Drosophila SUMM4 complex couples insulator function and DNA replication timing control

Evgeniya N. Andreyeva[1]†, Alexander V. Emelyanov[1]†, Markus Nevil[2], Lu Sun[3], Elena Vershilova[1], Christina A. Hill[4], Michael-C. Keogh[3], Robert J. Duronio[4-7], Arthur I. Skoultchi[1], Dmitry V. Fyodorov[1]*

[1]Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[2]UNC-SPIRE, University of North Carolina, Durham, NC 275999, USA

[3]Epicypher, Inc., Durham, NC 27709, USA

[4]Integrative Program for Biological and Genome Sciences, University of North Carolina, Chapel Hill, NC, 27599 USA

[5]Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, 27599 USA

[6]Department of Biology, University of North Carolina, Chapel Hill, NC, 27599 USA

[7]Department of Genetics, University of North Carolina, Chapel Hill, NC, 27599 USA

*Corresponding author: dmitry.fyodorov@einsteinmed.org

†These authors contributed equally to this work

Competing interest statement

Lu Sun and Michael-C Keogh are employed by Epicypher, Inc., a commercial developer and supplier of the EpiDyne® nucleosomes and associated remodeling assay platforms used in this study. The remaining authors declare no competing interests.

Abstract Asynchronous replication of chromosome domains during S phase is essential for eukaryotic genome function, but the mechanisms establishing which domains replicate early versus late in different cell types remain incompletely understood. *Drosophila* SNF2-related factor SUUR imparts under-replication of late-replicating intercalary heterochromatin in polytene chromosomes. SUUR negatively regulates DNA replication fork progression; however, its mechanism of action remains obscure. Here we developed a novel method termed MS-Enabled Rapid protein Complex Identification (MERCI) to isolate a stable stoichiometric native complex SUMM4 that comprises SUUR and a chromatin boundary protein Mod(Mdg4)-67.2. Mod(Mdg4) stimulates SUUR ATPase activity and is required for a normal spatiotemporal distribution of SUUR *in vivo*. SUUR and Mod(Mdg4)-67.2 together mediate the activities of *gypsy* insulator that prevent certain enhancer-promoter interactions and establish euchromatin-heterochromatin barriers in the genome. Furthermore, *SuUR* or *mod(mdg4)* mutations reverse under-replication of intercalary heterochromatin. Thus, SUMM4 can impart late replication of intercalary heterochromatin by attenuating the progression of replication forks through euchromatin/heterochromatin boundaries. Our findings reveal that DNA replication can be delayed by a chromatin barrier and uncover a critical role for architectural proteins in replication control. They suggest a mechanism for replication timing that does not depend on an asynchronous firing of replication origins.

41 ## Introduction

42 Replication of metazoan genomes occurs according to a highly coordinated spatiotemporal
43 program, where discrete chromosomal regions replicate at distinct times during S phase (Rhind &
44 Gilbert, 2013). The replication program follows the spatial organization of the genome in Megabase-
45 long constant timing regions interspersed by timing transition regions (Marchal, Sima, & Gilbert,
46 2019). The spatiotemporal replication program exhibits correlations with genetic activity, epigenetic
47 marks and features of 3D genome architecture and sub-nuclear localization. Yet the reasons for these
48 correlations remain obscure. Interestingly, the timing of firing for any individual origin of replication
49 is established during G1 before pre-replicative complexes (pre-RC) are assembled at origins
50 (Dimitrova & Gilbert, 1999), suggesting a mechanism that involves factors other than the core
51 replication machinery.

52

53 Most larval tissues of *Drosophila melanogaster* grow via G-S endoreplication cycles that duplicate
54 DNA without cell division resulting in polyploidy (Zielke, Edgar, & DePamphilis, 2013).
55 Endoreplicated DNA molecules frequently align in register to form giant polytene chromosomes
56 (Zhimulev et al., 2004). Importantly, in some cell types, genomic domains corresponding to the latest
57 replicated regions of dividing cells, specifically pericentric (PH) and intercalary (IH) heterochromatin,
58 fail to fully replicate during each endocycle resulting in underreplication (UR). In both dividing and
59 endoreplicating cells, these regions are depleted of sites for binding the Origin of Replication Complex
60 (ORC) and thus their replication primarily relies on forks progressing from external origins (Sher et
61 al., 2012). Although cell cycle programs are dissimilar between endoreplicating and mitotically
62 dividing cells, they share biochemically identical DNA replication machinery (Zielke et al., 2013).
63 Thus, UR provides a facile readout for late replication initiation and delayed fork progression.

64

65 The *Suppressor of Under-Replication* (*SuUR*) gene is essential for polytene chromosome UR in IH
66 and PH (Belyaeva et al., 1998). In *SuUR* mutants, the DNA copy number in UR regions is equivalent
67 to fully polyploidized regions of the genome. *SuUR* encodes a protein (SUUR) containing a helicase
68 domain with homology to that of the SNF2/SWI2 family. The occupancy of ORC in IH and PH is not
69 increased in *SuUR* mutants (Sher et al., 2012), and thus the increased replication of UR regions is
70 likely not due to the firing of additional origins. Rather, SUUR negatively regulates the rate of
71 replication fork progression (Nordman et al., 2014) by an unknown mechanism. It has been proposed
72 (Posukh, Maksimov, Skvortsova, Koryakov, & Belyakin, 2015) that retardation of the replisome by
73 SUUR takes place via simultaneous physical association with the components of the fork (*e.g.*, CDC45
74 and PCNA) (Kolesnikova et al., 2013; Nordman et al., 2014) and repressive chromatin proteins, such
75 as HP1a (Pindyurin et al., 2008).
76
77 Using a newly developed proteomics approach, we discovered that SUUR forms a stable complex
78 with a chromatin boundary protein Mod(Mdg4)-67.2. We demonstrate that SUUR and Mod(Mdg4)-
79 67.2 together are required for maximal UR of IH and full activity of the *gypsy* insulator, thereby
80 implicating insulators in obstructing replisome progression and the control of late DNA replication.

81 ## Results

82 **Identification of SUMM4, the native form of SUUR in *Drosophila* embryos**

83 To determine how SUUR functions in replication control we sought to identify its native complex.
84 Previous attempts to characterize the native form of SUUR by co-IP or tag-affinity purification gave
85 rise to multiple putative binding partners (Kolesnikova et al., 2013; Munden et al., 2018; Nordman et
86 al., 2014; Pindyurin et al., 2008). However, evaluating whether any of these proteins are present in a
87 native SUUR complex is problematic because of the low abundance of SUUR, which also precludes
88 its purification by conventional chromatography. Therefore, we developed a novel biochemical
89 approach using embryonic extracts (which can be obtained in large quantities) that relies on partial
90 purification by multi-step FPLC and shotgun proteomics of chromatographic fractions by quantitative
91 LCMS (***Figure 1A***). We term this technology MERCI for MS-Enabled Rapid protein Complex
92 Identification (*Materials and Methods*).

93

94 The depth of proteomic quantification is limited by the range of peptides identified in the
95 information-dependent acquisition (IDA), dubbed "ion library" (IL). Unfortunately, SUUR-specific
96 peptides could not be found in ILs obtained from acquisitions of crude nuclear extracts or fractions
97 from the first, phosphocellulose, step (IL1, ***Figure 1—figure supplement 1A, Data S1***). Thus, to
98 quantify SUUR in phosphocellulose fractions, we augmented IL1 with the IL obtained by acquisition
99 of recombinant SUUR (ILR, ***Figure 1A&B***). In ILs from subsequent chromatographic steps, peptides
100 derived from native SUUR were detected (***Figure 1—figure supplement 1A, Data S1***) and used for
101 quantification of cognate data-independent acquisitions (DIA/SWATH) (***Figure 1C***).

102

103 The final aspect of the MERCI algorithm calls for re-quantification of FPLC fraction SWATH
104 acquisitions with an IL from the last step (IL5) that is enriched for peptides derived from SUUR and

105 co-purifying polypeptides (***Figure 1A***) and includes only 140 proteins (***Figure 1—figure supplement***
106 ***1A, Data S1***). In this fashion, scarce polypeptides (including SUUR and, potentially, SUUR-binding
107 partners) that may not be detectable in earlier steps will not evade quantification. Purification profiles
108 of proteins quantified in all five FPLC steps (132) were then artificially stitched into 83-point arrays of
109 Z-scores (***Figure 1D, Data S2***). These profiles were Pearson-correlated with that of SUUR and ranked
110 down from the highest Pearson coefficient, PCC (***Figure 1E***). Whereas the PCC numbers for the
111 bottom 130 proteins lay on a smooth curve, the top two proteins, SUUR (PCC = 1.000) and
112 Mod(Mdg4) (PCC = 0.939) fell above the extrapolated (by polynomial regression) curve (***Figure 1F***).
113 Consistently, SUUR and Mod(Mdg4) exhibited nearly identical purification profiles in all five FPLC
114 steps (***Figure 1D***), unlike the next two top-scoring proteins, EGG (PCC = 0.881) and CG6700 (PCC =
115 0.874) (***Figure 1—figure supplement 1B&C***). Also, HP1a (PCC = 0.503), which had been proposed
116 to form a complex with SUUR (Pindyurin et al., 2008) did not co-purify with SUUR in any FPLC
117 steps (***Figure 1—figure supplement 1D***).
118
119 Mod(Mdg4) is a BTB/POZ domain protein that functions as an adaptor for architectural proteins
120 that promote various aspects of genome organization (Georgiev & Gerasimova, 1989; Gerasimova,
121 Gdula, Gerasimov, Simonova, & Corces, 1995). It is expressed as 26 distinct polypeptides generated
122 by splicing *in trans* of a common 5'-end precursor RNA with 26 unique 3'-end precursors (Buchner et
123 al., 2000). IL5 contained seven peptides derived from Mod(Mdg4) (99% confidence). Whereas four of
124 them mapped to the common N-terminal 402 residues, three were specific to the C-terminus of a
125 particular form, Mod(Mdg4)-67.2 (***Figure 1—figure supplement 1E-G***). Peptides specific to other
126 splice forms were not detected. We raised an antibody to the C-terminus of Mod(Mdg4)-67.2,
127 designated ModT antibody, and analyzed size exclusion column fractions by immunoblotting.
128 Consistent with SWATH analyses (***Figure 1C&D***) SUUR and Mod(Mdg4)-67.2 polypeptides

129 copurified as a complex with an apparent molecular mass of ~250 kDa (***Figure 1G***). Finally, we
130 confirmed that SUUR specifically co-immunoprecipitated with Mod(Mdg4)-67.2 from embryonic
131 nuclear extracts (***Figure 1H***). As a control, XNP co-immunoprecipitated with HP1a as shown
132 previously (Emelyanov, Konev, Vershilova, & Fyodorov, 2010), but did not with SUUR or
133 Mod(Mdg4) (***Figure 1H***). We conclude that SUUR and Mod(Mdg4) form a stable stoichiometric
134 complex that we term SUMM4.
135
136 **Biochemical activities of recombinant SUMM4 *in vitro***
137 We reconstituted recombinant SUMM4 complex by co-expressing FLAG-SUUR with
138 Mod(Mdg4)-67.2-His_6 in Sf9 cells and purified it by FLAG affinity chromatography (***Figure 2A***).
139 Mod(Mdg4)-67.2 is the predominant form of Mod(Mdg4) expressed in embryos (*e.g.*, ***Figure 1H***).
140 Thus, minor Mod(Mdg4) forms may have failed to be identified by IDA in IL5 (***Figure 1—figure***
141 ***supplement 1E***). We discovered that FLAG-SUUR did not co-purify with another splice form,
142 Mod(Mdg4)-59.1 (***Figure 1—figure supplement 1G, Figure 2A***). Whereas the identity of an ~100-
143 kDa Mod(Mdg4)-67.2-His_6 band copurifying with FLAG-SUUR was confirmed by mass-spec
144 sequencing, the FLAG-purified material from Sf9 cells expressing FLAG-SUUR and Mod(Mdg4)-
145 59.1 did not contain Mod(Mdg4)-specific peptides. Therefore, the shared N-terminus of Mod(Mdg4)
146 (1-402) is not sufficient for interactions with SUUR. However, this result does not exclude a
147 possibility that SUUR may form complex(es) with some of the other, low-abundance 24 splice forms
148 of Mod(Mdg4). The SUUR- Mod(Mdg4)-67.2 interaction is specific, as the second-best candidate
149 from our correlation analyses (*Drosophila* SetDB1 ortholog EGG; ***Figure 1F***) did not form a complex
150 with FLAG-SUUR (***Figure 2—figure supplement 1A***), although it associated with its known partner
151 WDE, an ortholog of hATF7IP/mAM (Wang et al., 2003).
152

153 The N-terminus of SUUR contains a region homologous with SNF2-like DEAD/H helicase
154 domains. Although SUUR requires its N-terminal domain to function *in vivo* (Munden et al., 2018), it
155 has been hypothesized to be inactive as an ATPase (Nordman & Orr-Weaver, 2015). We analyzed the
156 ability of recombinant SUUR and SUMM4 (***Figure 2A***) to hydrolyze ATP *in vitro* in comparison to
157 recombinant *Drosophila* ISWI (***Figure 2—figure supplement 1B***). Purified recombinant Mod(Mdg4)-
158 67.2 and a variant SUUR protein with a point mutation in the putative Walker A motif (K59A) were
159 used as negative controls (***Figure 2A, Figure 2—figure supplement 1B***). Contrary to the prediction,
160 both SUUR and SUMM4 exhibited strong ATPase activities (***Figure 2B***). SUMM4 was 1.4- to 2-fold
161 more active than SUUR alone, indicating that Mod(Mdg4)-67.2 stimulates SUUR enzymatic activity.
162 We then examined whether DNA and nucleosomes can stimulate the activity of SUUR. To this end,
163 we reconstituted oligonucleosomes on plasmid DNA (***Figure 2—figure supplement 1B-E***). Linker
164 histone H1-containing chromatin was also used as a substrate/cofactor, because SUUR has been
165 demonstrated to physically interact with H1 (Andreyeva et al., 2017). In contrast to ISWI, SUUR was
166 not stimulated by addition of DNA or nucleosomes and moderately (by about 70%) activated by H1-
167 containing oligonucleosomes (***Figure 2C***) consistent with its reported direct physical interaction with
168 H1 (Andreyeva et al., 2017).
169
170 We examined the nucleosome remodeling activities of SUUR and SUMM4; specifically, their
171 ability to expose a positioned DNA motif in the EpiDyne®-PicoGreen™ assay (*Materials and Methods*
172 and ***Figure 2—figure supplement 1F***). Centrally or terminally positioned mononucleosomes were
173 efficiently mobilized by ISWI and human BRG1 in a concentration- and time-dependent manner
174 (***Figure 2—figure supplement 1G-J***). In contrast, SUUR and SUMM4 did not reposition either
175 nucleosome (***Figure 2D***). Thus, SUUR and SUMM4 do not possess a detectable remodeling activity
176 and may resemble certain other SNF2-like enzymes (*e.g.*, RAD54) that utilize the energy of ATP

177 hydrolysis to mediate alternate DNA translocation reactions (Jaskelioff, Van Komen, Krebs, Sung, &
178 Peterson, 2003).
179
180 **The distribution of SUMM4 complex *in vivo***
181 We examined the positions of SUUR and Mod(Mdg4)-67.2 within polytene chromosomes by
182 indirect immunofluorescence (IF) and discovered that they overlap at numerous locations (***Figure 3A,***
183 ***Figure 3—figure supplement 1A&B***). In late endo-S phase when SUUR exhibited a characteristic
184 distribution, it co-localized with Mod(Mdg4)-67.2 at numerous (hundreds of) loci along the
185 chromosome arms (***Figure 3—figure supplement 1B***). Mod(Mdg4)-67.2 was present at classical
186 regions of SUUR enrichment, such as UR domains in 75C and 89E (***Figure 3—figure supplement***
187 ***1A***). The chromocenter, which consists of under-replicated PH, contains SUUR but did not show
188 occupancy by Mod(Mdg4)-67.2 (***Figure 3—figure supplement 1B***). Conversely, there were multiple
189 sites of Mod(Mdg4)-67.2 localization that were free of SUUR (***Figure 3—figure supplement 1B***).
190 Individual pixel intensities of IF signals for SUUR and Mod(Mdg4)-67.2 were plotted as a 2D scatter
191 plot (***Figure 3—figure supplement 1C***) and were found to exhibit a weak positive correlation
192 (R^2=0.278). Consistent with the possible multi-phasic relative distribution of SUUR and Mod(Mdg4)-
193 67.2 (***Figure 3—figure supplement 1B***), the 2D plot encompassed four distinct areas, where SUUR
194 and Mod(Mdg4)-67.2 were co-localized, enriched separately or absent (***Figure 3—figure supplement***
195 ***1D***). When regions of SUUR-alone and Mod(mdg4)-67.2-alone enrichment were excluded, and only
196 the regions of their apparent colocalization were considered, the anti-SUUR and anti-ModT signals
197 exhibited a strong positive correlation (R^2=0.568, ***Figure 3—figure supplement 1D***).
198
199 The existence of chromosome loci heavily enriched for Mod(Mdg4)-67.2 but devoid of SUUR
200 suggests that there are additional native form(s) of Mod(Mdg4)-67.2, either as an individual

201 polypeptide or in complex(es) other than SUMM4. When we fractionated *Drosophila* nuclear extract
202 using a different progression of FPLC steps (***Figure 3—figure supplement 1E***), we found that
203 Mod(Mdg4)-67.2 can form a megadalton-sized complex that did not contain SUUR (***Figure 3—figure***
204 ***supplement 1F-H***). Therefore, a more intricate pattern of Mod(Mdg4)-67.2 distribution likely reflects
205 loading of both SUMM4 and an alternative Mod(Mdg4)-67.2-containing complex.
206
207 We tested whether SUUR and Mod(Mdg4) loading into polytene chromosomes were mutually
208 dependent using mutant alleles of *SuUR* and *mod(mdg4)*. $SuUR^{ES}$ is a null allele of *SuUR* (Makunin et
209 al., 2002). $mod(mdg4)^{m9}$ is a null allele with a deficiency that removes gene regions of the shared 5'-
210 end precursor and eight specific 3'-precursors (Savitsky, Kim, Kravchuk, & Schwartz, 2016).
211 $mod(mdg4)^{u1}$ contains an insertion of a *Stalker* element in the last coding exon of Mod(Mdg4)-67.2 3'-
212 precursor (Gerasimova et al., 1995), and thus is predicted only to disrupt expression of this isoform.
213 $SuUR^{ES}$ and $mod(mdg4)^{u1}$ are homozygous viable, and $mod(mdg4)^{m9}$ is recessive adult pharate lethal.
214 We could not detect Mod(Mdg4)-67.2 expression in homozygous $mod(mdg4)^{m9}$ L3 salivary glands by
215 immunoblotting, whereas $mod(mdg4)^{u1}$ expressed a truncated polypeptide (*cf*, ~70 kDa and ~100 kDa,
216 ***Figure 3—figure supplement 1I***). The truncated 70-kDa polypeptide failed to load into polytene
217 chromosomes (***Figure 3B***, ***Figure 3—figure supplement 1J***). As shown previously, SUUR could not
218 be detected in $SuUR^{ES}$ chromosomes. Since homozygous $mod(mdg4)^{m9}$ L3 larvae were produced by
219 *inter se* crosses of heterozygous parents, the very low amounts of Mod(Mdg4)-67.2 in $mod(mdg4)^{m9}$
220 polytene chromosomes (barely above the detection limit) were presumably maternally contributed.
221
222 The absence (or drastic decrease) of Mod(Mdg4)-67.2 also strongly reduced the loading of SUUR
223 (***Figure 3B***, ***Figure 3—figure supplement 1J***). The normal distribution pattern of SUUR in polytene
224 chromosomes is highly dynamic (Andreyeva et al., 2017; Kolesnikova et al., 2013). SUUR is initially

225 loaded in chromosomes at the onset of endo-S phase and then re-distributes through very late endo-S,
226 when it accumulates in UR domains and PH. In both *mod(mdg4)* mutants, we observed a striking
227 absence of SUUR in polytene chromosomes during early endo-S (***Figure 3B***), which indicates that the
228 initial deposition of SUUR is dependent on its interactions with Mod(Mdg4). Although SUUR
229 deposition slightly recovered by late endo-S, it was still several fold weaker than that in wild type.
230 Potentially, in the absence of Mod(Mdg4), SUUR may be tethered to IH and PH loci by direct binding
231 with linker histone H1 as shown previously (Andreyeva et al., 2017). Finally, the gross subcellular
232 distribution of SUUR also strongly correlated with that of Mod(Mdg4): a mis-localization of truncated
233 Mod(Mdg4)-67.2 from nuclear to partially cytoplasmic was accompanied by a similar mis-localization
234 of SUUR (***Figure 3C***). This result indicates that the truncation of Mod(Mdg4) in *mod(mdg4)*^{u1} may
235 have an antimorphic effect by mis-localization and deficient chromatin binding of interacting
236 polypeptides, including SUUR (***Figure 3C***) and others (***Figure 3—figure supplement 1E-H***).
237
238 **The role of SUMM4 as an effector of the insulator/chromatin barrier function**
239 Mod(Mdg4)-67.2 does not directly bind DNA but instead is tethered by a physical association with
240 zinc finger factor Suppressor of Hairy Wing, Su(Hw). Su(Hw) directly binds to consensus sequences
241 that are present in *gypsy* transposable elements and are also widely distributed across the *Drosophila*
242 genome in thousands of copies (Adryan et al., 2007). Mod(Mdg4)-67.2 was previously shown to be
243 essential for the insulator activity of *gypsy* (Gerasimova et al., 1995), which functions *in vivo* to
244 prevent enhancer-promoter interactions and establish a barrier to the propagation of chromatin forms
245 (Cai & Levine, 1995; Roseman, Pirrotta, & Geyer, 1993). We therefore tested whether SUMM4
246 contributes to *gypsy* insulator functions.
247
248 The *ct*^{6} allele of *Drosophila* contains a *gypsy* element inserted between the wing enhancer and
249 promoter of the gene *cut* that inactivates *cut* expression and results in abnormal wing development

250 (***Figure 4A***). We discovered that both *mod(mdg4)*^{u1} and *SuUR*^{ES} mutations partially suppressed this
251 phenotype (***Figure 4A***) and significantly increased the wing size compared to *ct*^{6} allele alone (***Figure***
252 ***4B***). Thus, both subunits of SUMM4 are required to mediate the full enhancer-blocking activity of
253 *gypsy*. Another insulator assay makes use of a collection of *P{SUPor-P}* insertions that contain the
254 *white* reporter flanked by 12 copies of *gypsy* Su(Hw)-binding sites. When *P{SUPor-P}* is inserted in
255 heterochromatin, *white* is protected from silencing resulting in red eyes (Roseman et al., 1995). Both
256 *mod(mdg4)*^{u1} and *SuUR*^{ES} relieved the chromatin barrier function of Su(Hw) sites, causing repression
257 of *white* (***Figure 4C***). We conclude that SUMM4 is an insulator complex that contributes to the
258 enhancer-blocking and chromatin boundary functions of *gypsy* by a mechanism schematized in ***Figure***
259 ***4D&E***.
260
261 **The role of SUMM4 in regulation of DNA replication in polytene chromosomes**
262 A similar, chromatin partitioning-related mechanism may direct the function of SUUR in the
263 establishment of UR in late-replicating IH domains of polytene chromosomes (***Figure 4F***). It has been
264 long known that 3D chromosome partitioning maps show an "uncanny alignment" with replication
265 timing maps (Rhind & Gilbert, 2013). To examine the possible roles of SUMM4 in UR, we measured
266 DNA copy number genome-wide in salivary glands of L3 larvae by next generation sequencing
267 (NGS). In *w*^{1118} control salivary glands, the DNA copy profile revealed large (>100-kbp) domains of
268 reduced ploidy (***Figure 4—figure supplement 1A***), similar to a previous report (Andreyeva et al.,
269 2017). Excluding pericentric and sub-telomeric heterochromatin, we called 70 UR regions (***Table 1***) in
270 euchromatic arms, as described in *Materials and Methods*.
271
272 In both *SuUR* and *mod(mdg4)*^{m9} null larvae, we observed statistically significant suppression of
273 UR in IH (***Figure 4G, Figure 4—figure supplement 1B, Table 1***). Consistent with its lack of
274 accumulation within the chromocenter of polytene chromosomes (***Figure 3A***), Mod(Mdg4) was

275 dispensable for UR in PH. The NGS data strongly correlated with qPCR measurements of DNA copy
276 numbers (***Figure 4H, Figure 4—figure supplement 1C***). Furthermore, cytological evidence in the
277 75C region supported the molecular analyses in that both mutants exhibited a brighter DAPI staining
278 of the 75C1-2 band than that in w^{1118}, indicative of higher DNA content (***Figure 4—figure***
279 ***supplement 1C***).
280
281 Uniformly, *SuUR* mutation gave rise to a stronger relief of UR than that produced by the
282 *mod(mdg4)*^{m9} null allele (***Table 1***). This result can be explained by embryonic deposition of functional
283 Mod(Mdg4) proteins and RNA by heterozygous mothers, unlike the complete absence of SUUR
284 throughout the life cycle of the homozygous viable and fertile *SuUR*^{ES} animals. Although third instar
285 larvae are >1,000-fold larger, volume-wise, than the embryos, persistent Mod(Mdg4)-67.2 can still be
286 detected in polytene chromosomes of these larvae by IF despite its dilution and degradation (***Figure***
287 ***3C, Figure 3—figure supplement 1J***). In contrast, unlike L3, first instar larvae are nearly identical in
288 size to the embryos. Therefore, since the endoreplication cycles initiate in embryos and L1, in
289 *mod(mdg4)*^{m9} animals the first few out of 10-11 rounds of chromosome polytenization take place with
290 an almost normal amount of Mod(Mdg4) present, which may substantially limit the effect of
291 *mod(mdg4)*^{m9} mutation on UR as measured in L3.
292
293 Seemingly, there is a contradiction between a strong effect that *mod(mdg4)* null mutation has on
294 the loading of SUUR in polytene chromosomes (***Figure 3B***) and a weaker effect on UR (***Figure***
295 ***4G&H*** and ***Figure 4—figure supplement 1B-D***). However, the SUUR occupancy is examined in L3
296 after maternal *mod(mdg4)* product is nearly eliminated (***Figure 3B***). On the other hand, the DNA copy
297 number, although also measured in L3 (***Figure 4G&H*** and ***Figure 4—figure supplement 1B-D***), is a
298 product of multiple rounds of endoreplication that initiate before Mod(Mdg4) is exhausted. To validate

299 the putative effect of maternally contributed SUMM4 on the establishment of UR, we performed
300 qPCR measurements of DNA copy numbers in salivary glands of homozygous *SuUR* animals
301 produced by *inter se* crosses of heterozygous $SuUR^{ES}$/+ parents (***Figure 4H***, ***Figure 4—figure***
302 ***supplement 1C***, zygotic $SuUR^{ES}$). Similar to the maternal Mod(Mdg4), the initial maternal
303 contribution of SUUR partially limited the reversal of UR in cytological regions 4D and 75C. Thus,
304 when the *SuUR* and *mod(mdg4)* null mutant animals are similarly derived from heterozygous mothers
305 that deposit wild-type gene product into their progeny, the mutant UR phenotypes in the third instar
306 larval salivary gland are essentially indistinguishable.
307
308 We conclude that SUUR and Mod(Mdg4)-67.2 act together as subunits of stable SUMM4 complex,
309 which is required for the establishment of UR in the IH domains of *Drosophila* polytene chromosome.

310 ## Discussion

311 **MERCI is a powerful new approach to characterize stable stoichiometric protein complexes**

312 We present here a facile method, termed MERCI, to rapidly identify subunits of stable native
313 complexes by only partial chromatographic purification. It allows one to circumvent the conventional,
314 rate-limiting approach to purify proteins to apparent homogeneity. Since a multi-step FPLC scheme
315 invariably leads to an exponential loss of material, reducing the number of purification steps in the
316 MERCI protocol allows identification of rare complexes, such as SUMM4, which may be present in
317 trace amounts in native sources. On the other hand, MERCI obviates introduction of false-positives
318 frequently associated with tag purification of ectopically expressed targets that render results less
319 reliable. Notably, MERCI is not limited to analyses of known polypeptides, since it is readily amenable
320 to fractionation of native factors based on a correlation with their biochemical activities *in vitro*.

321

322 The dissection of protein interactome by extract fractionation on orthogonal FPLC columns and MS-
323 based approaches has been previously attempted (Havugimana et al., 2012; Shatsky et al., 2016).
324 However, unlike the newly developed MERCI approach, these studies were aimed at comprehensive,
325 proteome-wide analyses, which managed to only yield data for the most abundant complexes. The major
326 distinction of the MERCI protocol is that it is targeted towards a particular protein (SUUR in this study).
327 The crucial final stage of the MERCI algorithm is re-quantification of all acquired SWATH data using a
328 library acquired from fractions of the last column (IL5, ***Figure 1A&D***). The target protein and co-
329 purifying polypeptides are substantially enriched after several chromatographic steps and thus, yield a
330 greater number of detected peptides, which helps a more precise quantification. Although SWATH
331 allows reliable measurement of picogram amounts of proteins (***Figure 1—figure supplement 2A&B***),
332 the range of quantified polypeptides is always limited by those present in IDA acquisitions (ion
333 libraries). For low-abundance proteins, such as SUUR and Mod(Mdg4), specific peptides are not

334 detectable by IDA in earlier chromatographic steps (***Data S1***). Consequently, SWATH quantification
335 using only the cognate ion libraries would not discern the near perfect co-fractionation of SUUR and
336 Mod(Mdg4) in all five steps (***Figure 1D***), precluding identification of the SUUR-Mod(Mdg4) complex
337 (***Figure 1D-F***).
338
339 One limitation of the MERCI protocol is its failure to measure the absolute amounts of identified
340 polypeptides. For instance, quantification of SWATH data (***Figure 1C***) measures the relative (to
341 reference proteins and each other) amounts of SUUR across fractions. To measure the absolute levels of
342 SUUR, a semi-quantitative approach was used by building a titration curve from SWATH acquisitions
343 of known amounts of recombinant SUUR (***Figure 1—figure supplement 2A&B***). We estimated the
344 amount of SUUR in the nuclear extract (~140 pg in 25 µg total protein, ***Figure 1—figure supplement***
345 ***2B***) and in individual fractions from all chromatographic steps (***Figure 1—figure supplement 2C***).
346 Although in five FPLC steps we achieved >3,000-fold purification of SUUR, it remained only ~2% pure
347 (***Figure 1—figure supplement 2D***). A progressive loss of material precludes further purification (300 ng
348 of SUUR in 16 µg total protein). Thus, the SUMM4 complex would be nearly impossible to purify to
349 homogeneity from a substantial amount of starting material (~1 kg *Drosophila* embryos, ~2.5 g protein),
350 suggesting that SUMM4 could not be identified by the classical FPLC approach.
351
352 **SUMM4 regulates the function of *gypsy* insulator elements**
353 Both subunits of SUMM4 contribute to the known functions of *gypsy* insulator (***Figure 4A-C***).
354 Although a *SuUR* mutation decreased the insulator activity, the suppression was universally weaker than
355 that by *mod(mdg4)*^{u1}. It is possible that SUUR is not absolutely required for the establishment of the
356 insulator. For instance, the loss of SUMM4 may be compensated by the alternative complex of
357 Mod(Mdg4)-67.2 (***Figure 3—figure supplement 1E-H***). Furthermore, the *mod(mdg4)*^{u1} allele is

358 expected to have an antimorphic function, since it can mis-localize interacting partner proteins,
359 including SUUR itself (***Figure 3B&C***). Interestingly, *SuUR* has been previously characterized as a weak
360 suppressor of variegation of the *white*^{m4h} X chromosome inversion allele, which places the *white* gene
361 near PH (Belyaeva et al., 2003). In contrast, *SuUR* mutation enhances variegation in the context of
362 insulated, PH-positioned *white* (***Figure 3C***). Therefore, this phenotype is unrelated to the putative
363 *Su(var)* function of *SuUR* but, rather, is insulator-dependent.
364

365 **ATP-dependent motor proteins are required for the establishment of chromatin barrier and**
366 **chromosome partitioning**

367 Our discovery and analyses of SUMM4 provide a biochemical link between ATP-dependent motor
368 factors and the activity of insulators in regulation of gene expression and chromatin partitioning.
369 Insulator elements organize the genome into chromatin loops (Gerasimova et al., 1995) that are involved
370 in the formation of topologically associating domains, TADs (Peterson, Samuelson, & Hanlon, 2021;
371 Rowley et al., 2017; Szabo, Bantignies, & Cavalli, 2019). In mammals, CTCF-dependent loop formation
372 requires ATP-driven motor activity of SMC complex cohesin (Davidson et al., 2019). In contrast, CTCF
373 and cohesin are thought to be dispensable for chromatin 3D partitioning in *Drosophila* (Matthews &
374 White, 2019). Instead, the larger, transcriptionally inactive domains (canonical TADs) are interspersed
375 with smaller active compartmental domains, which themselves represent TAD boundaries (Rowley et
376 al., 2017). It has been proposed that in *Drosophila*, domain organization does not rely on architectural
377 proteins but is established by transcription-dependent, A-A compartmental (gene-to-gene) interactions
378 (Rowley et al., 2017). However, *Drosophila* TAD boundaries are enriched for architectural proteins
379 other than CTCF (Van Bortle et al., 2014), and their roles have not been tested in loss-of-function
380 models. Thus, it is possible that in *Drosophila*, instead of CTCF, the 3D partitioning of the genome is
381 facilitated by another group of insulator proteins, such as Su(Hw) and SUMM4 that together associate
382 with class 3 insulators (Schwartz et al., 2012).

383

384 Moreover, SUUR may provide the DNA motor function to promote a physical separation of active
385 and inactive loci and help establish chromosome contact domains (***Figure 4D-F***). The translocation
386 model is supported by observations of an asymmetric, selective propagation of SUUR from its initial
387 sites of deposition via Su(Hw)-Mod(Mdg4) binding towards inside of IH regions but not outside (***Figure***
388 ***4—figure supplement 1E***) (Filion et al., 2010), which may be facilitated by physical interactions
389 between SUUR and linker histone H1 enriched in IH (Andreyeva et al., 2017). It has been reported that
390 another *Drosophila* BTB/POZ domain insulator protein CP190 forms a complex with a DEAD-box
391 helicase Rm62 that contributes to the insulator activity (Lei & Corces, 2006). Thus, ATP-dependent
392 motor proteins may represent an obligatory component of the insulator complex machinery.

393

394 **SUMM4 mediates known biological functions of SUUR**

395 Our discovery explains previous observations about biological functions of SUUR. For instance, the
396 initial deposition of SUUR and its co-localization with PCNA has been proposed to depend on direct
397 physical interaction with components of the replisome (Kolesnikova et al., 2013). Our model indicates
398 that instead, the apparent co-localization of SUUR with PCNA throughout endo-S phase (***Figure***
399 ***3—figure supplement 1J***) may be caused by a replication fork retardation at insulator sites. SUUR is
400 deposited in chromosomes as a subunit of SUMM4 complex at thousands of loci by tethering via
401 Mod(Mdg4)-Su(Hw) interactions. As replication forks progress through the genome, they encounter
402 insulator complexes where replication machinery pauses for various periods of time before resolving the
403 obstacle. Thus, the increased co-residence time of PCNA and SUUR manifests cytologically as their
404 partial co-localization. With the progression of endo-S phase, some of the SUMM4 insulator complexes
405 are evicted and thus, the number of SUUR-positive loci is decreased, until eventually, the replication
406 fork encounters nearly completely impenetrable insulators demarcating the UR domain boundaries.

407

408 This mechanism is especially plausible given that boundaries of IH loci very frequently encompass
409 multiple, densely clustered Su(Hw) binding sites (*e.g.*, ***Figure 4G, Figure 4—figure supplement 1C***).
410 We examined the data from genome-wide proteomic analyses for Su(Hw) and SUUR performed by
411 DamID in Kc167 cells (Filion et al., 2010). Strikingly, Su(Hw) DamID-measured occupancy does not
412 exhibit a discrete pattern expected of a DNA-binding factor. Instead, it appears broadly dispersed,
413 together with SUUR, up to tens of kbp away from mapped Su(Hw) binding sites (***Figure 4—figure***
414 ***supplement 1E***). Interestingly, when Hidden Markov modeling was applied to the DamID data, Su(Hw),
415 Mod(Mdg4)-67.2 and SUUR occupancies were found to strongly correlate genome-wide in a novel
416 chromatin form ("malachite") that frequently demarcates the boundaries of IH (Khoroshko et al., 2016).
417 These observations strongly corroborate the translocation model for the mechanism of action of
418 SUMM4. According to this model, upon tethering to DNA-bound Su(Hw), SUMM4 traverses the UR
419 region, which helps to separate it in a contact domain. As DNA within the UR region is tracked by
420 SUUR (***Figure 4F***), it is brought into a transient close proximity with both SUMM4 and the associated
421 Su(Hw) protein, which is detected by DamID (or ChIP) as an expanded occupancy pattern.
422
423 The deceleration of SUUR-bound replication forks was also invoked as an explanation for the
424 apparent role of SUUR in the establishment of epigenetic marking of IH (Posukh et al., 2015). We
425 propose that global epigenetic modifications observed in the *SuUR* mutant likely do not directly arise
426 from derepression of the replisome as suggested but, rather, result from the coordinate insulator-
427 dependent regulatory functions of SUUR in both the establishment of a chromatin barrier and DNA
428 replication control (*cf* ***Figure 4E&F***).
429
430 **Architectural proteins can attenuate replication forks and regulate replication timing**
431 Our work demonstrates for the first time that insulator complexes assembled on chromatin can
432 attenuate the extent of replication in discrete regions of the salivary gland polylploid genome. Despite

433 distinct cell cycle programs in dividing and endoreplicating cells, the biochemical composition of
434 replisomes in both cell types is identical (Zielke et al., 2013). Therefore, similar insulator-driven
435 control mechanisms for DNA replication are likely conserved in mitotically dividing diploid cells. Our
436 data thus implicates insulator/chromatin boundary elements as a critical component of DNA
437 replication control. Our model suggests that delayed replication of repressed chromatin (*e.g.*, IH)
438 during very late S phase can be imposed in a simple, two-stroke mechanism (***Figure 4F***). First, it
439 requires that an extended genomic domain is completely devoid of functional origins of replication.
440 The assembly and licensing of proximal pre-RC complexes can be repressed epigenetically or at the
441 level of DNA sequence. And second, this domain is separated from flanking chromatin by a barrier
442 element associated with an insulator complex, such as SUMM4. This structural organization is capable
443 of preventing or delaying the entry of external forks fired from distal origins.
444
445 An important frequent feature of the partially suppressed UR in *mod(mdg4)* animals is its asymmetry
446 (***Figure 4—figure supplement 1C-D***), which is consistent with a unidirectional penetration of the UR
447 domain by a replication fork that fires from the nearest external origin (***Figure 4F***). The SUMM4-
448 dependent barrier may be created as a direct physical obstacle to MCM2-7 DNA-unwinding helicase or
449 other enzymatic activities of the replisome. Alternatively, SUMM4 may inhibit the replication
450 machinery indirectly by assembling at the insulator a DNA/chromatin structure that is incompatible with
451 replisome translocation. This putative inhibitory structure may involve epigenetic modifications of
452 chromatin as proposed earlier (Gaszner & Felsenfeld, 2006), linker histone H1 as shown previously
453 (Andreyeva et al., 2017) and may also be dependent on Rif1, a negative DNA replication regulator that
454 acts downstream of SUUR (Munden et al., 2018).
455
456 In conclusion, we used a newly developed MERCI approach to identify a stable stoichiometric
457 complex termed SUMM4 that comprises SUUR, a previously known negative effector of replication,

458 and Mod(Mdg4), an insulator protein. SUMM4 subunits cooperate to mediate transcriptional
459 repression and chromatin boundary functions of *gypsy*-like (class 3) insulators (Schwartz et al., 2012)
460 and inhibit DNA replication likely by slowing down replication fork progression through the boundary
461 element. Thus, SUMM4 is required for coordinate regulation of gene expression, chromatin
462 partitioning and DNA replication timing. The insulator-dependent regulation of DNA replication
463 offers a novel mechanism for the establishment of replication timing in addition to the currently
464 accepted paradigm of variable timing of replication origin firing.

465 # Materials and Methods

466 ## Recombinant proteins

467 Recombinant proteins were expressed in Sf9 cells using baculovirus system (SUUR, Mod(Mdg4),
468 EGG and WDE), in *E. coli* (ISWI, ModT antigen and LCMS reference proteins) or obtained from
469 EpiCypher Inc. (human BRG1/SMARCA4).

470

471 *Sf9 cells*

472 All baculovirus constructs were cloned by PCR with Q5 DNA polymerase (New England Biolabs)
473 and Gibson assembly with NEBuilder HiFi DNA Assembly Cloning kit (New England Biolabs) into
474 pFastBac vector (Thermo Fisher) under control of polyhedrin promoter. All constructs were validated by
475 Sanger sequencing. Baculoviruses were generated according to the protocol by Thermo Fisher. The
476 baculoviruses were isolated by plaque purification, amplified three times, and their titers were measured
477 by plaque assay. FLAG-SUUR construct was cloned from *SuUR-RA* cDNA (LD13959, DGRC). The
478 following open reading frame (ORF) was expressed: M**DYKDDDDK**H-SUUR-PA(1..962)-
479 VEACGTKLVEKY*. To generate ATPase-dead mutant, SUUR-PA(K59) codon was replaced with an
480 alanine codon by PCR and Gibson cloning. Mod(Mdg4)-67.2-V5-His_6 and Mod(Mdg4)-59.1-V5-His_6
481 constructs were cloned from cDNAs *mod(mdg4)-RT* and *mod(mdg4)-RI* synthesized as gBlocks by IDT,
482 Inc. The following ORFs were expressed: Mod(Mdg4)-67.2(1..610)-
483 GIL**EGKPIPNPLLGLDST**GASVEHHHHHH* and Mod(Mdg4)-59.1(1..541)-
484 GIL**EGKPIPNPLLGLDST**GASVEHHHHHH*. EGG-FLAG and EGG (untagged) were cloned by PCR
485 from *egg-RA* cDNA (IP14531). The following ORF was expressed: EGG-PA(1..1262)-**DYKDDDDK***
486 and EGG-PA(1..1262)-*. FLAG-WDE was cloned by PCR from *wde-RA* cDNA (LD26050). The
487 following ORF was expressed: M**DYKDDDDK**-WDE-PA(2..1420)-*. The sequences of FLAG and V5
488 tags are highlighted in bold typeface.

489
490 Cells, $2 \cdot 10^6$/ml in Sf-900 II SFM medium (Gibco), were infected at multiplicity of infection (MOI)
491 of ~10 in PETG shaker flasks (Celltreat, Inc.). After infection for 48-72 hours at 27°C, cells were
492 harvested, and recombinant proteins were purified by FLAG or Ni-NTA affinity chromatography
493 (Fyodorov & Kadonaga, 2003). Whereas, typically, amplified baculovirus stocks had titers above $5 \cdot 10^9$
494 pfu/ml, FLAG-SUUR viruses reached no more than $2\text{-}4 \cdot 10^8$ pfu/ml, presumably, due to the inhibitory
495 effect of over-expressed protein on viral DNA replication. Accordingly, whereas typical yields of
496 purified recombinant proteins were >100 µg from 1 L Sf9 cell culture, SUUR polypeptides were
497 produced at no more than 2 µg from 1 L culture, which also adversely affected the protein purity
498 (***Figure 1B&2A*** and ***Figure 2—figure supplement 1A&B***).
499
500 *E. coli*
501 The expression construct for untagged recombinant *Drosophila* ISWI was prepared from a full-
502 length ISWI cDNA (Ito et al., 1999). Human TXNRD1 sequence was cloned from a cDNA provided by
503 Addgene (#38863), and TXNRD2 was synthesized as a gBlock gene fragment by IDT, Inc. The ORFs
504 were inserted by Gibson cloning in a pET backbone vector in frame with a C-terminal intein-CBD
505 (chitin-binding domain) tag. Protein expression was induced by IPTG in Rosetta 2 cells, and proteins
506 were purified in non-denaturing conditions by chitin affinity chromatography and intein self-cleavage as
507 described (Emelyanov et al., 2014), followed by anion exchange chromatography (Source 15Q) on
508 FPLC. Detailed cloning and purification methods are available upon request. Note that the cloned
509 human thioredoxin reductase ORFs do not express the C-terminal selenocysteines. They were thus
510 presumed catalytically inactive (Arner, Sarioglu, Lottspeich, Holmgren, & Bock, 1999; Cheng & Arner,
511 2017) and designated hTXNRD1ci and hTXNRD2ci. They were used exclusively as spike-in mass
512 standards in LCMS acquisitions of *Drosophila* proteins.
513

514 Polypeptide corresponding to the C-terminal specific region of Mod(Mdg4)-67.2 was cloned in
515 pET24b vector in frame with a C-terminal His_6 tag. `M`-Mod(Mdg4)-67.2(403..610)-`GILEHHHHHH*` was
516 expressed in Rosetta 2 and purified by Ni-NTA affinity chromatography in non-denaturing conditions.
517 The polypeptide (ModT) was dialyzed into PBS (137 mM NaCl, 3 mM KCl, 8 mM NaH_2PO_4, 2 mM
518 KH_2PO_4) and used as an antigen for immunizations (see below). All recombinant proteins were
519 examined by SDS-PAGE along with Pierce BSA mass standards (Thermo Fisher), and their
520 concentrations were calculated from infrared scanning of Coomassie-stained gels (Odyssey Fc Imaging
521 System, LI-COR Biosciences).
522
523 **Crude cell extracts**
524 *Nuclear extract from* Drosophila *embryos*
525 ~1 kg or ~200 g wild-type (Oregon R) *Drosophila* embryos were collected 0-12 h after egg
526 deposition (AED) from population cages. The embryos were dechorionated, and nuclear extracts were
527 prepared as described (Kamakaka, Tyree, & Kadonaga, 1991). Protein concentration was measured by
528 Pierce BCA assay (Thermo Fisher). The extracts were fractionated by FPLC (***Figure 1B*** and ***Figure***
529 ***3—figure supplement 1E***) on AKTA PURE system (Cytiva Life Sciences). Aliquots of
530 chromatographic fractions were examined by quantitative shotgun proteomics or western blot analyses
531 as described below. Peak SUUR or Mod(Mdg4) fractions were diluted to appropriate ionic strength (if
532 applicable) and used as a starting material for the next chromatographic step. Details on FPLC column
533 sizes and run parameters are available upon request.
534
535 E. coli *lysate*
536 40-ml Rosetta 2 overnight culture was harvested by centrifugation, resuspended in 20 ml HEG (25
537 mM HEPES, pH 7.6, 0.1 mM EDTA, 10% glycerol) supplemented with 0.1 M KCl, 1 mM DTT and 2
538 mM $CaCl_2$. Cells were disrupted by sonication and centrifuged to remove insoluble material. Nucleic

539 acids were digested with 15 units micrococcal nuclease (Sigma Aldrich) for 20 min at 37°C, and the
540 proteins were precipitated with 2 M ammonium sulfate. The pellet was resuspended in 10 ml HEG + 0.1
541 M KCl + 1 mM DTT with protease inhibitors (0.5 mM benzamidine, 0.2 mM PMSF) and dialyzed
542 against the same buffer. After centrifugation, the concentration of soluble protein was measured by BCA
543 assay, the *E. coli* lysate was diluted to 1 mg/ml using 100 mM ammonium bicarbonate (ABC) and stored
544 at -80°C.
545
546 **Mass-spectroscopy samples**
547 *Column fractions*
548 For each chromatographic step, 14 to 20 fractions were selected based on the protein fractionation
549 profile according to the UV (A_{280}) absorbances measurements. 50-100 µl aliquots of chromatographic
550 fractions, starting material (SM) and column flow-through (FT, if applicable) were saved, and protein
551 concentrations were estimated based on their UV absorbances (1,000 mU A_{280} was considered to be
552 equivalent to 5 mg/ml total protein). Equal volumes of each fraction, SM and FT were used for MS
553 acquisitions, so that no more than 40 µg total protein was processed in each reaction. As a reference, the
554 reactions were supplemented with 1.5 µg each of purified recombinant human thioredoxin reductases 1
555 and 2 (hTXNRD1ci and hTXNRD2ci, catalytically inactive) expressed in *E. coli*. Dithiotreitol (DTT)
556 was added to the protein samples to 10 mM and NP-40 – to 0.02%. Reaction volumes were brought to
557 85 µl with 50 mM ammonium bicarbonate (ABC). All reagents, including water, were HPLC/MS grade.
558 The proteins were reduced for 1 h at 37°C and then alkylated with 30 mM iodoacetamide (IAA, 15 µl
559 200 mM IAA in water) for 45 min at room temperature in the dark. Alkylated proteins were desalted
560 into 50 mM ABC using ZebaSpin columns (40 kDa MWCO) and digested with 1 µg trypsin for 2 h at
561 37°C. 1 µg more trypsin was added, and the digestion progressed at 37°C overnight. Tryptic peptides
562 were lyophilized for 2 h on SpeedVac with heat and resuspended in 100 µl Sample Buffer: 1%

563 acetonitrile (ACN) and 0.1% formic acid (FA) in water. Equal volumes (23 µl) of samples were used for
564 IDA and SWATH acquisitions (in triplicate) as described below.
565
566 *Recombinant SUUR*
567 To generate the recombinant SUUR reference spectral library (ILR), ~0.5 µg purified recombinant
568 FLAG-SUUR (both 130 and 65 kDa bands, ***Figure 1A***) was mixed with 1.5 µg each of hTXNRD1ci and
569 hTXNRD2ci and processed for an IDA acquisition as described above, except for 0.5 µg trypsin was
570 used in each cleavage step, and the peptide sample was resuspended in 30 µl Sample Buffer. For
571 SWATH titration of SUUR (***Figure 1—figure supplement 1H&I***), 1 µg recombinant FLAG-SUUR was
572 mixed with 25 µg *E. coli* lysate protein and 1.5 µg each of hTXNRD1ci and hTXNRD2ci. 10-fold serial
573 dilutions down to 10 fg SUUR were also prepared using the mixture of *E. coli* lysate with reference
574 proteins. The samples were processed for SWATH acquisitions in triplicate as described above, 30 µl of
575 sample per injection.
576
577 *In-gel digestion of recombinant proteins for LCMS identification*
578 Recombinant SUUR or SUMM4 purified by FLAG immunoaffinity chromatography was resolved
579 on SDS-PAGE, stained with Coommassie Blue (***Figure 1A&2A***), and up to eight most prominent
580 protein bands were excised. The gel slices were transferred to 1.5-ml Eppendorf tubes, gently crushed
581 with a RotoDounce pestle and destained with 25 mM ABC in 50% methanol and then with 25 mM ABC
582 in 50% ACN (30 min each at room temperature). The proteins were reduced in 50 µl 10 mM DTT for 1
583 h at 55°C and alkylated with 30 mM IAA for 45 min at room temperature in the dark. The gel fragments
584 were washed with 25 mM ABC in 50% ACN, dehydrated with 100% ACN, dried in a SpeedVac,
585 rehydrated by addition of 50 µl 50 mM ABC and digested with 0.25 µg trypsin overnight at 37°C. The
586 peptides were extracted once with 50 µl 10% FA and once with 100 µl 3% FA in 60% ACN, both
587 extracts were combined, dried in a SpeedVac and resuspended in 50 µl Sample Buffer. 40 µl of each
588 sample was injected for IDA acquisitions as described below.

589

590 **Mass-spectroscopy acquisition methods**

591 LC-MS/MS analyses were performed on a TripleTOF 5600+ mass spectrometer (AB SCIEX) coupled
592 with M5 MicroLC system (AB SCIEX/Eksigent) and PAL3 autosampler.

593

594 *Instrument settings*

595 LC separation was performed in a trap-elute configuration, which consists of a trapping column
596 (LUNA C18(2), 100 Å, 5 µm, 20 × 0.3 mm cartridge, Phenomenex) and an analytical column (Kinetex
597 2.6 µm XB-C18, 100 Å, 50 × 0.3 mm microflow column, Phenomenex). The mobile phase consisted of
598 water with 0.1% FA (phase A) and 100% ACN containing 0.1% FA (phase B). 200 ng to 10 µg total
599 protein was injected for each acquisition. Peptides in Sample Buffer were injected into a 50-µl sample
600 loop, trapped and cleaned on the trapping column with 3% mobile phase B at a flow rate of 25 µl/min
601 for 4 min before being separated on the analytical column with a gradient elution at a flow rate of 5
602 µl/min. The gradient was set as follows: 0 to 48 min: 3% to 35% phase B, 48 to 54 min: 35% to 80%
603 phase B, 54 to 59 min: 80% phase B, 59 to 60 min: 80% to 3% phase B, and 60 to 65 min at 3% phase
604 B. An equal volume of each sample (23 µl) was injected four times, once for information-dependent
605 acquisition (IDA), immediately followed by data-independent acquisition (DIA/SWATH) in triplicate.
606 Acquisitions of distinct samples were separated by a blank injection to prevent sample carry-over. The
607 mass spectrometer was operated in positive ion mode with EIS voltage at 5,200 V, Source Gas 1 at 30
608 psi, Source Gas 2 at 20 psi, Curtain Gas at 25 psi and source temperature at 200°C.

609

610 *Information-dependent acquisitions (IDA) and data analyses*

611 IDA was performed to generate reference spectral libraries for SWATH data quantification. The
612 IDA method was set up with a 250-ms TOF-MS scan from 400 to 1250 Da, followed by MS/MS scans
613 in a high sensitivity mode from 100 to 1500 Da of the top 30 precursor ions above 100 cps threshold

614 (100 ms accumulation time, 100 ppm mass tolerance, rolling collision energy and dynamic
615 accumulation) for charge states (*z*) from +2 to +5. IDA files were searched using ProteinPilot (version
616 5.0.2, ABSciex) with a default setting for tryptic digest and IAA alkylation against a protein sequence
617 database. The *Drosophila* proteome FASTA file (21,970 protein entries, UniProt UP000000803,
618 3/21/2020) augmented with sequences for common contaminants as well as hTXNRD1 and hTXNRD2
619 was used as a reference for the search. Up to two missed cleavage sites were allowed. Mass tolerance for
620 precursor and fragment ions was set to 100 ppm. A false discovery rate (FDR) of 5% was used as the
621 cutoff for peptide identification.
622
623 *SWATH acquisitions and data analyses*
624 For SWATH acquisitions (Zhu, Chen, & Subramanian, 2014), one 50-ms TOF-MS scan from 400 to
625 1250 Da was performed, followed by MS/MS scans in a high sensitivity mode from 100 to 1500 Da (15-
626 ms accumulation time, 100 ppm mass tolerance, +2 to +5 *z*, rolling collision energy) with a variable-
627 width SWATH window (Zhang et al., 2015). DIA data were quantified using PeakView (version
628 2.2.0.11391, ABSciex) with SWATH Acquisition MicroApp (version 2.0.1.2133, ABSciex) against
629 selected spectral libraries generated in ProteinPilot. Retention times for individual SWATH acquisitions
630 were calibrated using 20 or more peptides for hTXNRD1ci and hTXNRD2ci. The following software
631 settings were utilized: up to 25 peptides per protein, 6 transitions per peptide, 95% peptide confidence
632 threshold, 5% FDR for peptides, XIC extraction window 20 minutes, XIC width 100 ppm. Protein peak
633 areas were exported as Excel files (***Data S2***) and processed as described below.
634
635 **MERCI**
636 MERCI is a novel approach for rapid identification of native protein complexes. It combines
637 enrichment for a target subunit of a putative complex by consecutive FPLC steps and quantitative
638 shotgun proteomics of chromatographic fractions. Crude nuclear extract from *Drosophila* embryos was

639 fractionated as in ***Figure 1A***. At every step, 40 µg or less total protein from each of 10-20 fractions
640 (equal volumes) was supplemented with a fixed amount (1.5 µg each) of exogenous reference proteins
641 (human thioredoxin reductases), reduced, alkylated and digested with trypsin (see above). MS1 and MS2
642 spectra of tryptic peptides were acquired by IDA, and relative SUUR abundance in fractions was
643 measured by data-independent acquisition (DIA/SWATH) in triplicate. SWATH data were quantified
644 using cognate IDA-derived ion libraries. Protein areas for all quantified proteins were normalized to the
645 sum of those for reference proteins. The relative numbers were averaged across triplicates, with standard
646 deviations calculated. The average numbers for all quantified proteins were further normalized by
647 converting them to Z-scores (see ***Data S2*** for an example of calculations). Peak SUUR fractions (one to
648 five) were then subjected to the next FPLC/MERCI step. After five column steps, the ion library from
649 the ultimate FPLC step (IL5) was used to re-quantify SWATH data from all steps. Z-scores for all
650 purification steps were stitched together, and the large array encompassing all data points for every
651 protein was analyzed by Pearson correlation with SUUR (***Data S2***). The most closely correlated
652 purification profiles served as an indication for protein co-purification, potentially, as subunits of a
653 stable complex.
654
655 **Biochemical assays with recombinant proteins**
656 *Oligonucleosome substrates*
657 Oligonucleosomes were reconstituted *in vitro* as described (Lu et al., 2013) from supercoiled
658 plasmid DNA (3.2 kb, pGIE-0), native core histones and H1 prepared from *Drosophila* embryos
659 (Fyodorov & Levenstein, 2002) by gradient salt dialysis in the presence of 0.2 mg/ml nuclease-free
660 bovine serum albumin (BSA, New England Biolabs). Quality of reconstitution was assessed by MNase
661 and chromatosome stop assays (***Figure 2—figure supplement 1D&E***).
662
663 *ATPase assay*

664 40 nM recombinant proteins were incubated in 25 µl reaction buffer containing 20 mM HEPES, pH
665 7.6, 0.15 M NaCl, 4 mM $MgCl_2$, 1 mM ATP, 0.1 mM EDTA, 0.02% (v/v) NP-40 and 0.1 mg/ml
666 nuclease-free BSA for 60 min at 27°C. Some reactions additionally contained 10 nM pGIE-0 plasmid
667 DNA or equivalent amounts of oligonucleosomes ±H1. ATPase assays were performed using ADP-Glo
668 Max kit (Promega). All reactions were performed in triplicate, the results were normalized to the ADP-
669 ATP titration curve according to the kit manual and converted to enzymatic rates (molecules of ATP
670 hydrolyzed per molecule of enzyme per minute). Averages and standard deviations were calculated.
671 Statistical differences were calculated by Mann-Whitney test.
672
673 *EpiDyne®-PicoGreen™ nucleosome remodeling assay*
674 EpiDyne®-PicoGreen™ is a restriction enzyme accessibility assay modified for increased throughput
675 and sensitivity (***Figure 2—figure supplement 1F***). In brief, a recombinant ATPase over a concentration
676 range (***Figure 2—figure supplement 1G-J***) was mixed with 10 nM EpiDyne biotinylated nucleosome
677 remodeling substrate (EpiCypher), terminally positioned 6-N-66 (219 bp fragment) or centrally
678 positioned 50-N-66 (263 bp) and 1 mM ATP in 20 µL remodeling buffer, 20 mM Tris-HCl, pH 7.5, 50
679 mM KCl, 3 mM $MgCl_2$, 0.01% (v/v) Tween-20, 0.01% (w/v) BSA. The remodeling reactions were
680 incubated at 23˚C in 384-well format. At indicated time points, the reactions were quenched, and
681 nucleosome substrates were immobilized on an equal volume of streptavidin-coated magnetic beads
682 (NEB), pre-washed and resuspended in 2x quench buffer, 20 mM Tris-HCl, pH 7.5, 600 mM KCl,
683 0.01% (v/v) Tween-20 and 0.01% (w/v) BSA. Beads were successively washed by collection on a
684 magnet (three times with wash buffer, 20 mM Tris-HCl, pH 7.5, 300 mM KCl, 0.01% (v/v) Tween-20)
685 and buffer replacement (once with RE buffer, 20 mM Tris-HCl, pH 7.5, 50 mM KCl, 3 mM $MgCl_2$,
686 0.01% (v/v) Tween-20). Beads were resuspended in 20 µl restriction enzyme mix, 50 units/ml Dpn II
687 (NEB) in RE buffer, and incubated at 23˚C for 30 min, collected on a magnet, and supernatants from all
688 wells were transferred to a new plate. They were mixed with an equal volume of Quant-iT™

689 PicoGreen™ dsDNA reagent (ThermoFisher, Component A) and 1 unit/ml thermolabile proteinase K
690 (NEB) in TE and incubated at 23˚C for 1 hr. Fluorescence intensity was detected on an Envision
691 microplate reader with excitation at 480 nm and emission at 531 nm, and data expressed as relative
692 fluorescence units (RFU) through the EnVision Workstation (version 1.13.3009.1409).
693
694 ***Drosophila* population culture, mutant stocks and genetics**
695 Wild-type (Oregon R) flies were maintained in population cages on agar-grape juice and yeast paste
696 plates at 26°C, 60% humidity with 12-h dark-light cycle. Mutant flies were reared, and crosses were
697 performed at 26°C on standard cornmeal/molasses medium with dry yeast added to the surface. $SuUR^{ES}$
698 was a gift of Igor Zhimulev, and $mod(mdg4)^{m9}$ was a gift of Yuri Schwartz. All other alleles were
699 obtained from the Bloomington Stock Center, Indiana. Combinations of alleles were produced either by
700 crosses with appropriate balancers and segregation of markers or by female germline meiotic
701 recombination. Intra-chromosomal recombination events were confirmed by PCR of genomic DNA.
702 Details and PCR primer sequences are available upon request.
703
704 Fly wings were dissected from ~5 days old adult males and transferred to the drop of PBS + 0.1%
705 Triton X-100 (PBST). The wings were soaked in 80% glycerol in PBST and photographed using Zeiss
706 AxioVert 200M microscope with EC Plan-Neofluar 2.5X/0.075 lens in bright field and CCD
707 monochrome camera AxioCam MRm. For wing area measurements, images were processed using
708 Fiji/ImageJ2 software package. Statistical differences were calculated by two-tailed t-test, assuming
709 unequal variances. Adult fly eye images were taken on live, CO_2-anesthetized 2-day-old females on
710 Zeiss stereomicroscope Discovery.V12 using CCD color camera AxioCam MRc.
711
712 **Antibodies, immunoblots and immunoprecipitation (IP)**

713 Polyclonal antibody (anti-ModT) was raised in Guinea pigs by Pocono Rabbit Farm & Lab. Rabbit
714 polyclonal antibody to the C-terminus of *Drosophila* XNP/ATRX (anti-XNP) was described previously
715 (Emelyanov et al., 2010). Rabbit and Guinea pig polyclonal antibodies to *Drosophila* SUUR were a gift
716 of Alexey Pindyurin (Nordman et al., 2014) and Igor Zhimulev (Pindyurin et al., 2008). Rabbit
717 polyclonal Mod(Mdg4)-FL antibody to full-length Mod(Mdg4)-67.2 that recognizes all splice forms of
718 Mod(Mdg4) was a gift of Jordan Rowley and Victor Corces. Mouse monoclonal anti-FLAG (M2, Sigma
719 Aldrich), anti-PCNA (PC10, Cell Signaling), anti-β-tubulin and anti-HP1a (E7 and C1A9,
720 Developmental Studies Hybridoma Bank) were obtained commercially.
721
722 Western blotting was performed using standard techniques. For FPLC column fraction analyses, 5-
723 10 µl of starting material and flow-through (if applicable) and 5-15 µl of column fractions were loaded
724 per lane. For expression analyses in salivary glands, 10 salivary glands from L3 larvae of indicated
725 genotype were frozen and thawed, boiled extensively in 40 µl 2x SDS-PAGE loading buffer,
726 centrifuged, and the material equivalent to four salivary glands was loaded per lane. The following
727 dilutions were used: 1:200,000 anti-ModT, 1:1,000 anti-Mod(Mdg4)-FL, 1:1,000 Guinea pig and rabbit
728 anti-SUUR, 1:1,000 anti-HP1a, 1:1,000 anti-β-tubulin and 1:2,000 anti-FLAG. Infrared-labeled
729 secondary antibodies: donkey anti-Guinea pig IRDye 800CW, goat anti-mouse IRDye 800CW, goat
730 anti-rabbit IRDye 800CW, goat anti-rabbit IRDye 680CW and goat anti-mouse IRDye 680RD – were
731 obtained from Li-COR Biosciences and used at 1:10,000. The blots were scanned on Odyssey Fc
732 Imaging System (LI-COR Biosciences).
733
734 Immunoprecipitation experiments were performed as described (Emelyanov et al., 2012). 400 µl
735 *Drosophila* embryonic nuclear extracts (~10 mg total protein) were incubated with 10 µl Guinea pig
736 anti-ModT, 30 µl rabbit anti-SUUR or 20 µl rabbit anti-XNP antibodies for 3 h at 4°C.
737 Immunocomplexes were collected by addition of 25 µl protein A-agarose plus (Thermo Fisher) for 2 h

738 at 4°C. After washing four times with 1 ml of buffer HEG (25 mM HEPES, pH 7.6, 0.1 mM EDTA,
739 10% glycerol) + 0.15 M NaCl, the immunoprecipitated proteins were eluted with 80 µl 2x SDS-PAGE
740 loading buffer and analyzed by SDS-PAGE and western blot using Guinea pig or rabbit anti-SUUR and
741 anti-Mod(Mdg4) and mouse anti-HP1a antibodies. For Mod(Mdg4) and HP1a, 8 µl of
742 immunoprecipitated material (equivalent to 1 mg nuclear extract proteins) and 5% input (2 µl nuclear
743 extract, 50 µg total protein) were analyzed. For SUUR, 20 µl of immunoprecipitated material
744 (equivalent to 2.5 mg nuclear extract proteins) and 10% input (10 µl nuclear extract, 250 µg total
745 protein) were analyzed.
746
747 **Polytene chromosomes and indirect immunofluorescence (IF) analyses**
748 For all cytological experiments, larvae were reared and collected at 18°C. Polytene chromosomes
749 and whole-mount salivary glands were prepared and analyzed as described previously (Andreyeva et al.,
750 2017). Briefly, salivary glands from wandering third instar larvae were dissected in PBS. Glands were
751 transferred into a formaldehyde-based fixative (one ~15-µl drop of 3% lactic acid, 45% acetic acid,
752 3.7% formaldehyde on a coverslip) for 2 min, squashed, and frozen in liquid N_2. The coverslips were
753 removed, and slides were placed in 70% ethanol for 20 min and stored at −20°C. The slides were
754 washed three times for 5 min in PBST. Primary antibodies were incubated overnight at 4°C in PBST +
755 0.1% BSA and washed three times for 5 min each with PBST. Secondary antibodies were incubated for
756 2 h at room temperature in PBST + 0.1% BSA and washed three times for 5 min each with PBST.
757
758 DNA was stained with 0.1 µg/ml DAPI in PBST for 3 min, and squashes were mounted in Prolong
759 Glass anti-fade mountant (Molecular Probes). Primary and secondary antibodies were used at the
760 following dilutions: Guinea pig anti-ModT, 1:50,000; rabbit anti-SUUR, 1:100; mouse anti-PCNA,
761 1:1,000; mouse anti-FLAG, 1:100; Alexa Fluor 488 highly cross-absorbed (HCA) goat anti-mouse,
762 Alexa Fluor 568 HCA goat anti-Guinea pig and Alexa Fluor 647 plus HCA goat anti-rabbit (all Thermo

763 Fisher), all 1:800. Indirect immunofluorescence (IF) images were obtained with Zeiss AxioVERT 200M
764 microscope and AxioCam MRm mono microscopy camera using a 40x/1.3 Plan-Neofluar or 63x/1.40
765 Plan-Apochromat lenses with oil immersion. Images were acquired using AxioVision software.
766
767 For whole-mount IF staining, L3 larvae were reared at 26°C, and salivary glands were dissected in
768 PBS and fixed in 3.7% formaldehyde (Sigma Aldrich) for 20 min at room temperature. The glands were
769 washed in PBS + 0.3% Triton X-100 and permeabilized for 30 min at 37°C in PBS + 1% Triton X-100.
770 Blocking was performed for 30 min at room temperature in PBS+ 0.3% Triton X-100 supplemented
771 with 10% fetal calf serum and 1% BSA. The glands were incubated with primary antibodies diluted in
772 blocking solution for 48 h at 4°C, washed three times with PBS + 0.3% Triton X-100 for 30 min, and
773 incubated with secondary antibodies in blocking solution overnight at 4°C. The stained glands were
774 washed three times with PBS + 0.3% Triton X-100 for 30 min, stained with DAPI (0.1 µg/ml) for 30
775 min, and mounted in Prolong Gold anti-fade (Invitrogen). IF images were obtained on a Leica SP8
776 confocal microscope using a 20X/0.75 PLAPO lens and processed using Fiji/ImageJ software.
777
778 To quantify the putative colocalization of SUUR and Mod(Mdg4)-67.2 in polytene chromosomes
779 (***Figure 3A***), the image resolution was reduced to 1,388 x 1,040. Pixel intensities (1,443,520) for SUUR
780 and ModT channels were extracted from Bitmap files (ImageJ), normalized to Z-scores and plotted as an
781 X-Y scatter plot (***Figure 3—figure supplement 1C***). For colocalization analyses, the plot regions
782 (Z_{ModT}>1 AND Z_{SUUR}<3, green) and (Z_{ModT}<1 AND Z_{SUUR}>3, red) were excluded from consideration
783 (***Figure 3—figure supplement 1D***).
784
785 **Next generation sequencing analyses (NGS)**
786 Salivary glands from female wandering third-instar larvae were isolated and flash-frozen in liquid N_2
787 until all samples were collected. Genomic DNA for sequencing was prepared from 25 L3 salivary gland
788 pairs or 10 mg embryos (0-6 h AED) using DNeasy Blood and Tissue kit (Qiagen). Each sample was

789 prepared in triplicate. The tissues were soaked in 180 µl buffer ATL + 20 µl proteinase K (15 mg/ml)
790 and lysed for 2-3 h at 55°C. The reactions were cooled to room temperature, supplemented with 4 µl
791 RNase A, ~40 mg/ml (Sigma Aldrich), and RNA was digested for 10 min. The genomic DNA was
792 fragmented with 0.002 units DNase I (Thermo Fisher) in 100-µl reactions containing 10 mM Tris-HCl,
793 pH 7.5, 10 mM $MnCl_2$, 0.1 mM $CaCl_2$, 0.1 mg/ml RNase A and 0.2 mg/ml nuclease-free BSA (1x
794 reaction buffer) for 15 min at 37°C. (DNAse I dilutions were prepared using 1x reaction buffer.)
795 Reactions were stopped by adding 5 µl 0.5 M EDTA, and DNase I was inactivated for 20 min at 65°C.
796 The fragmented DNA was purified on QiaQuick columns using PCR purification kit (Qiagen) and eluted
797 in 40 µl 10 mM Tris-HCl, pH 8.0. The size distribution of DNA fragments (200-600 bp, average ~400
798 bp) was confirmed and DNA concentration was measured on 2100 BioAnalyzer (Agilent). Libraries
799 were prepared from 20 ng of fragmented genomic DNA with the ThruPLEX DNA-seq kit using
800 SMARTer® DNA Unique Dual Indexes (TakaraBio) and sequenced 150-bp paired-end reads on an
801 NovaSeq 6000 (Novagene).
802
803 The sequencing quality of each sample was assessed using FASTQC version 0.11.7 (Andrews,
804 2010). Raw paired-end reads were trimmed of adapters using BBDuk from the BBTools software
805 version 38.71 using the parameters: `ktrim=r ref=adapters rcomp=t tpe=t tbo=t`
806 `hdist=1 mink=11` (Bushnell, 2014). Reads were aligned to the BDGP Release 6 of the *Drosophila*
807 *melanogaster* genome (dm6) (dos Santos et al., 2015) using Bowtie2 version 2.3.4.1 (Langmead &
808 Salzberg, 2012) and parameters `-q --local --very-sensitive-local --no-unal --`
809 `no-mixed --no-discordant --phred33 -I 10 -X 700`. Duplicate reads were marked
810 using Picard 2.2.4 (BroadInstitute) and SAM files were converted to BAM format, filtered for quality (`-`
811 `bq 5`), and removed of duplicates (`-bF 0x400`) using Samtools version 1.9 (Danecek et al., 2021). To
812 examine replicate concordance, a principal component analysis (PCA) was performed using the
813 deepTools package. Replicates clustered indicating high genome-wide similarity within genotypes (*not*

814 *shown*). For visualization, replicates were merged (`samtools merge`) and coverage was calculated
815 across 50-bp bins and normalized to counts per million (CPM) using deeptools version 3.2.0:
816 `bamCoverage -bs 50 –normalizeUsing CPM` (Ramirez et al., 2016). Each genotype was
817 scaled to the diploid Oregon R embryo signal in 5-kb bins: `bigWigCompare —-operation`
818 `first -bs 5000`. DamID-chip data for SUUR and Su(Hw) were retrieved from GSE22069 (Filion
819 et al., 2010). ChIP-chip data for Su(Hw) insulator elements were also used (Negre et al., 2010). UR
820 domains were called using a custom R script to identify regions at least 100 kb in length that fell below
821 the average chromosomal read count as described (Andreyeva et al., 2017). Visualization of all data was
822 performed on the UCSC Genome browser using the dm6 release of the *Drosophila* genome (Kent et al.,
823 2002). Each data set was auto-scaled to its own min and maximum and the data were windowed by
824 mean with 16-pixel smoothing applied.
825

826 **Quantitative real-time PCR**

827 Genomic DNA samples prior to DNase I fragmentation (see above) were diluted to ~0.25 ng/µl.
828 Real-time PCR was performed using 0.5 ng genomic DNA on a ViiA7 thermocycler (Applied
829 Biosystems) with a three-step protocol (95°C 15 sec, 60°C 30 sec, 68°C 60 sec) and iTaq Universal
830 SYBR Green Supermix (Bio-Rad). Primer sequences are provided in ***Table 2***. Each reaction was
831 performed in three technical replicates for each of the three biological samples (N=9). For each
832 amplicon, the average Ct value (`<Ct>`) was calculated and normalized to the average Ct value for a
833 random intergenic genomic sequence as a loading control. Further, for each template, the ΔCt was
834 normalized to the average Ct value for embryonic DNA (diploid control). Standard deviation (σ_{Ct}) for
835 each reaction in triplicate was also calculated. The following ΔΔCt formula was used: $\langle\Delta\Delta Ct\rangle = (\langle Ct_{target}\rangle - \langle Ct_{intergenic86D}\rangle)_{SG} - (\langle Ct_{target}\rangle - \langle Ct_{intergenic86D}\rangle)_{embryo}$.
836 Standard deviations
837 for <ΔΔCt> were calculated as $\sigma_{\Delta\Delta Ct} = \text{square root of } (\sigma^2_{target} + \sigma^2_{intergenic86D})/2$. ΔΔCt's

838 were converted to DNA copy numbers as $2^{-<\Delta\Delta Ct>}$. The confidence interval was calculated in the range

839 between $2^{-<\Delta\Delta Ct>-\sigma}$ and $2^{-<\Delta\Delta Ct>+\sigma}$.

840

841 To examine the putative zygotic function(s) of *SuUR*, heterozygous $SuUR^{ES}$ parents were produced

842 by balancing with *TM6B, Tb* who were crossed *inter se*. L3 salivary glands were dissected from

843 homozygous *SuUR* mutant progeny, and DNA copy numbers were measured by qPCR as described

844 above.

Acknowledgments

This paper is dedicated to the memory of Jonathan R. Warner who participated in initial discussions that have led to the development of MERCI technology. We thank B. Bartholdy, K. Beirit, B. Birshtein, V. Elagin, M. Gamble, T. Kolesnikova, A. Lusser, A. Pindyurin, C. Schildkraut, Y. Schwartz, S. Sidoli and I. Zhimulev for helpful discussions and critical reading of the manuscript. We thank Y. Schwartz, I. Zhimulev and Bloomington Stock Center for fly stocks and V. Corces, A. Pindyurin, J. Rowley and I. Zhimulev for antibodies. We are grateful to A. Aravin and B. Godneeva for cloning EGG and WDE baculovirus constructs. We thank A. Kumar and N. Baker for help with confocal microscopy, and M. Rogers and J. Secombe for the use of Zeiss Discovery.V12. Confocal images were obtained at the Analytical Imaging Facility (Einstein). We thank P. Schultes for help with maintaining the LCMS instrument.

Additional Information

Funding

NIH Office of the Director	NIGMS R01 GM074233	Dmitry V Fyodorov
NIH Office of the Director	NIGMS R01 GM129244	Arthur I Skoultchi
NIH Office of the Director	NIGMS R01 GM124201	Robert J Duronio
NIH Office of the Director	NIGMS R44 GM123869	Michael-C Keogh (EpiCypher)
NIH Office of the Director	NCI T32 CA217824 Training Grant	Markus Nevil
NIH Office of the Director	NIGMS K12 GM000678 Training Grant	Markus Nevil
NIH Office of the Director	NIGMS Equipment Supplement	Dmitry V Fyodorov, Arthur I Skoultchi, and Charles C Query

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Evgeniya N Andreyeva, Investigation, Methodology, Validation, Visualization, Interpretation, Writing – review & editing; Alexander V Emelyanov, Conceptualization, Investigation, Methodology, Validation, Visualization, Interpretation, Writing – review & editing; Markus Nevil, Funding Acquisition, Methodology, Validation, Visualization, Interpretation, Writing – review & editing; Lu Sun, Investigation, Methodology, Validation, Visualization, Interpretation, Writing – review & editing; Elena Vershilova, Investigation, Methodology; Christina A Hill, Investigation, Methodology; Michael-C Keogh, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Interpretation, Writing – review & editing; Robert J Duronio, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Interpretation, Writing – review & editing; Arthur I Skoultchi, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Interpretation, Writing – review & editing; Dmitry V. Fyodorov, Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Interpretation, Writing – original draft, Writing – review & editing

Additional Files

Supplementary Files

Data file S1

Data file S2

Data Availability

893 NGS data has been submitted to Gene Expression Omnibus (GEO, accession number

894 GSE189421). *Reviewers may access the data via:*

895 *https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE189421 using the token upirgumcxngljul*

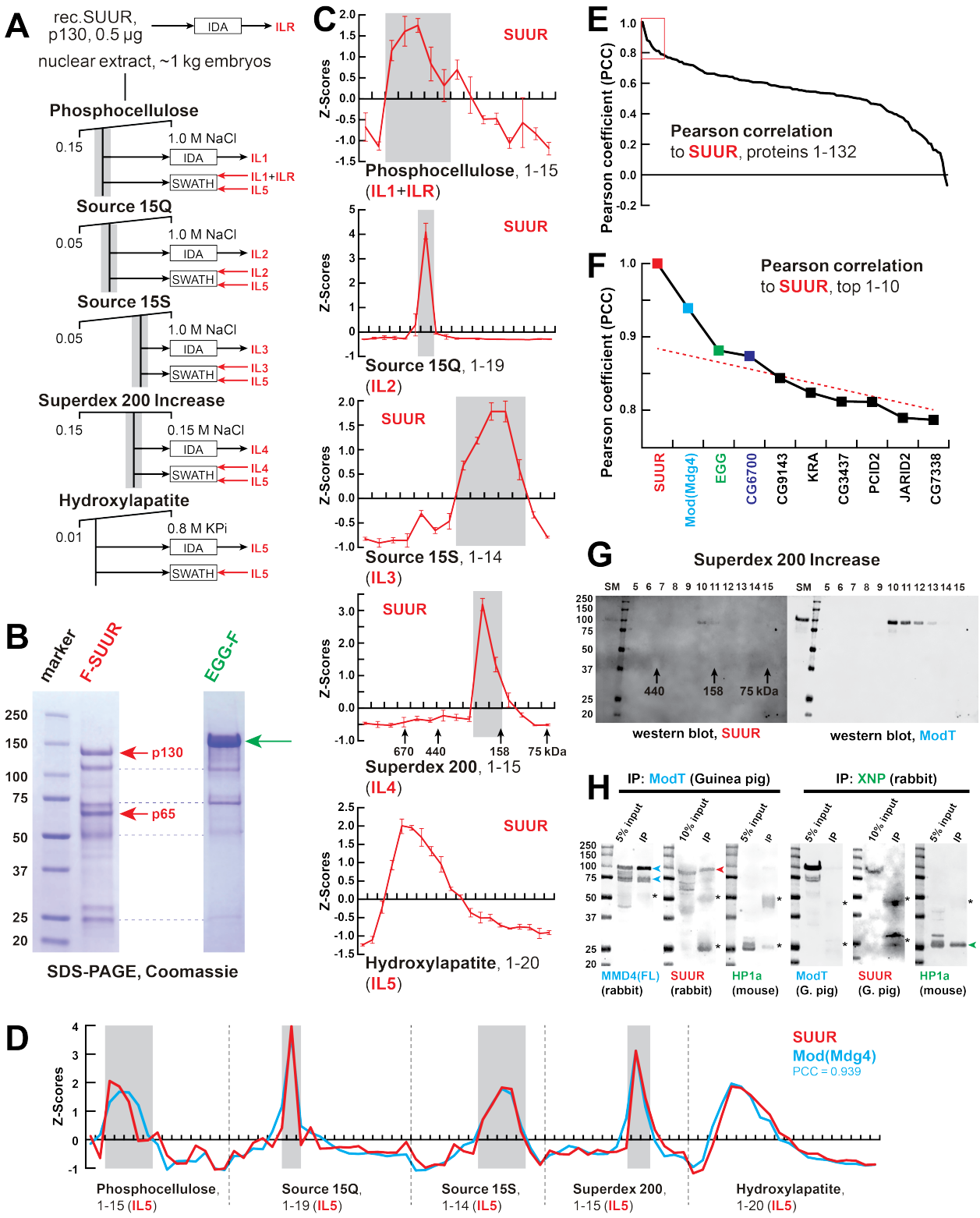


Figure 1. Identification of the SUMM4 complex by MERCI. (A) Schematic of FPLC purification of the native form of SUUR using MERCI approach. ILR, ion library obtained by IDA of recombinant FLAG-SUUR; IL1-5, ion libraries obtained by IDA of FPLC fractions from chromatographic steps 1-5. KPi,

900 potassium phosphate, pH 7.6. (B) Recombinant FLAG-SUUR expressed in Sf9 cells. Identities of eight
901 most prominent bands were determined by mass-spectroscopy. p130 and p65 correspond to full-length
902 and C-terminally truncated FLAG-SUUR, respectively (red arrows). Other bands represent common
903 Sf9-specific contaminants purified by FLAG chromatography (blue dashed lines), *cf* purified EGG-F
904 (green arrow). Molecular mass marker bands are indicated (kDa). (C) SWATH quantitation profiles of
905 SUUR fractionation across individual FPLC steps. Ion libraries (IL) used for SWATH quantitation are
906 shown at the bottom. Z-scores across indicated column fractions are plotted; error bars, standard
907 deviations (N=3). Gray rectangles, fraction ranges used for the next FPLC step; in Superdex 200 step,
908 black arrows, expected peaks of globular proteins with indicated molecular masses in kDa. (D) SWATH
909 quantitation profiles of SUUR (red) and Mod(Mdg4) (cyan) fractionation across five FPLC steps. IL5
910 ion library was used for SWATH quantification. (E) Pearson correlation of fractionation profiles for
911 individual 132 proteins to that of SUUR, sorted from largest to smallest. Red box, the graph portion
912 shown in (F). (F) Top ten candidate proteins with the highest Pearson correlation to SUUR. Red dashed
913 line, trend line extrapolated by polynomial regression (n = 5) from the bottom 130 proteins. (G) Western
914 blot analyses of Superdex 200 fractions with SUUR and ModT antibodies. Molecular mass markers are
915 shown on the left (kDa). (H) Co-IP experiments. SUUR (red arrowhead) co-purifies from nuclear
916 extracts with Mod(Mdg4)-67.2 (cyan arrowheads) but not HP1a (green arrowhead). Anti-XNP co-IPs
917 HP1a but not SUUR of Mod(Mdg4)-67.2. Asterisks, IgG heavy and light chains detected due to
918 antibody cross-reactivity. Mod(Mdg4)-67.2(FL) antibody recognizes all splice forms of Mod(Mdg4).
919
920 **Figure supplement 1.** Identification of the SUMM4 complex by MERCI.
921
922 **Figure supplement 2.** Quantification of SUUR in chromatographic fractions.
923
924 **Figure 1—table 1.** UR domains and UR suppression in SUMM4 subunit mutant alleles.

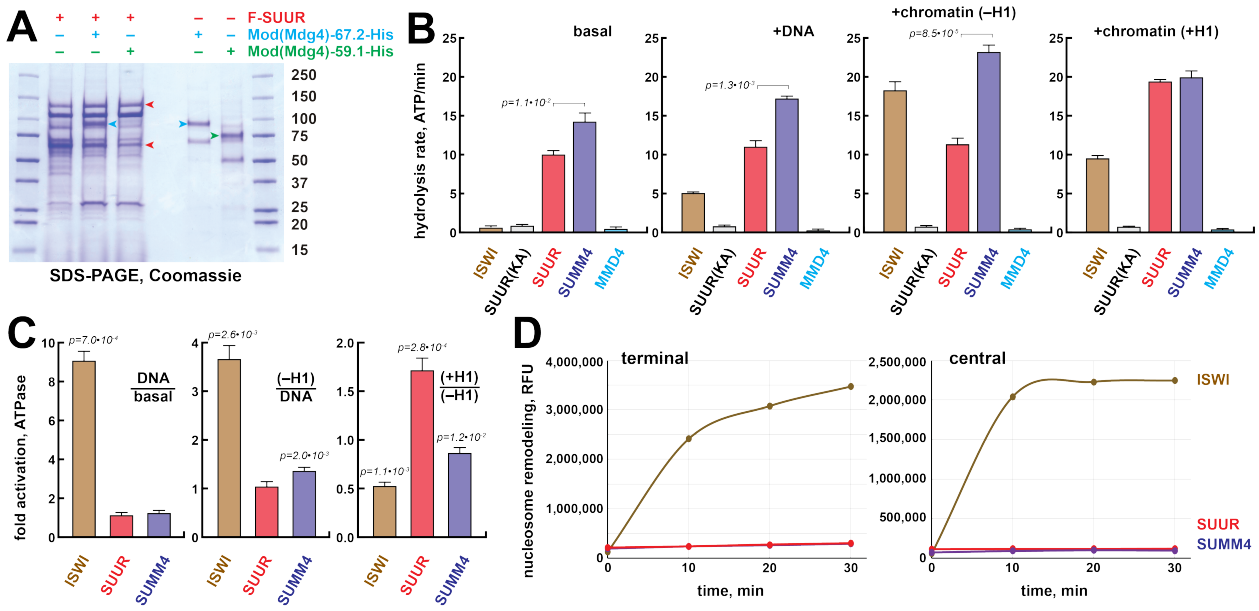


Figure 2. Biochemical activities of SUMM4. (A) Recombinant SUMM4. Mod(Mdg4)-His$_6$, 67.2 (p100, cyan arrowhead) and 59.1 (p75, green arrowhead) splice forms were co-expressed with FLAG-SUUR (red arrowheads, p130 and p65) or separately in Sf9 cells and purified by FLAG or Ni-NTA affinity chromatography. Mod(Mdg4)-67.2 forms a specific complex with SUUR. Identities of the 130-, 100-, 75- and 65-kDa protein bands from FLAG- and Ni-NTA-purified material were determined by mass-spectroscopy. (B) ATPase activities of recombinant ISWI (brown bars), FLAG-SUUR (red bars) and SUMM4 (FLAG-SUUR + Mod(Mdg4)-67.2-His$_6$, purple bars). Equimolar amounts of proteins were analyzed in reactions in the absence or presence of plasmid DNA or equivalent amounts of reconstituted oligonucleosomes, ±H1. SUUR(KA) and MMD4, ATPases activities of K59A mutant of SUUR (gray bars) and Mod(Mdg4)-67.2-His$_6$ (cyan bars). Hydrolysis rates were converted to moles ATP per mole protein per minute. All reactions were performed in triplicate, error bars represent standard deviations. *p*-values for statistically significant differences are indicated (Mann-Whitney test). (C) DNA- and nucleosome-dependent stimulation or inhibition of ATPase activity. The activities were analyzed as in (B). Statistically significant differences are shown (Mann-Whitney test). (D) Nucleosome sliding activities by EpiDyne®-PicoGreen™ assay (see *Materials and Methods*) with 5 nM of recombinant ISWI, SUUR or SUMM4. Reaction time courses are shown for terminally (6-N-66) and centrally (50-N-

942 66) positioned mononucleosomes (***Figure 2—figure supplement 1G-J***). RFU, relative fluorescence

943 units produced by PicoGreen fluorescence.

944

945 **Figure supplement 1.** Recombinant proteins, substrates and biochemical assays.

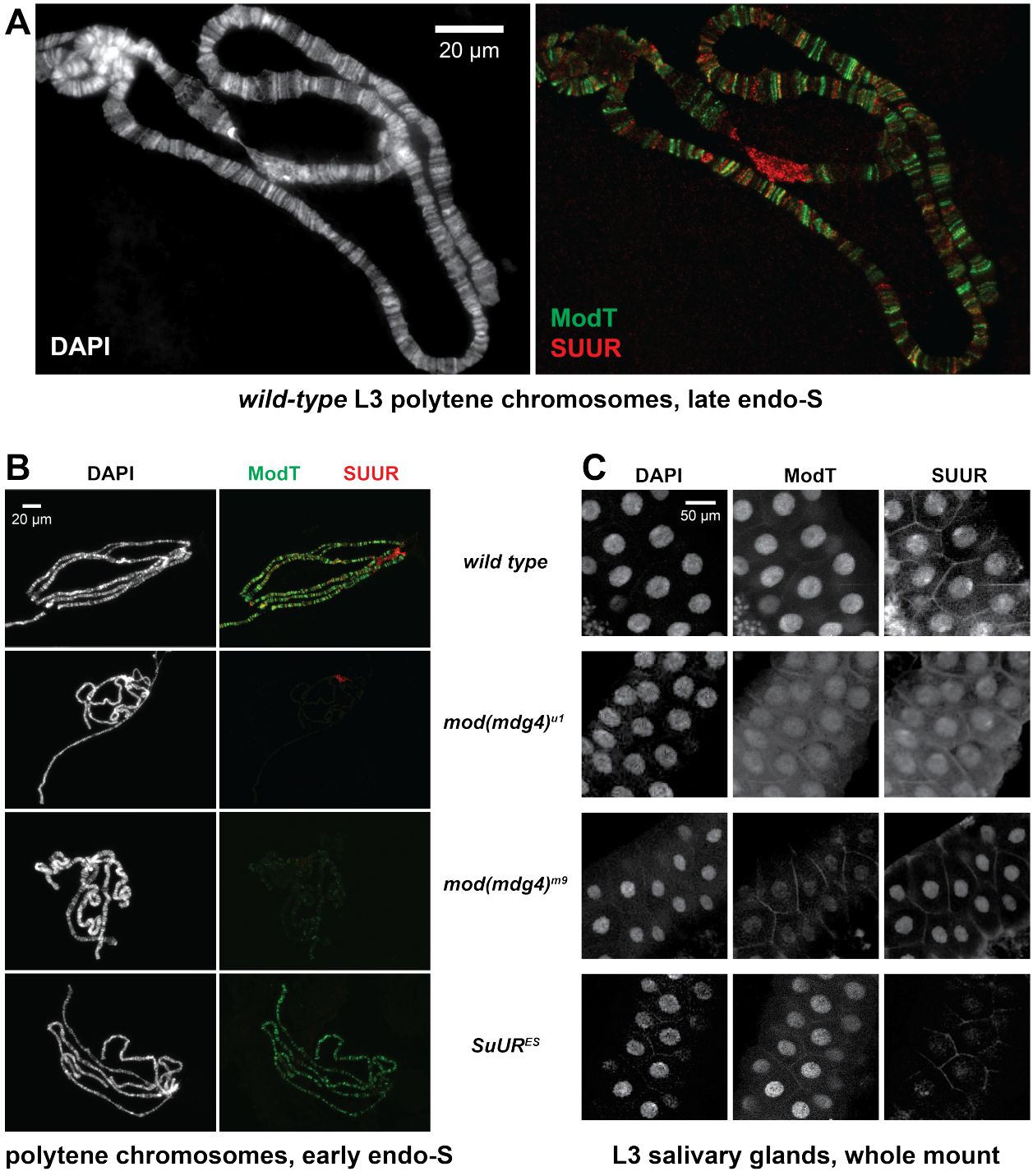


Figure 3. Spatiotemporal distribution of SUMM4 *in vivo*. (A) Colocalization of SUUR and Mod(Mdg4)-67.2 in *wild-type* polytene chromosomes. Localization patterns of Mod(Mdg4)-67.2 and SUUR in L3 polytene chromosomes were analyzed by indirect IF staining. The polytene spread fragment (3L and 3R arms) corresponds to a nucleus in late endo-S phase, according to PCNA staining (***Figure 3—figure supplement 1A***). ModT (green) and SUUR (red) signals overlap extensively in

952 euchromatic arms. Note the additional strong ModT IF loci that are SUUR-free as well as Mod(Mdg4)-
953 67.2-free SUUR in pericentric 3LR. DAPI staining shows the overall chromosome morphology. (B)
954 SUUR loading into chromosomes during early endo-S phase is compromised in *mod(mdg4)* mutants.
955 *SuUR* mutation does not appreciably change the distribution of Mod(Mdg4)-67.2. Endo-S timing was
956 established by PCNA staining (***Figure 3—figure supplement 1J***). (C) Abnormal subcellular distribution
957 of SUMM4 subunits in *mod(mdg4)* and *SuUR* mutants. L3 salivary glands were fixed and whole-mount-
958 stained with DAPI, ModT and SUUR antibodies. Whereas both polypeptides are mostly nuclear in wild
959 type, they are partially mis-localized to the cytoplasm in *mod(mdg4)*^{u1} mutant.
960
961 **Figure supplement 1.** Spatiotemporal distribution of SUUR and Mod(Mdg4) *in vivo* and alternative
962 complex(es) of Mod(Mdg4).

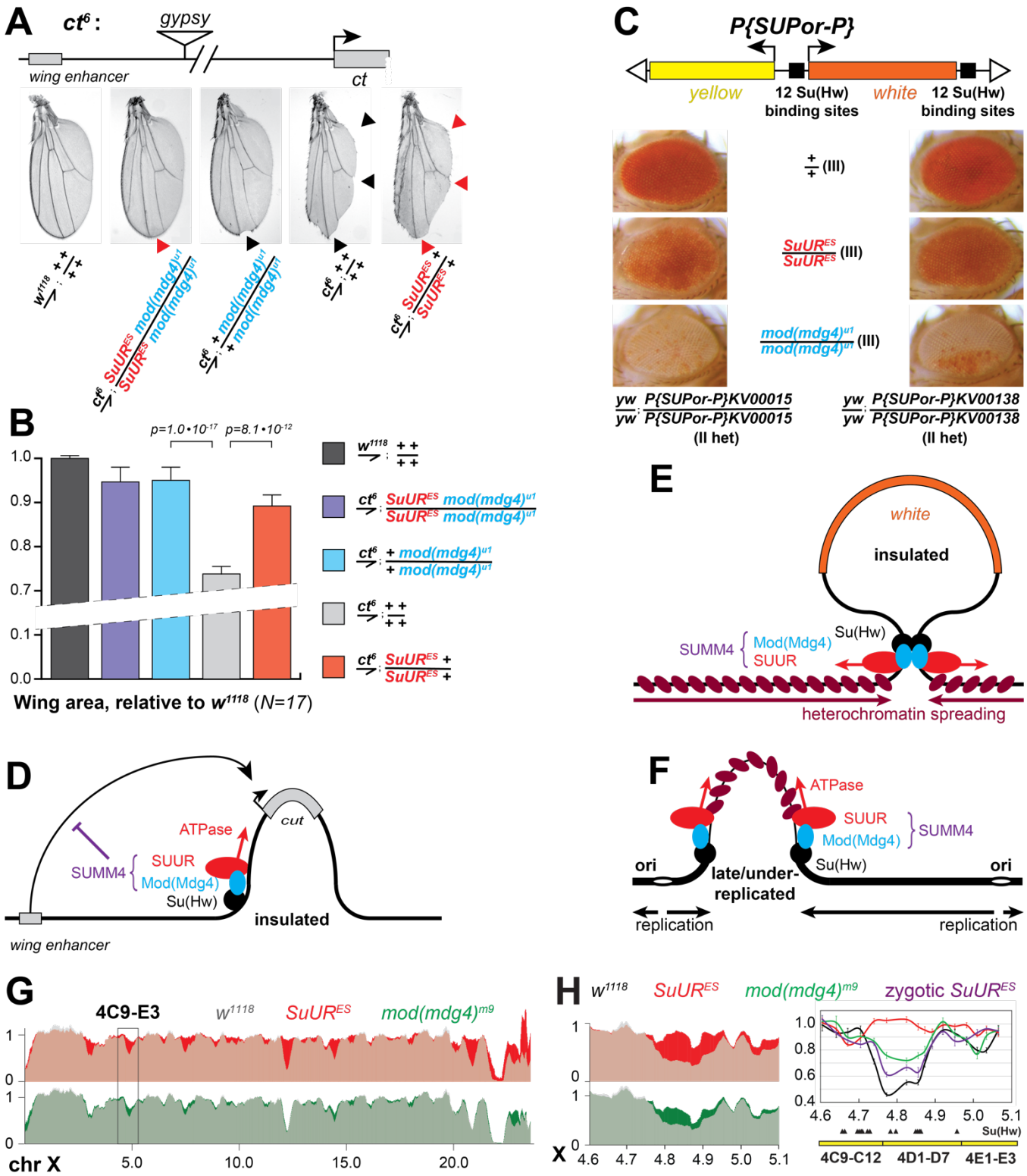


Figure 4. Biological functions of SUMM4 in regulation of gene expression and DNA replication. (A) SUMM4 subunits are required for the enhancer-blocking activity in *ct*^{6}. Top: schematic diagram of the *ct*^{6} reporter system; the *gypsy* retrotransposon is inserted in between the wing enhancer and promoter

967 of *cut* (Bag, Dale, Palmer, & Lei, 2019). Bottom left: the appearance of wild type adult wing; bottom
968 right: the appearance of ct^6 adult wing in the wild-type background. $SuUR^{ES}$ and $mod(mdg4)^{u1}$ alleles
969 are recessive suppressors of the ct^6 phenotype. Red and black arrowheads point to distinct anatomical
970 features of the wing upon *SuUR* mutation. (B) Relative sizes (areas) of wings in adult male flies of the
971 indicated phenotypes were measured as described in *Materials and Methods*. *p*-values for statistically
972 significant differences are indicated (t- test). (C) SUMM4 subunits are required for the chromatin
973 barrier activity of Su(Hw) binding sites. Top: schematic diagram of the *P{SUPor-P}* reporter system
974 (Bellen et al., 2004); clustered 12 copies of *gypsy* Su(Hw) binding sites flanks the transcription unit of
975 *white*. *KV00015* and *KV00138* are *P{SUPor-P}* insertions in pericentric heterochromatin of 2L.
976 $SuUR^{ES}$ and $mod(mdg4)^{u1}$ alleles are recessive suppressors of the boundary that insulates *white* from
977 heterochromatin encroachment. (D) Schematic model for the function of SUMM4 in blocking
978 enhancer-promoter interactions in the ct^6 locus. (E) Schematic model for the function of SUMM4 in
979 establishing a chromatin barrier in heterochromatin-inserted *P{SUPor-P}* elements. (F) Schematic
980 model for a putative function of SUMM4 in blocking/retardation of replication fork progression in IH
981 domains. (G) Analyses of DNA copy numbers in *Drosophila* salivary gland cells. DNA from L3
982 salivary glands was subjected to high-throughput sequencing. DNA copy numbers (normalized to
983 diploid embryonic DNA) are shown across the X chromosome. Genomic coordinates in Megabase
984 pairs are indicated at the bottom. The control trace (w^{1118} allele) is shown as semitransparent light gray
985 in the foreground; $SuUR^{ES}$ (homozygous null) and $mod(mdg4)^{m9}$ (zygotic null from crosses of
986 heterozygous parents) traces are shown in the background in red and green, respectively; their overlaps
987 with w^{1118} traces appear as lighter shades of colors. Black box, 4C9-E3 cytological region. (H) Close-
988 up view of DNA copy numbers in region 4C9-E3 from high-throughput sequencing data are presented
989 as in (G). DNA copy numbers were also measured independently by real-time qPCR. The numbers
990 were calculated relative to embryonic DNA and normalized to a control intergenic region. The X-axis
991 shows chromosome positions (in Megabase pairs) of target amplicons. Black, w^{1118}; red, $SuUR^{ES}$

992 (homozygous null); green, *mod(mdg4)*^{m9} (zygotic null from crosses of heterozygous parents); purple,

993 *SuUR*^{ES} (zygotic null from crosses of heterozygous parents). Error bars represent the confidence

994 interval (see *Materials and Methods*). Black arrowheads, positions of mapped Su(Hw) binding sites

995 (Negre et al., 2010). Yellow boxes show approximate boundaries of cytogenetic bands.

996

997 **Figure supplement 1.** Biological functions of SUMM4 in regulation of under-replication.

998

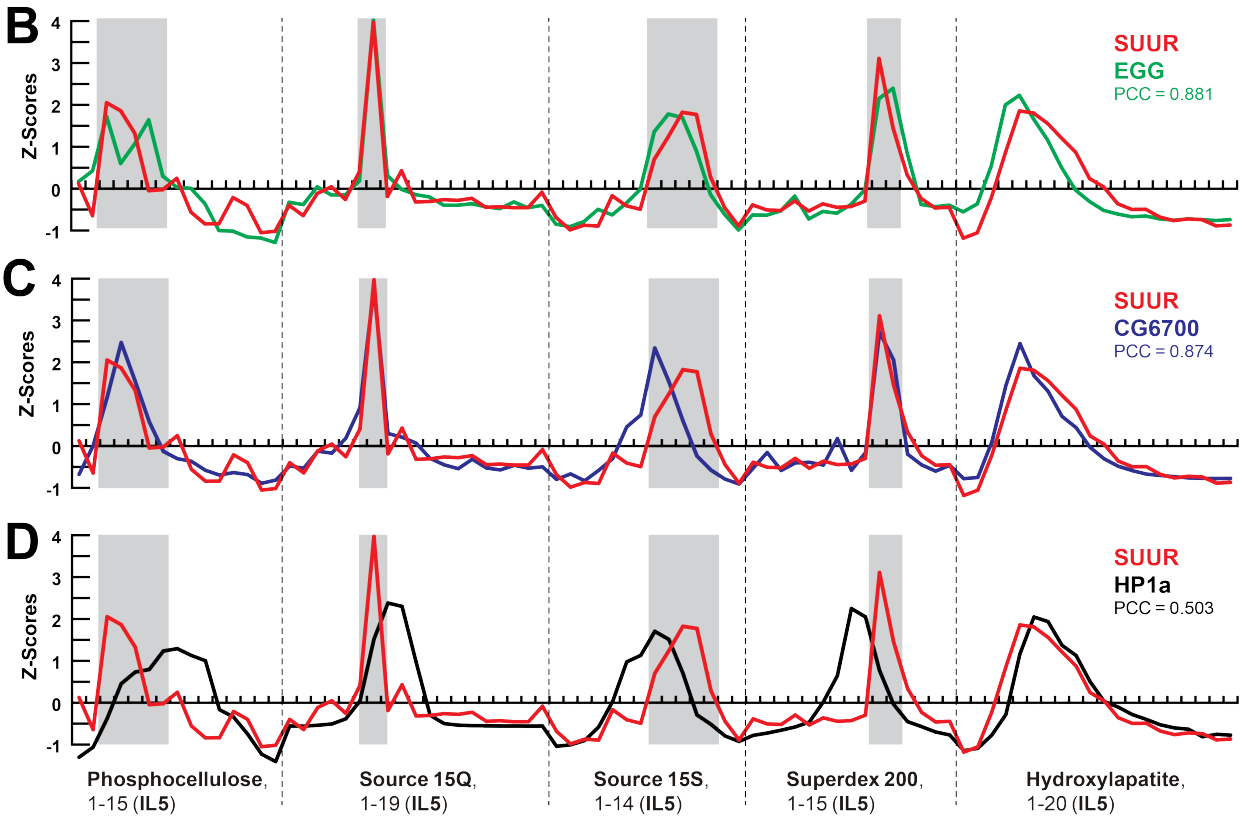
999 **Figure 4—table 2.** Primer sequences used for qPCR.

A

step (ion library)	fractions	µg on IDA, average	rank / proteins	peptides SUUR / total
nuclear extract	1	8	none / 1,448	0 / 7,787
rec. SUUR (ILR)	1	1	1 / 3	24 / 26
Phosphocellulose (IL1)	15	8	none / 2,128	0 / 25,343
Source 15Q (IL2)	19	8	1,377 / 1,813	2 / 23,227
Source 15S (IL3)	14	5	166 / 622	7 / 7,016
Superdex 200 (IL4)	15	2	134 / 318	9 / 3,494
Hydroxylapatite (IL5)	20	1	48 / 140	10 / 1,891

E

sequence	confidence	theor. *m/z*	obs. *m/z*	z	RT, min	MS2 counts
SSLTPASSSAGVK	>99%	596.314	596.315	2	21.4	2,444.44
TSVVSAAEAK	>99%	481.761	481.762	2	22.1	786.02
LHQQSPQQVR	>99%	407.555	407.555	3	14.7	3,335.51
QEEAEYIDLPMELPTK	>99%	953.461	953.460	2	45.3	1,522.16
LQAATLNEEASEPAVY	>99%	853.418	853.418	2	42.1	633.45
VFPYEGEHVHFMQASDK	>99%	505.986	505.986	4	41.8	2,216.17
SCLPSQFMPGESGVISSLSPSK	>99%	765.705	765.703	3	45.2	1,011.73



F Mod(Mdg4)-PT (67.2)

mod(mdg4): exons 2-4 *pre-mod(mdg4)-T*: exons 2-3

```
  1 MADDEQFSLCWNNFNTNLSAGFHESLCRGDLVDVSLAAEGQIVKAHRLVLSVCSPFFRKMFTQMPSNTHAIVFLNNVSHSALKDLIQFMYCG  92
 93 EVNVKQDALPAFISTAESLQIKGLTDNDPAPQPPQESSPPPAAPHVQQQQIPAQRVQRQQPRASARYKIETVDDGLGDEKQSTTQIVIQTTA 184
185 APQATIVQQQQPQQAAQQIQSQQLQTGTTTTATLVSTNKRSAQRSSLTPASSSAGVKRSKTSTSANVMDPLDSTTETGATTTAQLVPQQITV 276
277 QTSVVSAAEAKLHQQSPQQVRQEEAEYIDLPMELPTKSEPDYSEDHGDAAGDAEGTYVEDDTYGDMRYDDSYFTENEDAGNQTAANTSGGGV 368
369 TATTSKAVVKQQSQNYSESSFVDTSGDQGNTEAQAATSASATKIPPRKRGRPKTKVEDQTPKPKLLEKLQAATLNEEASEPAVYASTTKGGV 460
461 KLIFNGHLFKFSFRKADYSVFQCCYREHGEECKVRVVCDQKRVFPYEGEHVHFMQASDKSCLPSQFMPGESGVISSLSPSKELLMKNTTKLE 552
553 EADDKEDEDFEEFEIQEIDEIELDEPEKTPAKEEEVDPNDFREKIKRRLQKALQNKKK                                  610
```

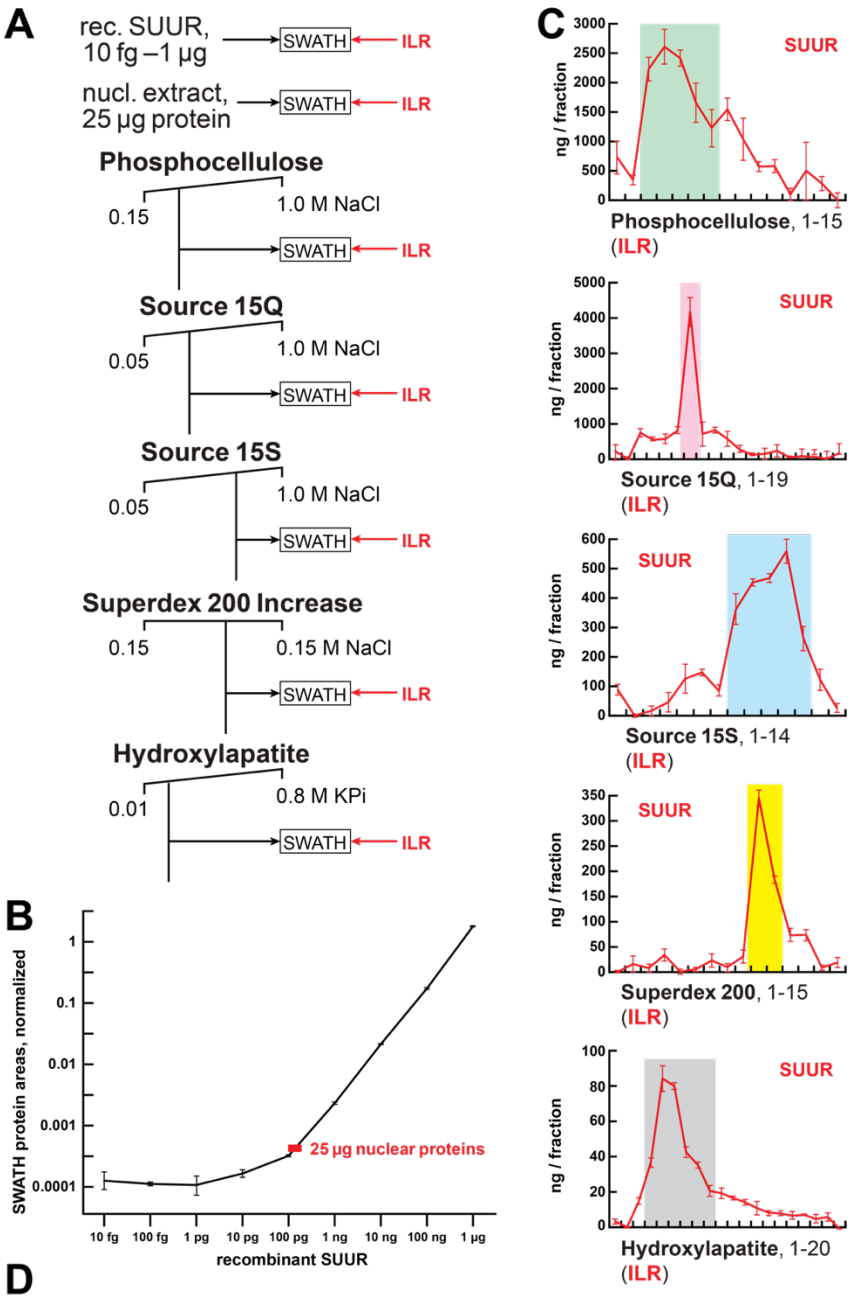
G Mod(Mdg4)-PI (59.1)

mod(mdg4): exons 2-4 *pre-mod(mdg4)-I*: exons 4-5

```
  1 MADDEQFSLCWNNFNTNLSAGFHESLCRGDLVDVSLAAEGQIVKAHRLVLSVCSPFFRKMFTQMPSNTHAIVFLNNVSHSALKDLIQFMYCG  92
 93 EVNVKQDALPAFISTAESLQIKGLTDNDPAPQPPQESSPPPAAPHVQQQQIPAQRVQRQQPRASARYKIETVDDGLGDEKQSTTQIVIQTTA 184
185 APQATIVQQQQPQQAAQQIQSQQLQTGTTTTATLVSTNKRSAQRSSLTPASSSAGVKRSKTSTSANVMDPLDSTTETGATTTAQLVPQQITV 276
277 QTSVVSAAEAKLHQQSPQQVRQEEAEYIDLPMELPTKSEPDYSEDHGDAAGDAEGTYVEDDTYGDMRYDDSYFTENEDAGNQTAANTSGGGV 368
369 TATTSKAVVKQQSQNYSESSFVDTSGDQGNTEAQVCDDLDDMKGAIKHSLLTFIRGQRGCKLLAFNGHNYVRNRRSNLKTYWICSKKGSTKC 460
461 NARVVTNVVEGVHKIVLESCHHTCLNTERKKRLSVTNVVGKARSKSEKSVSTGFIKEEGDEDLTLELRTLNLSIEDLNNLQ            541
```

1000

1001 **Figure 1—figure supplement 1.** Identification of the SUMM4 complex by MERCI. **(**A) Representation
1002 of SUUR in ion libraries ILR and IL1-5 (***Figure 1B, Table 1***). Total number of identified proteins and the
1003 confidence rank of SUUR among them as well as the total number of detected peptides (95% confidence)
1004 and the number of SUUR-specific peptides are shown. (B-D) SWATH quantitation profiles of SUUR
1005 (red), EGG (green), CG6700 (blue) and HP1a (black) fractionation across five FPLC steps as in ***Figure***
1006 ***1D***. Pearson coefficients (PCC) are shown (***Figure 1E***). (E) Mod(Mdg4)-specific peptides from ion library
1007 IL5 (***Data S1***). Gray shading, peptides specific to the common part (coding exons 2-4) of Mod(Mdg4);
1008 cyan shading, peptides specific to polypeptide Mod(Mdg4)-67.2 encoded by *pre-mod(mdg4)-T*, exons 2-
1009 3. (F) Mod(Mdg4)-67.2 polypeptide sequence. The common part is shaded in gray, splice form-specific
1010 part is shaded in cyan. Peptides from ion library IL5 (as in E) are highlighted in bold red. (G) Mod(Mdg4)-
1011 59.1 polypeptide sequence. The common part is shaded in gray, splice form-specific part is shaded in
1012 green.



D

step	SUUR	total protein	purity	purification factor	cumulative
nuclear extract	14 ± 1 µg	2.5 g	0.0006%	N/A	N/A
Phosphocellulose	7.9 ± 0.3 µg	80 mg	0.010%	17.6	18
Source 15Q	4.2 ± 0.4 µg	6.3 mg	0.07%	6.8	119
Source 15S	2.1 ± 0.1 µg	1.4 mg	0.15%	2.3	268
Superdex 200	530 ± 12 ng	70 µg	0.8%	5.0	1,352
Hydroxylapatite	300 ± 4 ng	16 µg	1.9%	2.5	3,348

Figure 1—figure supplement 2. Quantification of SUUR in chromatographic fractions. (A) Schematic of SWATH quantification of recombinant SUUR, nuclear extract (starting material) and FPLC fractions for SUUR using ion library ILR. (B) SUUR titration curve obtained by SWATH quantitation of 10 fg – 1 µg recombinant FLAG-SUUR in the presence of 25 µg *E. coli* lysate; both axes are logarithmic ($\log_{10}$).

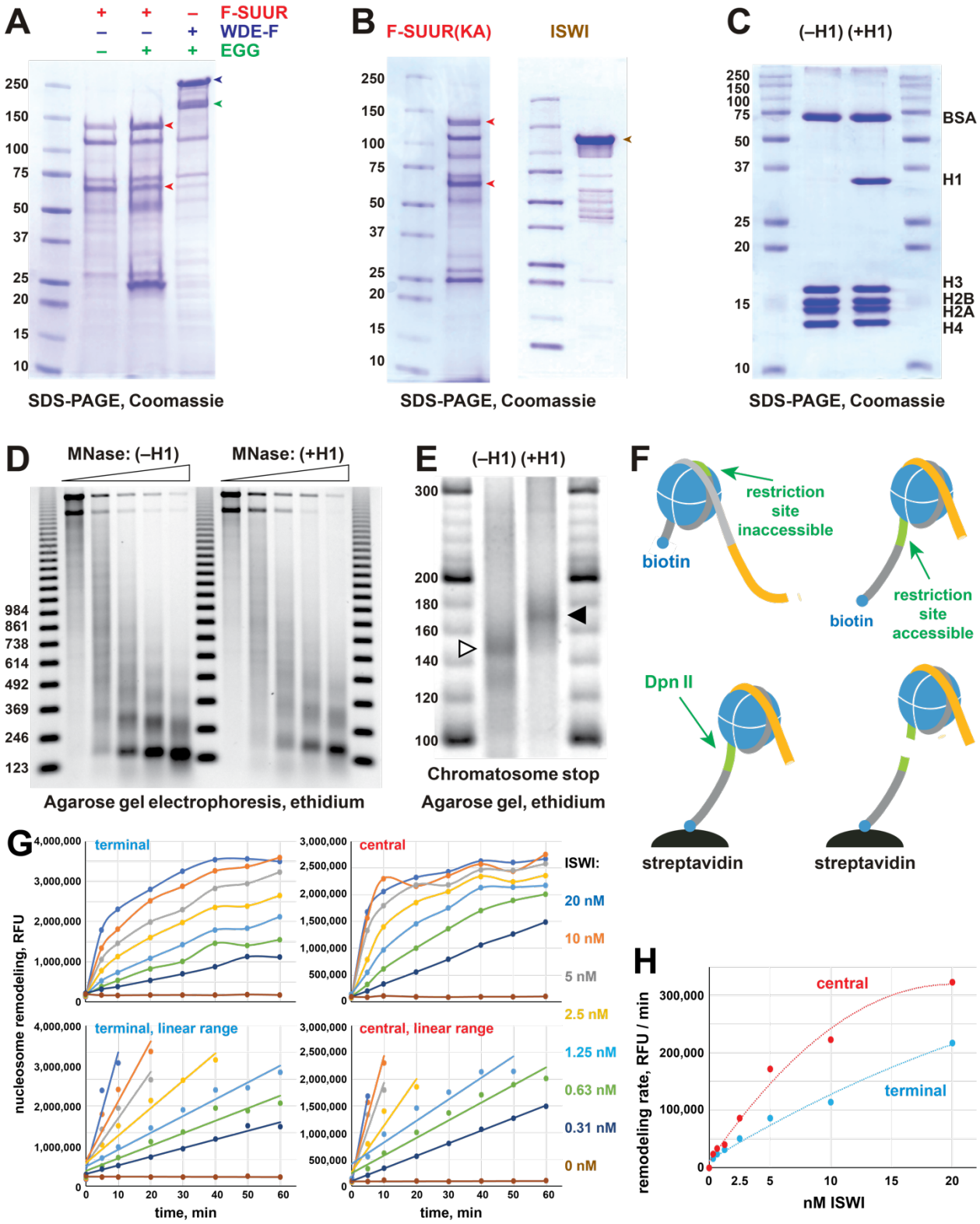
1018 Red rectangle, SUUR quantification in 25 μg nuclear extract; error bars, standard deviations (*N*=3). (C)
1019 SWATH quantitation profiles of SUUR fractionation across individual FPLC steps. Ion library ILR was
1020 used for SWATH quantification, and relative amounts were converted to estimated ng SUUR per fraction.
1021 Error bars, standard deviations (*N*=3); colored boxes, peak fractions of SUUR. (D) SUUR purification by
1022 FPLC. Total protein was measured by BCA assay, and SUUR was measured as in (C). Relative purity,
1023 purification factor in each step and cumulative purification factor are shown.

1024 **Table 1.** UR domains and UR suppression in SUMM4 subunit mutant alleles. Domains of underreplication
1025 (UR) in euchromatic arms of polytene chromosomes were called in w^{1118} as described in *Methods*. Their
1026 genomic coordinates, approximate cytological location ("Cyto band") and average DNA copy numbers
1027 ("<CN>") in homozygous w^{1118}, $SuUR^{ES}$ and $mod(mdg4)^{m9}$ L3 larvae are shown. <CN> numbers were
1028 normalized to the average DNA copy numbers across euchromatic genome. Percent UR recovery levels were
1029 calculated as $(\langle CN\rangle_{mut} - \langle CN\rangle_{w1118}) / (1 - \langle CN\rangle_{w1118})$; negative numbers indicate increased UR. UR *p*-values
1030 were calculated using the DESeq2 package by averaging the Wald test *p*-values of each 5-kbp bin significantly
1031 different than the w^{1118} signal. UR was called as suppressible by a mutant if $p < 0.01$; regions that do not exhibit
1032 a statistically significant recovery of UR are marked in red. Averages of <CN> across all called UR domains
1033 and averages of percent Recovery across all suppressible UR domains ("<Recovery>", bottom row) were
1034 adjusted for each UR domain length; calculation errors = standard deviations.

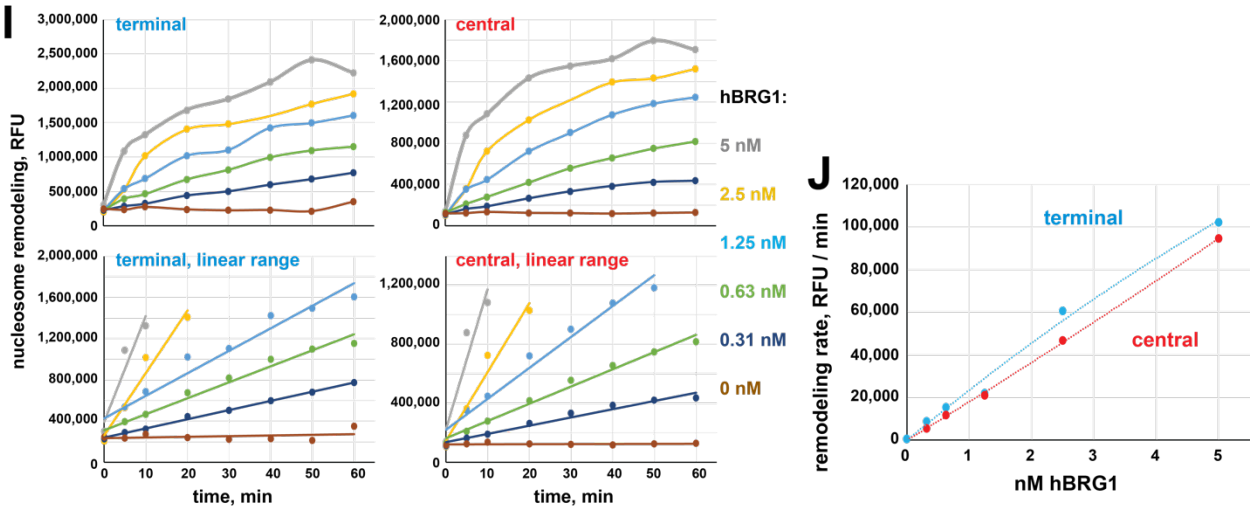
N	chromosome coordinates				Length	UR, w^{1118}	UR, $SuUR^{ES}$			UR, $mod(mdg4)^{m9}$		
	arm	left	right	Cyto band		<CN>	<CN>	Recovery	*p*-value	<CN>	Recovery	*p*-value
1	X	2,950,001	3,140,000	3C3-C7	190,000	0.51	0.93	86%	**7.3E-05**	0.58	14%	1.1E-02
2	X	4,710,001	4,900,000	4C15-D5	190,000	0.56	0.96	92%	**3.9E-04**	0.81	57%	**6.9E-05**
3	X	4,965,001	5,070,000	4E1-E2	105,000	0.72	0.86	50%	**5.6E-04**	0.80	28%	1.4E-02
4	X	6,415,001	6,525,000	6A1-B1	110,000	0.71	0.90	65%	**1.4E-03**	0.80	29%	**7.3E-03**
5	X	7,335,001	7,560,000	7B1-B4	225,000	0.65	0.98	95%	**1.2E-03**	0.79	40%	**2.8E-03**
6	X	7,750,001	7,865,000	7B7-C1	115,000	0.64	0.94	84%	**3.0E-09**	0.84	55%	**5.2E-07**
7	X	8,880,001	9,005,000	8B5-C2	125,000	0.73	0.86	50%	**5.5E-03**	0.76	9%	**4.6E-03**
8	X	9,405,001	9,555,000	8D12-E7	150,000	0.72	0.91	67%	**3.6E-04**	0.85	47%	**3.6E-03**
9	X	11,170,001	11,325,000	10A10-B3	155,000	0.67	0.84	53%	**3.2E-03**	0.78	35%	**2.6E-03**
10	X	12,040,001	12,430,000	11A2-A10	390,000	0.38	0.97	94%	**1.4E-08**	0.42	6%	**6.8E-03**
11	X	13,950,001	14,100,000	12D1-E1	150,000	0.69	0.72	10%	1.0E-02	0.73	14%	1.4E-02
12	X	14,290,001	14,565,000	12E7-F1	275,000	0.51	0.94	87%	**4.1E-04**	0.69	36%	**8.1E-04**
13	X	17,925,001	18,030,000	16F3-F5	105,000	0.67	0.99	98%	**1.7E-15**	0.90	68%	**3.4E-05**
14	X	20,000,001	20,105,000	19A4-B1	105,000	0.79	1.12	157%	**1.4E-13**	0.82	12%	**6.1E-03**
15	X	20,525,001	21,020,000	19D2-E7	495,000	0.50	0.97	93%	**1.3E-07**	0.51	2%	**4.9E-03**
16	X	21,630,001	22,450,000	20A5-C1	820,000	0.04	0.32	29%	**1.8E-03**	0.06	2%	**6.4E-03**
17	X	22,550,001	22,995,000	20C2-F3	445,000	0.48	0.81	64%	**7.8E-05**	0.74	51%	**3.5E-04**
18	2L	3,920,001	4,025,000	24D1-D4	105,000	0.63	0.93	81%	**7.9E-07**	0.80	46%	**5.9E-05**
19	2L	4,585,001	4,790,000	25A2-A5	205,000	0.66	0.99	98%	**1.9E-08**	0.78	36%	**1.3E-03**
20	2L	5,400,001	5,510,000	25E1-E4	110,000	0.82	0.99	95%	**4.0E-08**	0.90	45%	**8.3E-03**
21	2L	6,155,001	6,320,000	26B9-C2	165,000	0.74	1.08	130%	**7.3E-14**	0.88	54%	**4.7E-04**
22	2L	9,030,001	9,150,000	29F8-30A2	120,000	0.76	0.98	93%	**1.5E-04**	0.95	79%	**3.3E-03**
23	2L	11,535,001	11,795,000	32F2-33A1	260,000	0.44	0.90	83%	**2.9E-04**	0.57	24%	**1.5E-03**
24	2L	12,215,001	12,340,000	33D3-E1	125,000	0.58	0.86	66%	**3.6E-11**	0.75	40%	**1.1E-04**
25	2L	12,765,001	12,970,000	33F5-34A3	205,000	0.55	0.91	79%	**8.8E-04**	0.73	40%	**7.0E-05**
26	2L	14,685,001	15,010,000	35B4-B8	325,000	0.41	0.88	80%	**5.7E-04**	0.54	23%	**7.2E-04**
27	2L	15,295,001	15,735,000	35D1-D4	440,000	0.49	0.76	53%	**2.3E-05**	0.54	9%	**4.0E-03**
28	2L	15,770,001	15,900,000	35D4-D6	130,000	0.54	0.87	71%	**4.5E-08**	0.68	31%	**6.7E-04**

29	2L	15,925,001	16,240,000	35D6-F1	315,000	0.29	0.90	87%	**6.7E-07**	0.38	12%	**1.4E-05**
30	2L	16,925,001	17,375,000	36B4-C7	450,000	0.23	0.89	85%	**1.4E-04**	0.26	4%	**4.3E-03**
31	2L	17,515,001	18,100,000	36C10-E4	585,000	0.34	0.87	80%	**5.0E-06**	0.36	2%	**3.7E-03**
32	2L	18,160,001	18,300,000	36E6-F2	140,000	0.67	0.99	97%	**3.3E-06**	0.90	69%	**3.1E-06**
33	2L	20,110,001	20,290,000	38C1-C4	180,000	0.48	0.69	41%	**8.9E-04**	0.46	-5%	**1.8E-03**
34	2L	20,485,001	20,620,000	38C8-D1	135,000	0.77	0.98	93%	**1.0E-06**	0.99	97%	**2.1E-05**
35	2L	21,400,001	21,550,000	39D3-E2	150,000	0.10	0.15	5%	**3.2E-03**	0.14	3%	**4.4E-03**
36	2L	21,805,001	22,125,000	40A4-E4	320,000	0.53	0.94	87%	**6.9E-05**	0.54	1%	**9.5E-03**
37	2R	4,875,001	5,050,000	41C4-D1	175,000	0.35	0.86	78%	**2.3E-10**	0.34	-1%	**4.0E-03**
38	2R	5,410,001	5,535,000	41F1-F3	125,000	0.58	0.79	50%	**1.1E-03**	0.52	-13%	**2.2E-03**
39	2R	6,290,001	6,505,000	42A14-B1	215,000	0.13	0.50	42%	**9.3E-04**	0.14	1%	**2.7E-03**
40	2R	13,620,001	13,760,000	50B6-C3	140,000	0.63	0.95	88%	**4.1E-18**	0.78	41%	**1.3E-05**
41	2R	20,355,001	20,540,000	56F17-57A5	185,000	0.56	0.92	83%	**2.0E-06**	0.71	35%	**8.2E-04**
42	2R	21,830,001	21,945,000	58A2-A4	115,000	0.72	0.95	83%	**1.1E-05**	0.71	-3%	2.2E-02
43	2R	23,145,001	23,320,000	59D1-D6	175,000	0.62	1.04	110%	**1.3E-22**	0.67	13%	**7.7E-03**
44	3L	4,840,001	5,100,000	64C1-C5	260,000	0.38	0.92	87%	**3.5E-08**	0.40	3%	**6.6E-03**
45	3L	5,385,001	5,510,000	64C15-D3	125,000	0.51	0.88	76%	**1.9E-22**	0.73	45%	**6.0E-09**
46	3L	6,290,001	6,485,000	65A11-B3	195,000	0.52	0.89	77%	**4.9E-05**	0.71	38%	**1.2E-04**
47	3L	9,180,001	9,300,000	67A1-A7	120,000	0.67	0.97	90%	**6.5E-09**	0.73	20%	1.0E-02
48	3L	10,000,001	10,195,000	67D3-D10	195,000	0.62	0.97	93%	**4.4E-13**	0.79	44%	**5.7E-06**
49	3L	13,085,001	13,220,000	70A1-A2	135,000	0.66	1.01	104%	**3.6E-09**	0.89	66%	**2.9E-06**
50	3L	13,550,001	13,855,000	70B6-C4	305,000	0.26	0.95	94%	**1.8E-06**	0.39	18%	**7.3E-04**
51	3L	15,175,001	15,500,000	71B7-D3	325,000	0.39	0.94	89%	**5.6E-04**	0.46	10%	**3.7E-03**
52	3L	17,115,001	17,240,000	73F1-74A1	125,000	0.71	1.02	106%	**4.3E-05**	0.84	45%	**2.7E-03**
53	3L	18,175,001	18,525,000	75B11-75D2	350,000	0.45	0.87	76%	**6.8E-05**	0.47	4%	**4.6E-03**
54	3L	20,555,001	20,695,000	77D1-77E3	140,000	0.60	1.02	106%	**2.2E-22**	0.84	61%	**3.6E-11**
55	3R	6,060,001	6,310,000	83D2-E4	250,000	0.70	0.92	72%	**7.6E-04**	0.63	-22%	1.0E-02
56	3R	6,495,001	6,635,000	83F1-84A1	140,000	0.53	0.96	91%	**7.8E-08**	0.71	39%	**2.2E-04**
57	3R	6,915,001	7,055,000	84B1-B2	140,000	0.64	0.93	80%	**3.9E-04**	0.82	49%	**1.9E-05**
58	3R	7,550,001	7,785,000	84D9-84E2	235,000	0.44	0.80	65%	**8.0E-06**	0.51	12%	**4.2E-03**
59	3R	10,450,001	10,660,000	86B6-C4	210,000	0.55	0.98	97%	**8.1E-11**	0.66	25%	**7.6E-04**
60	3R	10,910,001	11,140,000	88C15-86D4	230,000	0.45	0.94	89%	**2.3E-10**	0.46	2%	**2.3E-03**
61	3R	12,050,001	12,165,000	87A5-B1	115,000	0.63	0.96	88%	**9.9E-24**	0.81	49%	**5.9E-09**
62	3R	12,745,001	12,935,000	87C8-D4	190,000	0.67	0.89	68%	**7.5E-05**	0.60	-21%	1.1E-02
63	3R	14,935,001	15,055,000	88D8-D10	120,000	0.70	0.88	61%	**7.6E-06**	0.84	47%	**1.0E-04**
64	3R	16,670,001	16,970,000	89D6-E5	300,000	0.40	0.92	87%	**2.7E-09**	0.47	10%	**3.2E-03**
65	3R	17,160,001	17,355,000	89F1-90A2	195,000	0.62	0.94	84%	**1.0E-03**	0.86	64%	**2.8E-04**
66	3R	20,085,001	20,290,000	92C4-E1	205,000	0.61	0.81	53%	**1.5E-03**	0.71	26%	**3.6E-03**
67	3R	20,340,001	20,525,000	92E4-E12	185,000	0.58	0.96	91%	**5.0E-05**	0.79	50%	**7.2E-04**
68	3R	22,110,001	22,295,000	94A2-A4	185,000	0.61	0.93	83%	**3.4E-11**	0.76	39%	**3.0E-04**
69	3R	28,005,001	28,295,000	98B7-C3	290,000	0.40	0.91	85%	**2.5E-05**	0.60	32%	**6.9E-04**
70	3R	28,370,001	28,480,000	98C5-D2	110,000	0.73	0.98	94%	**1.2E-09**	0.91	66%	**4.3E-07**
UR domains: **70** <Length>: 216 ± 64 kbp Average <CN> across all UR domains: 0.49 ± 0.08							Suppressed UR domains: **69** <Length>: 217 ± 64 kbp <Recovery>: 78 ± 11%			Suppressed UR domains: **60** <Length>: 225 ± 67 kbp <Recovery>: 26 ± 9%		

1035


A
F-SUUR
WDE-F
EGG
SDS-PAGE, Coomassie
B
F-SUUR(KA)
ISWI
SDS-PAGE, Coomassie
C
(–H1) (+H1)
BSA
H1
H3
H2B
H2A
H4
SDS-PAGE, Coomassie
D
MNase: (–H1)
MNase: (+H1)
Agarose gel electrophoresis, ethidium
E
(–H1) (+H1)
Chromatosome stop
Agarose gel, ethidium
F
restriction site inaccessible
biotin
restriction site accessible
Dpn II
streptavidin
streptavidin
G
terminal
central
terminal, linear range
central, linear range
nucleosome remodeling, RFU
time, min
ISWI:
20 nM
10 nM
5 nM
2.5 nM
1.25 nM
0.63 nM
0.31 nM
0 nM
H
central
terminal
remodeling rate, RFU / min
nM ISWI

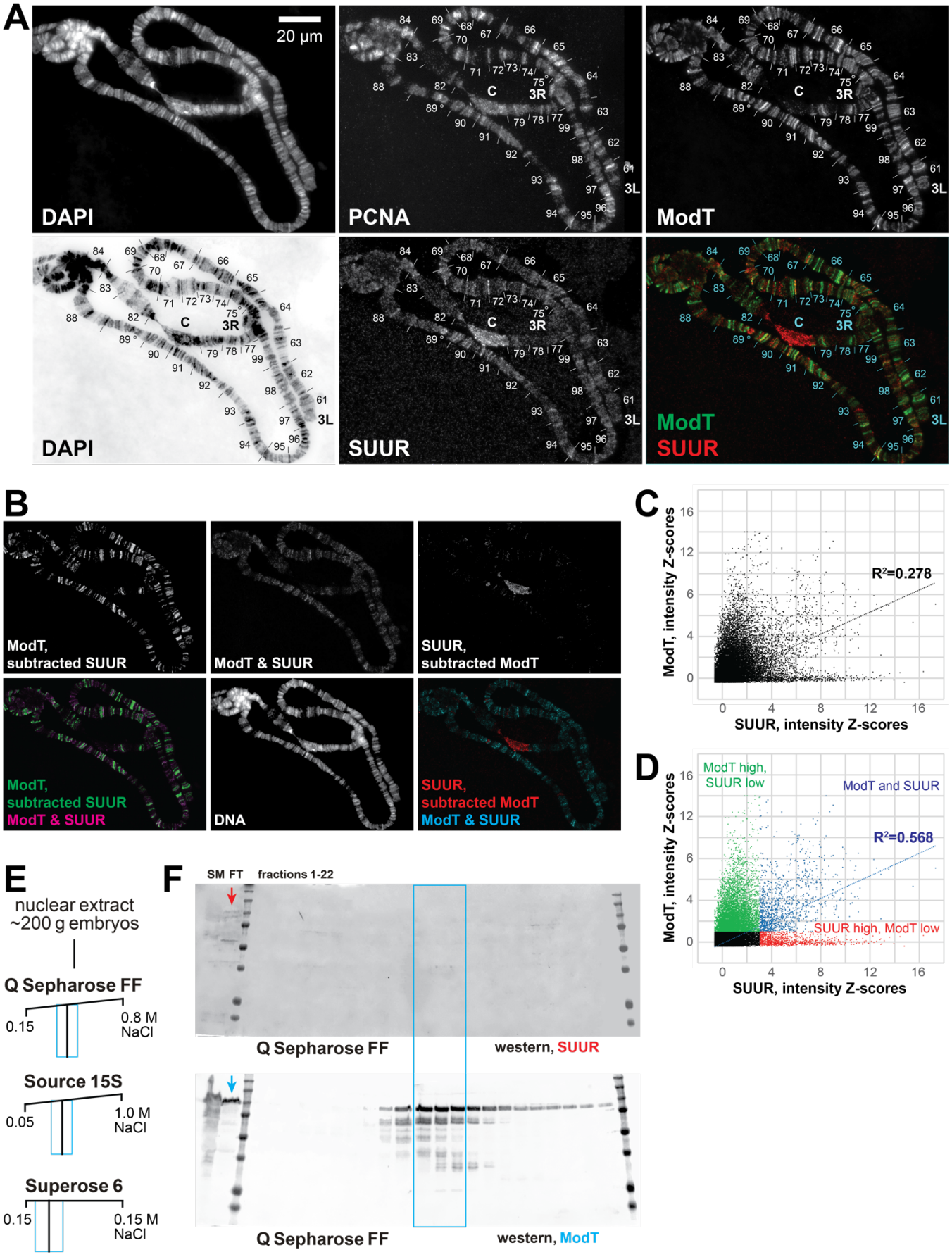
1036



1037

1038 **Figure 2—figure supplement 1.** Recombinant proteins, substrates and biochemical assays. (A) Physical
1039 interactions of recombinant EGG, SUUR and WDE. Untagged EGG (green arrowhead) was co-expressed
1040 with FLAG-SUUR (red arrowheads, p130 and p65) or WDE-FLAG (purple arrowhead) in Sf9 cells and
1041 purified by FLAG affinity chromatography. EGG forms a specific complex with WDE but not SUUR. (B)
1042 Recombinant FLAG-SUUR(K59A) expressed in Sf9 cells and ISWI expressed in *E. coli*. See legend to
1043 ***Figure 1B***. (C) Protein composition of *in vitro* reconstituted chromatin. Oligonucleosomes prepared from
1044 plasmid DNA and core histones with (+H1) or without H1 (–H1) were analyzed by SDS-PAGE and
1045 Coomassie staining. Positions of BSA, H1 and core histone bands are indicated on the right; molecular
1046 mass markers (kDa) are shown on the left. (D) Micrococcal nuclease (MNase) analysis of reconstituted
1047 chromatin. Partial digestion with five different dilutions of MNase was performed on H1-free (–H1) and
1048 H1-containing (+H1) oligonucleosomes. Deproteinated DNA fragments were analyzed by agarose gel
1049 electrophoresis and stained with ethidium. Note the increased nucleosome repeat length in (+H1) lanes
1050 consistent with H1 incorporation. Triangles at the top indicate increasing MNase concentrations; 123 bp
1051 ladder was used as a molecular mass marker. (E) Chromatosome stop assay. Oligonucleosomes assembled
1052 with or without H1 were subjected to partial MNase digestion, and DNA was analyzed by agarose gel
1053 electrophoresis and ethidium bromide staining. Positions of the core particle and chromatosome DNA are
1054 indicated by arrowheads. DNA fragment sizes in the 20-bp DNA ladder marker are shown. (F)

1055 EpiCypher® EpiDyne®-PicoGreen™ assay design. EpiDyne nucleosomes encompass a restriction site
1056 shielded by the initial nucleosome position but exposed for Dpn II cleavage upon remodeling (sliding or
1057 displacement). Biotinylated substrates are immobilized on streptavidin magnetic beads. Digest by Dpn II
1058 releases the substrates from beads, and supernatant is quantified by PicoGreen™ (dsDNA detection
1059 reagent) fluorescence. (G) Titration of *Drosophila* ISWI remodeling activity using terminally (6-N-66) or
1060 centrally (50-N-66) positioned mononucleosomes. Early reaction time points were separately plotted to
1061 indicate linear ranges. RFU, relative fluorescence units. (H) Early remodeling rates for ISWI were
1062 calculated by linear regression analyses of data in respective linear ranges. ISWI exhibits a stronger
1063 remodeling activity with a centrally positioned nucleosome substrate. (I) Titration of human BRG1
1064 remodeling activity. Data are presented as in (G). (J) Early remodeling rates for BRG1 were calculated
1065 and plotted as in (H). BRG1 does not exhibit a bias towards remodeling centrally or terminally positioned
1066 nucleosomes.


A
20 µm
DAPI
PCNA
ModT
DAPI
SUUR
ModT
SUUR
C
3R
3L
B
ModT,
subtracted SUUR
ModT & SUUR
SUUR,
subtracted ModT
ModT,
subtracted SUUR
ModT & SUUR
DNA
SUUR,
subtracted ModT
ModT & SUUR
C
ModT, intensity Z-scores
SUUR, intensity Z-scores
R²=0.278
D
ModT high,
SUUR low
ModT and SUUR
R²=0.568
SUUR high, ModT low
ModT, intensity Z-scores
SUUR, intensity Z-scores
E
nuclear extract
~200 g embryos
Q Sepharose FF
0.15
0.8 M
NaCl
Source 15S
0.05
1.0 M
NaCl
Superose 6
0.15
0.15 M
NaCl
F
SM FT
fractions 1-22
Q Sepharose FF
western, SUUR
Q Sepharose FF
western, ModT

1067

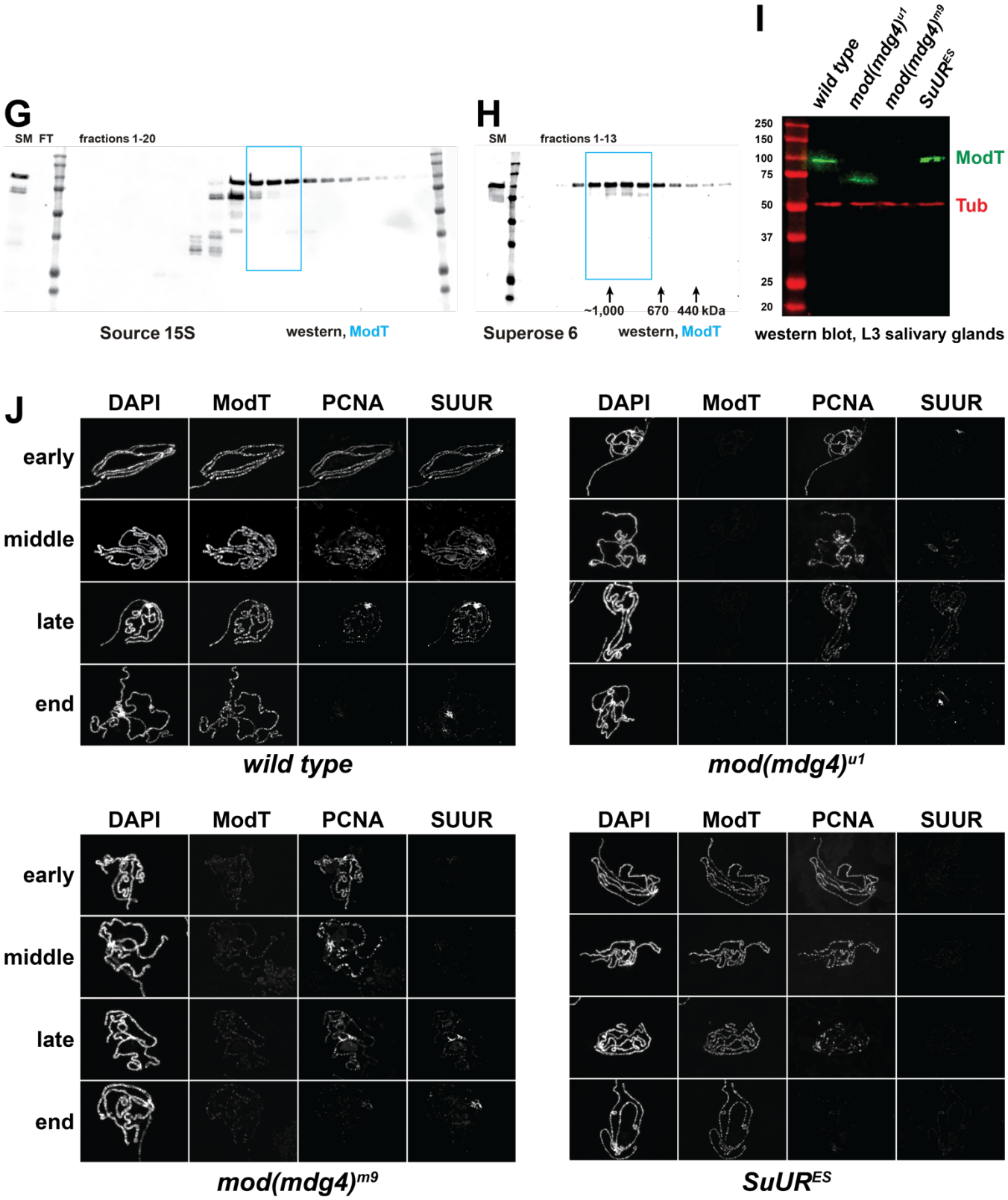
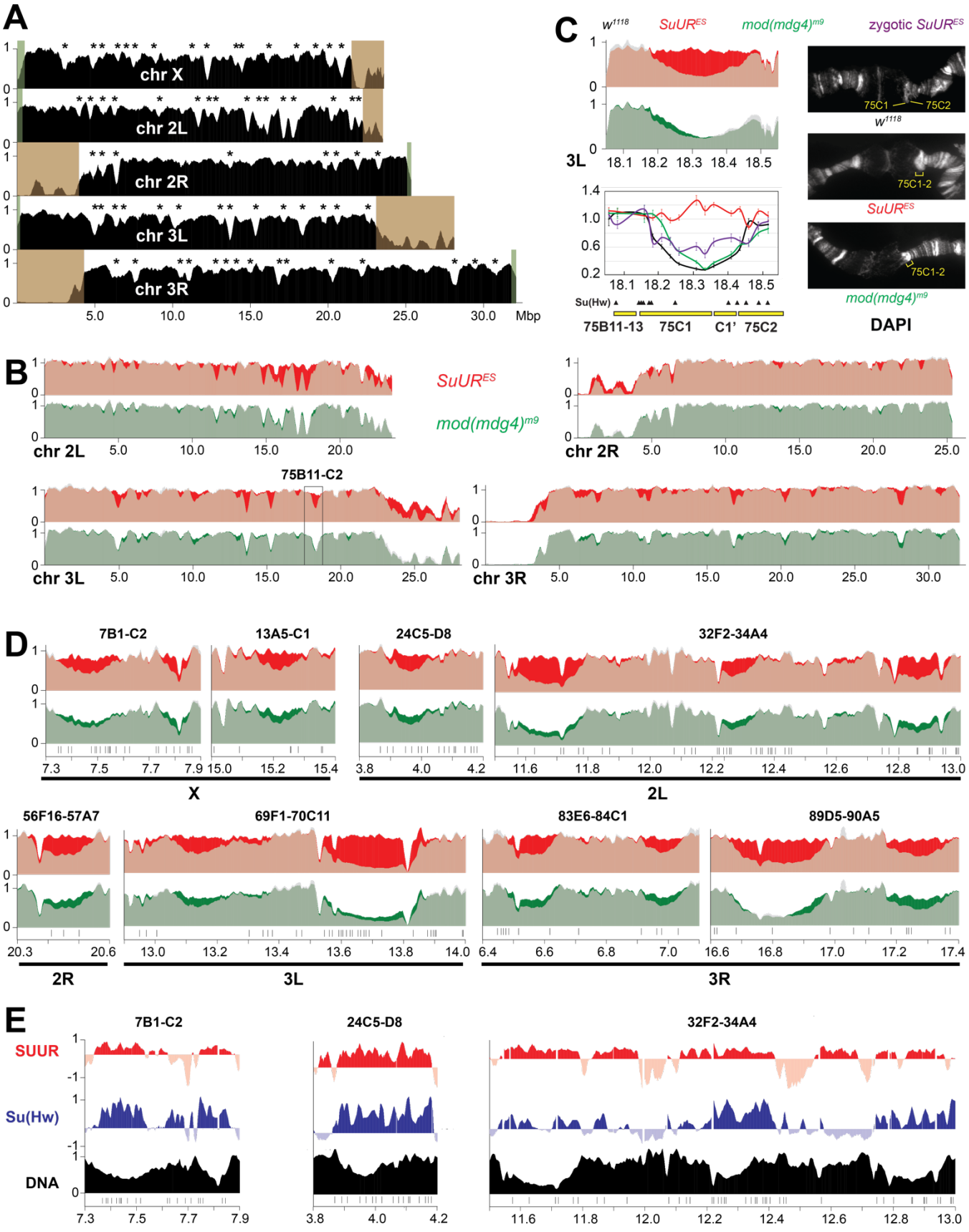


Figure 3—figure supplement 1. Spatiotemporal distribution of SUUR and Mod(Mdg4) *in vivo* and alternative complex(es) of Mod(Mdg4). (A) Colocalization of SUUR and Mod(Mdg4)-67.2 in *wild-type*

1071 polytene chromosomes. See legend to ***Figure 3A***. 3L and 3R telomeres are marked; approximate
1072 boundaries of cytological regions are shown according to (Lefevre, 1976); positions of IH regions 75C
1073 and 89E that are underreplicated and responsive to *SuUR* mutation are marked by circles. (B) The patterns
1074 of colocalization and independent loading of SUUR and Mod(Mdg4)-67.2 in *wild-type* polytene
1075 chromosomes. Subtracted and overlapping images were produced in ImageJ (*Supplemental Methods*).
1076 Green, enriched Mod(Mdg4)-67.2 and low SUUR; red, enriched SUUR and low Mod(Mdg4)-67.2;
1077 magenta or cyan, overlapping enriched Mod(Mdg4)-67.2 and SUUR. (C) Quantification of the overlap
1078 between SUUR and Mod(Mdg4)-67.2 in *wild-type* polytene chromosomes (***Figure 3A***). Individual pixel
1079 intensities of anti-SUUR and anti-ModT IF signals are normalized to Z-scores and plotted on *x*- and *y*-
1080 axes, respectively (*Supplemental Methods*); they exhibit a weak positive correlation (R^2>0.2). (D)
1081 Visually, the 2D plot (C) is split in four separate areas demarcated by Z_{ModT}= 1 and Z_{SUUR}= 3. When pixels
1082 representing ModT-only and SUUR-only areas (green and red, respectively) are removed, the remaining
1083 pixels that are simultaneously enriched for Mod(Mdg4)-67.2 and SUUR (blue) exhibit a strong positive
1084 correlation (R^2>0.5). (E) Schematic of partial FPLC purification of an alternative complex of Mod(Mdg4)-
1085 67.2. Cyan boxes, fraction ranges used for the next chromatographic step. (F) Western blot analyses of Q
1086 Sepharose FF fractions with SUUR and ModT antibodies. SUUR and ~25% total Mod(Mdg4)-67.2
1087 present in the starting material (SM) fractionate in the flow-through (FT, arrows), whereas Mod(Mdg4)-
1088 67.2 also fractionates as an additional, SUUR-free peak (cyan box). Molecular mass markers are as in
1089 ***Figure 1G***. (G) Western blot analysis of Source 15S fractions with the ModT antibody. (H) Western blot
1090 analyses of Superose 6 fractions with the ModT antibody. Black arrows, expected peaks of globular
1091 proteins with indicated molecular masses in kDa. (I) Western blot analyses of lysates of whole salivary
1092 glands. L3 salivary glands from homozygous animals of indicated genotypes were probed with ModT
1093 (green) and β-tubulin antibodies (red, loading control). Mass marker sizes (kDa) are shown on the left. (J)
1094 Spatiotemporal distribution of SUUR in polytene chromosomes. See legend to ***Figure 3B***. Although

1095 SUUR is not properly loaded into *mod(mdg4)* chromosomes during early endo-S phase (as in wild type),

1096 its deposition partially recovers during late endo-S.


A
chr X
chr 2L
chr 2R
chr 3L
chr 3R
5.0
10.0
15.0
20.0
25.0
30.0
Mbp
C
w1118
SuURES
mod(mdg4)m9
zygotic SuURES
3L
18.1
18.2
18.3
18.4
18.5
Su(Hw)
75B11-13
75C1
C1'
75C2
75C1
75C2
75C1-2
DAPI
B
SuURES
mod(mdg4)m9
chr 2L
chr 2R
75B11-C2
chr 3L
chr 3R
D
7B1-C2
13A5-C1
24C5-D8
32F2-34A4
X
2L
56F16-57A7
69F1-70C11
83E6-84C1
89D5-90A5
2R
3L
3R
E
SUUR
Su(Hw)
DNA
7B1-C2
24C5-D8
32F2-34A4

1097

1098 **Figure 4—figure supplement 1.** Biological functions of SUMM4 in regulation of under-replication. (A)
1099 Genome-wide analyses of DNA copy numbers in *Drosophila* salivary gland cells (w^{1118} control).
1100 Chromosome arms are indicated in white. Brown- and green-shades boxes, mapped pericentric and
1101 telomeric heterochromatin regions (Hoskins et al., 2015), respectively. Asterisks, positions of UR domains
1102 (***Table 1***). See legend to ***Figure 4G*** for other designations. (B) Genome-wide analyses of DNA copy
1103 numbers in *Drosophila* salivary gland cells in chromosome arms 2L, 2R, 3L and 3R. The data were
1104 obtained and presented as for the X chromosome (***Figure 4G***). Black box, 75B11-C2 cytological region.
1105 (C) Close-up view of DNA copy numbers by high-throughput sequencing and by qPCR for region 75B11-
1106 C2 (data are presented as in ***Figure 4H***) and DAPI-stained polytene chromosome segments around
1107 cytological regions 75B-75C. (D) Close-up view of DNA copy numbers by high-throughput sequencing
1108 for additional genomic regions. Approximate cytogenetic locations are indicated at the top of each panel.
1109 Short vertical bars at the bottom, positions of mapped Su(Hw) binding sites (Negre et al., 2010). See (C)
1110 for other designations. (E) Sample plots of DamID profiles for SUUR (red) and Su(Hw) (purple), $\log_2$
1111 enrichment over Dam-only control (Filion et al., 2010). Positive values are plotted in dark colors and
1112 negative values in light colors for contrast. DNA copy numbers in salivary gland cells (black) indicate UR
1113 IH domains. Vertical bars, Su(Hw) binding sites (Negre et al., 2010).

1114 **Table 2.** Primer sequences used for qPCR. Genomic coordinates indicate full amplicons, including the

1115 length of each primer. Coordinates refer to the BDGP R6/dm3 assembly.

Cytological location	4C9-E3
Sequence	**Genomic coordinates**
CCTCGATCGGTTTACATTCG CCATAAACCCAAACGAGCTG	X:4,607,333..4,607,433
CACATGGTGTCCTTGCATTC GCCTAAACCAGCGATTCAAC	X:4,643,535..4,643,626
GGGATGTGCTGCCTTTTATG AGTTGCCACGACCAAAACTC	X:4,673,491..4,673,595
TGAAGGCCCTGGATGATAAG TGGCATAGATATCGGTGTGC	X:4,706,888..4,706,995
GGCTTGATTTTCGACTGCTC AAAGGAAACAGCTCCGTGTG	X:4,742,052..4,742,153
TTGCAGTGCCTCAAAGTCAG ACCGACCAAAATCGAGACTG	X:4,774,644..4,774,740
CCTATCACCTGCCCATTTTG TTACGTCCCTGGTTTCTTGC	X:4,826,332..4,826,430
AGCCATCCTGTTGCATCTTC GCGCCAACAAATTCTCTCAG	X:4,856,456..4,856,547
ACCTCGCCAACATTACCAAC AAACAACACGACGGCTCTTC	X:4,873,801..4,873,880
AACTGCCCAAAGTGAAGGTG GTTCAAGTGCAGCCAATGTG	X:4,893,272..4,893,370
CGGCAAACACGACTACAATG CAGTCGGATGCTGGTAGATATG	X:4,920,840..4,920,943
AGCATGGACCCATCGATTAC TTTCCCTGGGTAGCATTCAC	X:4,951,780..4,951,879
GAGATGCAAGATGCCACAAG CCTTAGAGCGCTTCAATTCG	X:4,982,299..4,982,391
AGGCAACCTGCAACTGAAAC ACAATTGCGTACGTGAGCTG	X:5,009,757..5,009,859
GTCTTGGAGTTGCCGTTTTG TGCGCTGATCTCGTTAGATG	X:5,033,854..5,033,945
CTAACCATCGCCAAATCCTC CGTCCACAATTAGCTTGCAG	X:5,064,863..5,064,959
TCCCTGCGACAACCTTTAAC CTCCGTGACATGCTTGATTC	X:5,097,851..5,097,941

Cytological location	75B11-C2
Sequence	**Genomic coordinates**
ATTTGGACTGGGGCAGTTTC CTGAAACACGGAAGTTGAGTCC	3L:18,062,106..18,062,230
AAAACACAAGCACATAGGCAAC AGTTTCTGGCGTTGTATCCG	3L:18,087,066..18,087,175
GTGCACGGACGCGTATAATC AAGTTAGCTCACGTGAGATGATG	3L:18,164,427..18,164,499
ACTATTATTTCTGGCTGGCTACG GCCGGCTGCTACTTATGGC	3L:18,188,845..18,188,948
ATACAGATACAGCTCGCACTGG AGTGGTGCCGATGGAAAAAC	3L:18,214,103..18,214,210
ACCACGCCCCTAAGCAAATAG ATCTCGCCAGCTAAAGATCTCG	3L:18,238,935..18,239,021
TGGGGCATTTTTGACGGTAG GCTTTTAGCCTCGAGAAACCG	3L:18,263,954..18,264,043
CTTGGCTCAGGTTTCCCTTC AAAGGACGCCACAACAATGC	3L:18,313,914..18,314,025
ATCTCTCTGGGGCATCCAAG CGCCAGCGCAGTTAAAAGTAAC	3L:18,338,911..18,339,046
TGCACCAAGCTACACAATGG CACAGGACTCCAAATTCTGCAC	3L:18,364,090..18,364,232
AGTGATAGCGGAGTAACAGTGG GTGGCGTGGATCCAACTTTATG	3L:18,414,106..18,414,187
TGCGCTAGTTCTCACCAACG ACCAACTTAAGCACCAACTAAGG	3L:18,439,417..18,439,489
ACGGGTGCCCTTAATGTTTAC GGTCGTTGCCCATGTCTTTG	3L:18,464,296..18,464,376
CAACCCTATCCATCCATCCATG CAATCGGCCTAATTCACCCATG	3L:18,491,978..18,492,057
ACATATTCGCCGACCAAGTG ACACTAACACGTGCCCCTAAC	3L:18,520,543..18,520,680

Cytological location	86D9
Sequence	**Genomic coordinates**
TGGCGCCGCTTTCTTATTAG AGAACAGGTTTGTGCGCTTG	3R:11,261,333..11,261,450

1116

1117 ## References

1118 Adryan, B., Woerfel, G., Birch-Machin, I., Gao, S., Quick, M., Meadows, L., . . . White, R. (2007).
1119 Genomic mapping of Suppressor of Hairy-wing binding sites in Drosophila. *Genome Biol, 8*(8),
1120 R167. doi:10.1186/gb-2007-8-8-r167

1121 Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Retrieved from
1122 http://www.bioinformatics.babraham.ac.uk/projects/fastqc

1123 Andreyeva, E. N., Bernardo, T. J., Kolesnikova, T. D., Lu, X., Yarinich, L. A., Bartholdy, B. A., . . .
1124 Fyodorov, D. V. (2017). Regulatory functions and chromatin loading dynamics of linker histone H1
1125 during endoreplication in Drosophila. *Genes Dev, 31*(6), 603-616. doi:10.1101/gad.295717.116

1126 Arner, E. S., Sarioglu, H., Lottspeich, F., Holmgren, A., & Bock, A. (1999). High-level expression in
1127 Escherichia coli of selenocysteine-containing rat thioredoxin reductase utilizing gene fusions with
1128 engineered bacterial-type SECIS elements and co-expression with the selA, selB and selC genes. *J*
1129 *Mol Biol, 292*(5), 1003-1016. doi:10.1006/jmbi.1999.3085

1130 Bag, I., Dale, R. K., Palmer, C., & Lei, E. P. (2019). The zinc-finger protein CLAMP promotes gypsy
1131 chromatin insulator function in Drosophila. *J Cell Sci, 132*(5). doi:10.1242/jcs.226092

1132 Bellen, H. J., Levis, R. W., Liao, G., He, Y., Carlson, J. W., Tsang, G., . . . Spradling, A. C. (2004). The
1133 BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila*
1134 genes. *Genetics, 167*(2), 761-781.

1135 Belyaeva, E. S., Boldyreva, L. V., Volkova, E. I., Nanayev, R. A., Alekseyenko, A. A., & Zhimulev, I.
1136 F. (2003). Effect of the Suppressor of Underreplication (SuUR) gene on position-effect variegation
1137 silencing in Drosophila melanogaster. *Genetics, 165*(3), 1209-1220.
1138 doi:10.1093/genetics/165.3.1209

1139 Belyaeva, E. S., Zhimulev, I. F., Volkova, E. I., Alekseyenko, A. A., Moshkin, Y. M., & Koryakov, D.
1140 E. (1998). Su(UR)ES: a gene suppressing DNA underreplication in intercalary and pericentric

1141 heterochromatin of Drosophila melanogaster polytene chromosomes. *Proc Natl Acad Sci U S A,*
1142 *95*(13), 7532-7537. doi:10.1073/pnas.95.13.7532
1143 BroadInstitute. "Picard Tools." Broad Institute, GitHub repository. version 2.2.4. Retrieved from
1144 http://broadinstitute.github.io/picard/
1145 Buchner, K., Roth, P., Schotta, G., Krauss, V., Saumweber, H., Reuter, G., & Dorn, R. (2000). Genetic
1146 and molecular complexity of the position effect variegation modifier mod(mdg4) in Drosophila.
1147 *Genetics, 155*(1), 141-157. doi:10.1093/genetics/155.1.141
1148 Bushnell, B. (2014). BBTools software package. Retrieved from http://bbtools.jgi.doe.gov
1149 Cai, H., & Levine, M. (1995). Modulation of enhancer-promoter interactions by insulators in the
1150 Drosophila embryo. *Nature, 376*(6540), 533-536. doi:10.1038/376533a0
1151 Cheng, Q., & Arner, E. S. (2017). Selenocysteine Insertion at a Predefined UAG Codon in a Release
1152 Factor 1 (RF1)-depleted Escherichia coli Host Strain Bypasses Species Barriers in Recombinant
1153 Selenoprotein Translation. *J Biol Chem, 292*(13), 5476-5487. doi:10.1074/jbc.M117.776310
1154 Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., . . . Li, H. (2021). Twelve
1155 years of SAMtools and BCFtools. *Gigascience, 10*(2). doi:10.1093/gigascience/giab008
1156 Davidson, I. F., Bauer, B., Goetz, D., Tang, W., Wutz, G., & Peters, J. M. (2019). DNA loop extrusion
1157 by human cohesin. *Science, 366*(6471), 1338-1345. doi:10.1126/science.aaz3418
1158 Dimitrova, D. S., & Gilbert, D. M. (1999). The spatial position and replication timing of chromosomal
1159 domains are both established in early G1 phase. *Mol Cell, 4*(6), 983-993. doi:10.1016/s1097-
1160 2765(00)80227-0
1161 dos Santos, G., Schroeder, A. J., Goodman, J. L., Strelets, V. B., Crosby, M. A., Thurmond, J., . . .
1162 FlyBase, Consortium. (2015). FlyBase: introduction of the Drosophila melanogaster Release 6
1163 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Res,*
1164 *43*(Database issue), D690-697. doi:10.1093/nar/gku1099

1165 Emelyanov, A. V., Konev, A. Y., Vershilova, E., & Fyodorov, D. V. (2010). Protein complex of
1166 Drosophila ATRX/XNP and HP1a is required for the formation of pericentric beta-heterochromatin
1167 in vivo. *J Biol Chem, 285*(20), 15027-15037. doi:M109.064790 [pii] 10.1074/jbc.M109.064790
1168 Emelyanov, A. V., Rabbani, J., Mehta, M., Vershilova, E., Keogh, M. C., & Fyodorov, D. V. (2014).
1169 Drosophila TAP/p32 is a core histone chaperone that cooperates with NAP-1, NLP, and
1170 nucleophosmin in sperm chromatin remodeling during fertilization. *Genes Dev, 28*(18), 2027-2040.
1171 doi:10.1101/gad.248583.114
1172 Emelyanov, A. V., Vershilova, E., Ignatyeva, M. A., Pokrovsky, D. K., Lu, X., Konev, A. Y., &
1173 Fyodorov, D. V. (2012). Identification and characterization of ToRC, a novel ISWI-containing ATP-
1174 dependent chromatin assembly complex. *Genes Dev, 26*(6), 603-614. doi:10.1101/gad.180604.111
1175 Filion, G. J., van Bemmel, J. G., Braunschweig, U., Talhout, W., Kind, J., Ward, L. D., . . . van Steensel,
1176 B. (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila
1177 cells. *Cell, 143*(2), 212-224. doi:10.1016/j.cell.2010.09.009
1178 Fyodorov, D. V., & Kadonaga, J. T. (2003). Chromatin assembly in vitro with purified recombinant
1179 ACF and NAP-1. *Methods Enzymol, 371*, 499-515. doi:10.1016/S0076-6879(03)71037-4
1180 Fyodorov, D. V., & Levenstein, M. E. (2002). Chromatin assembly in *Drosophila* systems. In *Current*
1181 *Protocols in Molecular Biology* (pp. 21.27.21-21.27.27). New York: Wiley & Sons.
1182 Gaszner, M., & Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms.
1183 *Nat Rev Genet, 7*(9), 703-713. doi:10.1038/nrg1925
1184 Georgiev, P. G., & Gerasimova, T. I. (1989). Novel genes influencing the expression of the yellow locus
1185 and mdg4 (gypsy) in Drosophila melanogaster. *Mol Gen Genet, 220*(1), 121-126.
1186 doi:10.1007/BF00260865
1187 Gerasimova, T. I., Gdula, D. A., Gerasimov, D. V., Simonova, O., & Corces, V. G. (1995). A
1188 Drosophila protein that imparts directionality on a chromatin insulator is an enhancer of position-
1189 effect variegation. *Cell, 82*(4), 587-597. doi:10.1016/0092-8674(95)90031-4

1190 Havugimana, P. C., Hart, G. T., Nepusz, T., Yang, H., Turinsky, A. L., Li, Z., . . . Emili, A. (2012). A
1191 census of human soluble protein complexes. *Cell, 150*(5), 1068-1081. doi:10.1016/j.cell.2012.08.011
1192 Hoskins, R. A., Carlson, J. W., Wan, K. H., Park, S., Mendez, I., Galle, S. E., . . . Celniker, S. E. (2015).
1193 The Release 6 reference sequence of the Drosophila melanogaster genome. *Genome Res, 25*(3), 445-
1194 458. doi:10.1101/gr.185579.114
1195 Ito, T., Levenstein, M. E., Fyodorov, D. V., Kutach, A. K., Kobayashi, R., & Kadonaga, J. T. (1999).
1196 ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent
1197 catalysis of chromatin assembly. *Genes Dev, 13*(12), 1529-1539.
1198 Jaskelioff, M., Van Komen, S., Krebs, J. E., Sung, P., & Peterson, C. L. (2003). Rad54p is a chromatin
1199 remodeling enzyme required for heteroduplex DNA joint formation with chromatin. *Journal of*
1200 *Biological Chemistry, 278*(11), 9212-9218.
1201 Kamakaka, R. T., Tyree, C. M., & Kadonaga, J. T. (1991). Accurate and efficient RNA polymerase II
1202 transcription with a soluble nuclear fraction derived from *Drosophila* embryos. *Proceedings of the*
1203 *National Academy of Sciences of the United States of America, 88*(3), 1024-1028.
1204 Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D.
1205 (2002). The human genome browser at UCSC. *Genome Res, 12*(6), 996-1006.
1206 doi:10.1101/gr.229102
1207 Khoroshko, V. A., Levitsky, V. G., Zykova, T. Y., Antonenko, O. V., Belyaeva, E. S., & Zhimulev, I. F.
1208 (2016). Chromatin Heterogeneity and Distribution of Regulatory Elements in the Late-Replicating
1209 Intercalary Heterochromatin Domains of Drosophila melanogaster Chromosomes. *PLoS One, 11*(6),
1210 e0157147. doi:10.1371/journal.pone.0157147
1211 Kolesnikova, T. D., Posukh, O. V., Andreyeva, E. N., Bebyakina, D. S., Ivankin, A. V., & Zhimulev, I.
1212 F. (2013). Drosophila SUUR protein associates with PCNA and binds chromatin in a cell cycle-
1213 dependent manner. *Chromosoma, 122*(1-2), 55-66. doi:10.1007/s00412-012-0390-9

1214 Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods, 9*(4),
1215 357-359. doi:10.1038/nmeth.1923
1216 Lefevre, G. (1976). A photographic representation and interpretation of the polytene chromosomes of
1217 Drosophila melanogaster salivary glands. In Ashburner & Novitski (Ed.), *The Genetics and Biology*
1218 *of Drosophila.* (Vol. Ia, pp. 31-66). London: Academic Press.
1219 Lei, E. P., & Corces, V. G. (2006). RNA interference machinery influences the nuclear organization of a
1220 chromatin insulator. *Nature Genetics, 38*(8), 936-941.
1221 Lu, X., Wontakal, S. N., Kavi, H., Kim, B. J., Guzzardo, P. M., Emelyanov, A. V., . . . Skoultchi, A. I.
1222 (2013). Drosophila H1 regulates the genetic activity of heterochromatin by recruitment of Su(var)3-
1223 9. *Science, 340*(6128), 78-81. doi:10.1126/science.1234654
1224 Makunin, I. V., Volkova, E. I., Belyaeva, E. S., Nabirochkina, E. N., Pirrotta, V., & Zhimulev, I. F.
1225 (2002). The Drosophila suppressor of underreplication protein binds to late-replicating regions of
1226 polytene chromosomes. *Genetics, 160*(3), 1023-1034. doi:10.1093/genetics/160.3.1023
1227 Marchal, C., Sima, J., & Gilbert, D. M. (2019). Control of DNA replication timing in the 3D genome.
1228 *Nat Rev Mol Cell Biol, 20*(12), 721-737. doi:10.1038/s41580-019-0162-y
1229 Matthews, N. E., & White, R. (2019). Chromatin Architecture in the Fly: Living without CTCF/Cohesin
1230 Loop Extrusion?: Alternating Chromatin States Provide a Basis for Domain Architecture in
1231 Drosophila. *Bioessays, 41*(9), e1900048. doi:10.1002/bies.201900048
1232 Munden, A., Rong, Z., Sun, A., Gangula, R., Mallal, S., & Nordman, J. T. (2018). Rif1 inhibits
1233 replication fork progression and controls DNA copy number in Drosophila. *Elife, 7*.
1234 doi:10.7554/eLife.39140
1235 Negre, N., Brown, C. D., Shah, P. K., Kheradpour, P., Morrison, C. A., Henikoff, J. G., . . . White, K. P.
1236 (2010). A comprehensive map of insulator elements for the Drosophila genome. *PLoS Genet, 6*(1),
1237 e1000814. doi:10.1371/journal.pgen.1000814

1238 Nordman, J. T., Kozhevnikova, E. N., Verrijzer, C. P., Pindyurin, A. V., Andreyeva, E. N., Shloma, V.
1239 V., . . . Orr-Weaver, T. L. (2014). DNA copy-number control through inhibition of replication fork
1240 progression. *Cell Rep, 9*(3), 841-849. doi:10.1016/j.celrep.2014.10.005
1241 Nordman, J. T., & Orr-Weaver, T. L. (2015). Understanding replication fork progression, stability, and
1242 chromosome fragility by exploiting the Suppressor of Underreplication protein. *Bioessays, 37*(8),
1243 856-861. doi:10.1002/bies.201500021
1244 Peterson, S. C., Samuelson, K. B., & Hanlon, S. L. (2021). Multi-Scale Organization of the Drosophila
1245 melanogaster Genome. *Genes (Basel), 12*(6). doi:10.3390/genes12060817
1246 Pindyurin, A. V., Boldyreva, L. V., Shloma, V. V., Kolesnikova, T. D., Pokholkova, G. V., Andreyeva,
1247 E. N., . . . Zhimulev, I. F. (2008). Interaction between the Drosophila heterochromatin proteins
1248 SUUR and HP1. *J Cell Sci, 121*(Pt 10), 1693-1703. doi:10.1242/jcs.018655
1249 Posukh, O. V., Maksimov, D. A., Skvortsova, K. N., Koryakov, D. E., & Belyakin, S. N. (2015). The
1250 effects of SUUR protein suggest its role in repressive chromatin renewal during replication in
1251 Drosophila. *Nucleus, 6*(4), 249-253. doi:10.1080/19491034.2015.1074366
1252 Ramirez, F., Ryan, D. P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., . . . Manke, T. (2016).
1253 deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res,*
1254 *44*(W1), W160-165. doi:10.1093/nar/gkw257
1255 Rhind, N., & Gilbert, D. M. (2013). DNA replication timing. *Cold Spring Harb Perspect Biol, 5*(8),
1256 a010132. doi:10.1101/cshperspect.a010132
1257 Roseman, R. R., Johnson, E. A., Rodesch, C. K., Bjerke, M., Nagoshi, R. N., & Geyer, P. K. (1995). A P
1258 element containing suppressor of hairy-wing binding regions has novel properties for mutagenesis in
1259 Drosophila melanogaster. *Genetics, 141*(3), 1061-1074. doi:10.1093/genetics/141.3.1061
1260 Roseman, R. R., Pirrotta, V., & Geyer, P. K. (1993). The su(Hw) protein insulates expression of the
1261 Drosophila melanogaster white gene from chromosomal position-effects. *EMBO J, 12*(2), 435-442.
1262 doi:10.1002/j.1460-2075.1993.tb05675.x

1263 Rowley, M. J., Nichols, M. H., Lyu, X., Ando-Kuri, M., Rivera, I. S. M., Hermetz, K., . . . Corces, V. G.
1264 (2017). Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Mol Cell, 67*(5),
1265 837-852 e837. doi:10.1016/j.molcel.2017.07.022
1266 Savitsky, M., Kim, M., Kravchuk, O., & Schwartz, Y. B. (2016). Distinct Roles of Chromatin Insulator
1267 Proteins in Control of the Drosophila Bithorax Complex. *Genetics, 202*(2), 601-617.
1268 doi:10.1534/genetics.115.179309
1269 Schwartz, Y. B., Linder-Basso, D., Kharchenko, P. V., Tolstorukov, M. Y., Kim, M., Li, H. B., . . .
1270 Pirrotta, V. (2012). Nature and function of insulator protein binding sites in the Drosophila genome.
1271 *Genome Res, 22*(11), 2188-2198. doi:10.1101/gr.138156.112
1272 Shatsky, M., Dong, M., Liu, H., Yang, L. L., Choi, M., Singer, M. E., . . . Biggin, M. D. (2016).
1273 Quantitative Tagless Copurification: A Method to Validate and Identify Protein-Protein Interactions.
1274 *Mol Cell Proteomics, 15*(6), 2186-2202. doi:10.1074/mcp.M115.057117
1275 Sher, N., Bell, G. W., Li, S., Nordman, J., Eng, T., Eaton, M. L., . . . Orr-Weaver, T. L. (2012).
1276 Developmental control of gene copy number by repression of replication initiation and fork
1277 progression. *Genome Res, 22*(1), 64-75. doi:10.1101/gr.126003.111
1278 Szabo, Q., Bantignies, F., & Cavalli, G. (2019). Principles of genome folding into topologically
1279 associating domains. *Sci Adv, 5*(4), eaaw1668. doi:10.1126/sciadv.aaw1668
1280 Van Bortle, K., Nichols, M. H., Li, L., Ong, C. T., Takenaka, N., Qin, Z. S., & Corces, V. G. (2014).
1281 Insulator function and topological domain border strength scale with architectural protein
1282 occupancy. *Genome Biol, 15*(6), R82. doi:10.1186/gb-2014-15-5-r82
1283 Wang, H., An, W., Cao, R., Xia, L., Erdjument-Bromage, H., Chatton, B., . . . Zhang, Y. (2003). mAM
1284 facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause
1285 transcriptional repression. *Mol Cell, 12*(2), 475-487. doi:10.1016/j.molcel.2003.08.007

1286 Zhang, Y., Bilbao, A., Bruderer, T., Luban, J., Strambio-De-Castillia, C., Lisacek, F., . . . Varesio, E.
1287 (2015). The Use of Variable Q1 Isolation Windows Improves Selectivity in LC-SWATH-MS
1288 Acquisition. *J Proteome Res, 14*(10), 4359-4371. doi:10.1021/acs.jproteome.5b00543
1289 Zhimulev, I. F., Belyaeva, E. S., Semeshin, V. F., Koryakov, D. E., Demakov, S. A., Demakova, O. V., .
1290 . . Andreyeva, E. N. (2004). Polytene chromosomes: 70 years of genetic research. *Int Rev Cytol, 241*,
1291 203-275. doi:10.1016/S0074-7696(04)41004-3
1292 Zhu, X., Chen, Y., & Subramanian, R. (2014). Comparison of information-dependent acquisition,
1293 SWATH, and MS(All) techniques in metabolite identification study employing ultrahigh-
1294 performance liquid chromatography-quadrupole time-of-flight mass spectrometry. *Anal Chem,*
1295 *86*(2), 1202-1209. doi:10.1021/ac403385y
1296 Zielke, N., Edgar, B. A., & DePamphilis, M. L. (2013). Endoreplication. *Cold Spring Harb Perspect*
1297 *Biol, 5*(1), a012948. doi:10.1101/cshperspect.a012948
1298