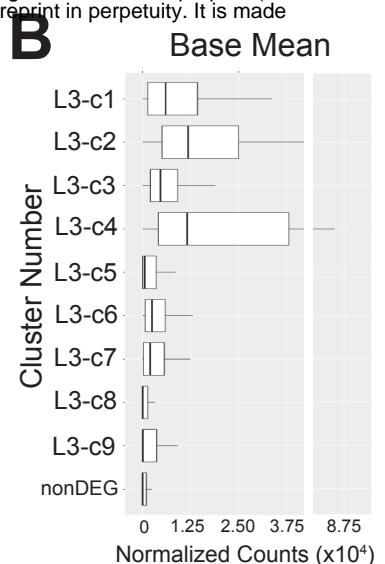


Figure 4

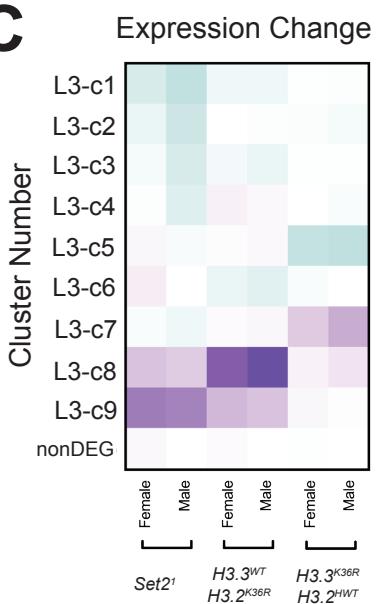
A



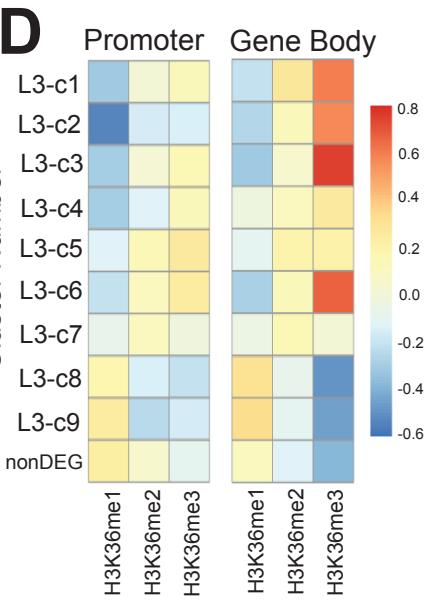
B



C



D



E

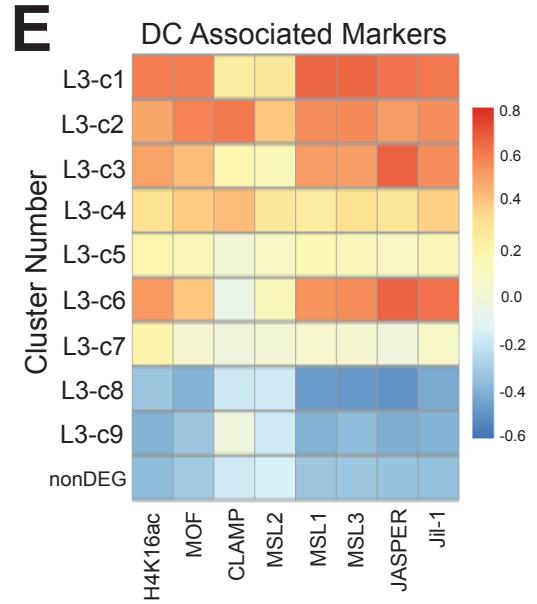


Figure 5

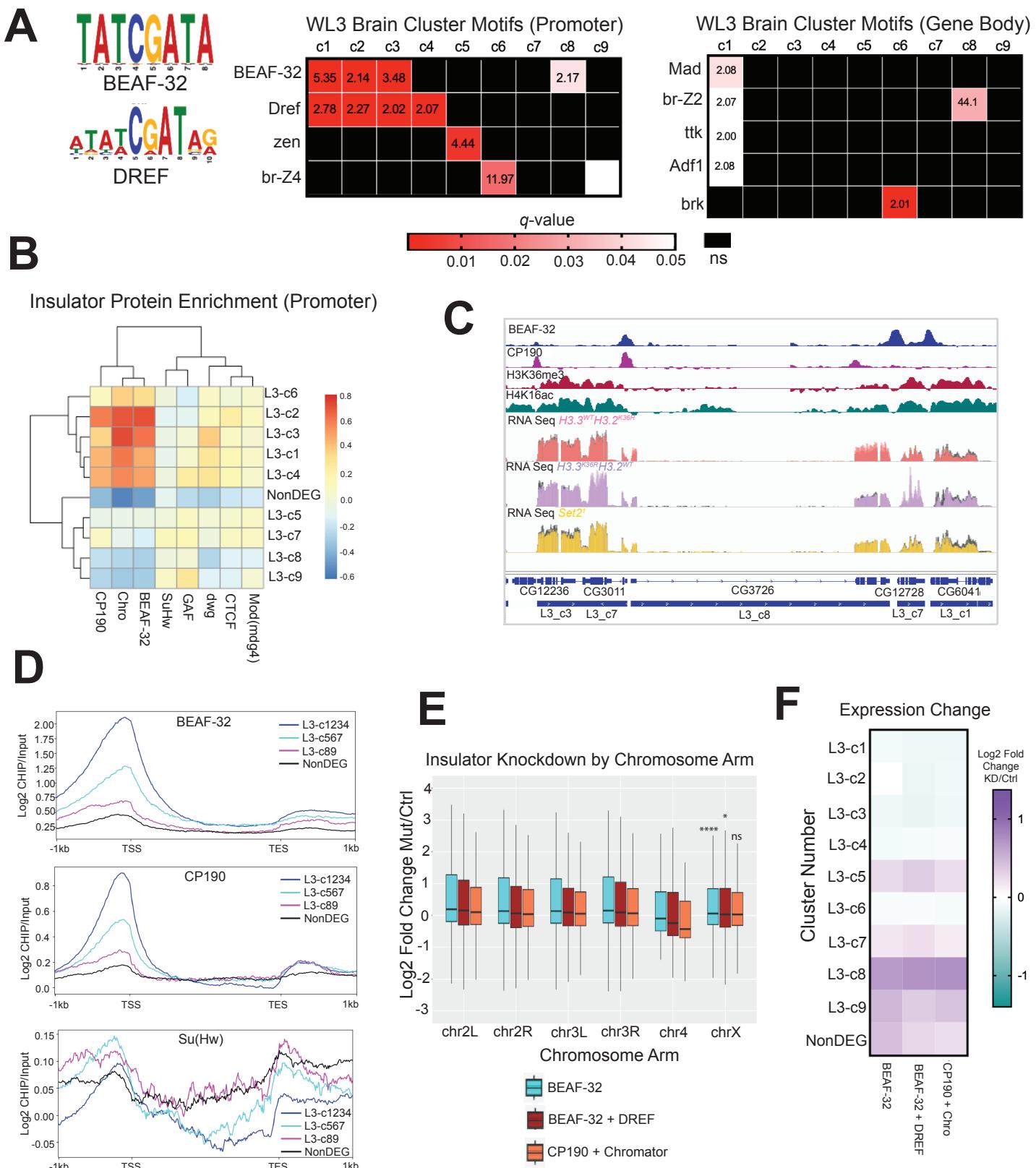


Figure 6

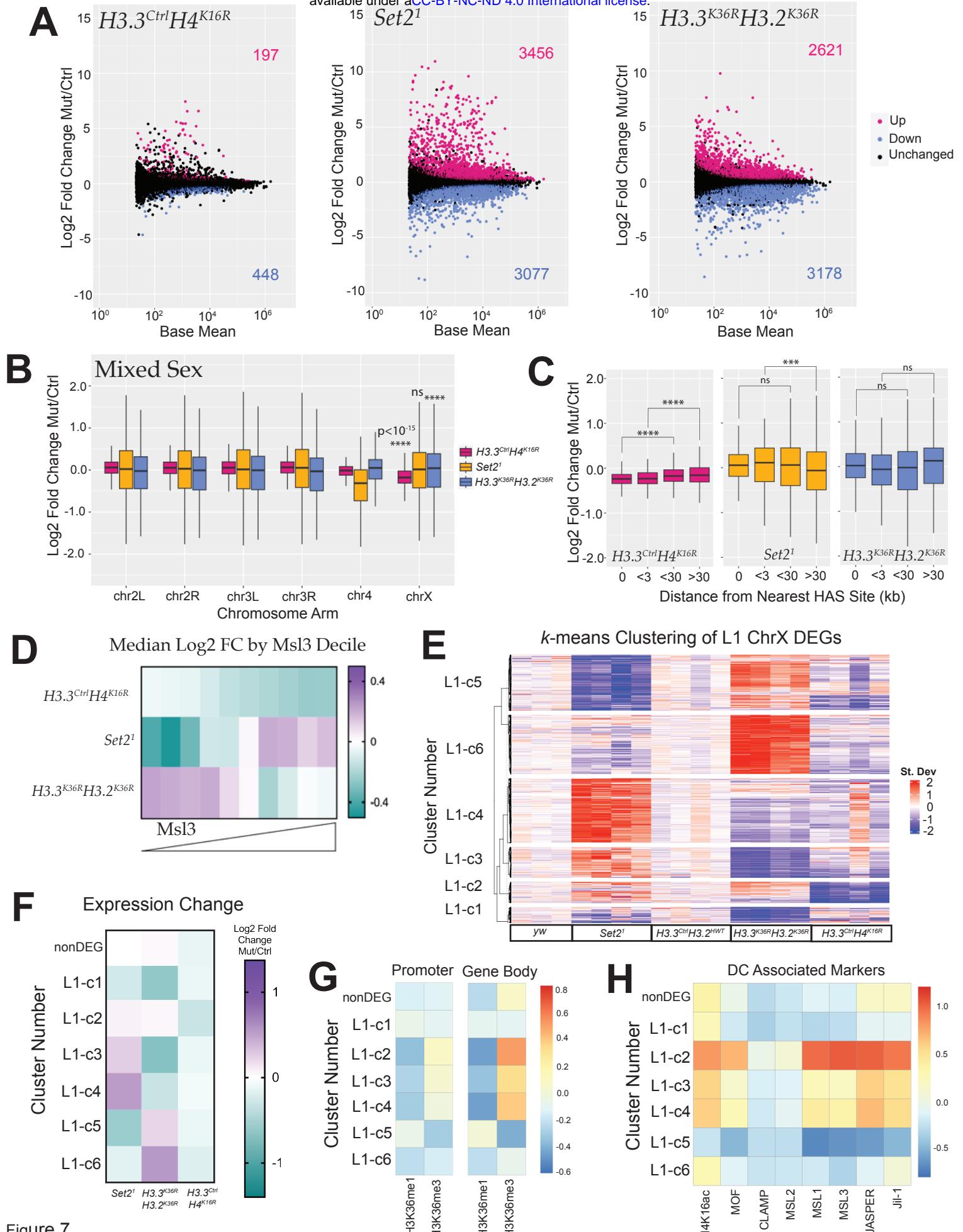


Figure 7