SUPPLEMENTAL INFORMATION

Regulatory Functions and Chromatin Loading Dynamics of Linker Histone H1 during Endoreplication in *Drosophila*

Evgeniya N. Andreyeva, Travis J. Bernardo, Tatyana D. Kolesnikova, Xingwu Lu, Lyubov A. Yarinich, Boris A. Bartholdy, Xiaohan Guo, Olga V. Posukh, Sean Healton, Michael A. Willcockson, Alexey V. Pindyurin, Igor F. Zhimulev, Arthur I. Skoultchi, and Dmitry V. Fyodorov

Supplemental Materials and Methods

Fly genetics and H1 knockdown in vivo

Flies were maintained on standard corn meal, sugar and yeast medium with Tegosept at 18-29°C as indicated. *His1* RNAi allele *pINT-H1^{5M}* was described in (Lu et al. 2009), *SuUR^{ES}* was described in (Belyaeva et al. 1998). *UAS-His1-dsRNA-10-3* stock (Siriaco et al. 2009) was a generous gift of Giorgia Siriaco (UC Santa Cruz). *da-GAL4* and *Tub-GAL4* drivers and various balancer chromosomes were obtained from the Bloomington Stock Center. For a strong depletion of H1 protein in salivary glands (to <5% normal level), we used heterozygous *pINT-H1^{5M}* RNAi allele in combination with heterozygous *Tub-GAL4* driver at 29°C (**Fig. 1I**) as in (Lu et al. 2009); *pINT-Nau/+; Tub-GAL4/+* animals were used as controls. For moderate to strong depletion of H1 (to ~15% normal level), the same RNAi-driver allele combination was used at 25°C (**Fig. 1A-F**, **Supplemental Fig. S1A-B**,**H** and **Supplemental Tables S1**, **S2**); *pINT-Nau/+; Tub-GAL4/+* animals were used as controls. For moderate depletion of H1 (to ~30% normal level), the following RNAi-driver allele combination was used: *UAS-His1-dsRNA-10-3/+* and *da-GAL4/+* at 25°C (**Figs. 1J, 2C**,**D**, **Supplemental Figs. S2B**, **S6A** and **Table 1**) or 29°C (**Fig. 2E**); *UAS-His1-dsRNA-10-3*, *da-GAL4* or wild-type (Oregon R) animals were used as controls. In several control experiments, a moderate depletion of H1 was produced by a combination of *pINT-H1^{2M}* RNAi allele and *Tub-GAL4* driver at 29°C, and the results were similar (*data not shown*).

Genomic DNA for high throughput sequencing

Freshly dissected salivary glands from wandering L3 larvae (300 pairs for the control, 600 pairs for H1 knockdown) were cross-linked in 1 % formaldehyde for 15 min at room temperature. Glycine was added to the final concentration of 125 mM, the glands were pelleted by centrifugation (10 min, 470 *g* at room temperature), washed with 1x PBS and stored at -80°C.

The glands were resuspended in 300 µl room-temperature lysis buffer (50 mM Tris–HCl, pH 8.0, 10 mM EDTA, 1% SDS, Sigma P8340 protease inhibitor mix diluted 1:100, 1 mM PMSF, 20 mM Na-butyrate). Cells were lysed by Dounce homogenization (tight pestle) and kept on ice for 5 min. The samples were sonicated to shear chromatin to an average DNA fragment size of 200 bp using a Covaris sonicator. The lysates were transferred to microcentrifuge tubes and centrifuged (10 min, 12,000 *g* at 4°C). 1.6 ml RIPA ChIP buffer (10 mM Tris–HCl, pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate, Sigma P8340 protease inhibitor mix diluted 1:100, 1 mM PMSF, 20 mM Na-butyrate) was added to chromatin supernatants, and the mixture was centrifuged (10 min, 12,000 *g* at 4°C). The chromatin supernatants were stored at -80°C in 200-µl aliquots. 200 µl chromatin sample was mixed with 400 µl elution buffer (20 mM Tris–HCl, pH 7.5, 5 mM EDTA, 50 mM NaCl, 20 mM Na-butyrate, 1% SDS, 50 µg/ml proteinase K), and DNA cross-link reversal and protein digestion was performed at 68°C for 2 h. DNA was isolated by phenol-chloroform-isoamyalcohol extraction and ethanol precipitation with acrylamide carrier (0.0025% wt/vol). Sequencing libraries were prepared using the Illumina TruSeq ChIP Sample Prep Kit following the manufacturer's protocol.

Subsampling analyses of UR regions

To rule out any potential effects of sequencing depth, the analysis was repeated using a sub-sampled control with approximately equivalent uniquely aligning reads as the H1 knockdown. The identified UR regions were highly similar between the two sampling methods in the control knockdown (89.1% of UR nucleotides identified in fully sampled control were identified as UR in subsampled control). In addition, no appreciable differences were observed with respect to the fold change of copy number at UR regions in the H1 knockdown when compared to the control knockdown using either sampling method (*not shown*).

Validation of underreplication by quantitative PCR

Quantitative PCR was performed on genomic DNA prepared from salivary glands of L3 larvae of the following genotypes: *yw*; $SuUR^{ES}$ (Belyaeva et al. 1998), *pINT-Nau/+; Tub-GAL4/+* (Wei et al. 2007) and *pINT-*$H1^{5M}$*/+; Tubulin-GAL4/+* reared at 29°C (Lu et al. 2009). Primer sequences are provided in **Supplemental Table S3**. Approximately 40 pairs of salivary glands were dissected and collected in 1x PBS on ice. The glands were pelleted at 100 *g* for 3 min, and genomic DNA was purified using the *Quick-gDNA* MiniPrep kit (Zymo) following the manufacturer's protocol. Genomic DNA from non-polytenized, 0-12 h *yw* embryos was purified using the same protocol. For each genotype, at least two independent biological collections were performed. Real-time PCR was performed using 1 ng genomic DNA on a ViiA7 thermocycler (Applied Biosystems) with a three-step protocol (95°C 15 sec, 60°C 30 sec, 68°C 60 sec) and iTaq Universal SYBR Green Supermix (Bio-Rad). For each amplicon, the copy number in salivary gland DNA was calculated relative to embryonic DNA and normalized to the copy number of a random genomic sequence as a loading control. To avoid bias, we selected a random 118 bp interval from an intergenic region in cytological region 86D9 of chromosome 3R (*intergenic86D9*, **Supplemental Table S3**) that exhibited genome-average copy numbers in our sequence data for both control and H1-depleted salivary glands. The following ΔΔCt formula was used: $(Ct_{sg} - Ct_{embryo})_{target} - (Ct_{sg} - Ct_{embryo})_{intergenic86D}$. Alternatively, qPCR with primers from the rp49/*RpL32* locus (**Supplemental Table S3**) was used as normalization/loading control, and the results obtained were very similar (**Fig. 1H**,**I** and *data not shown*).

Recombinant proteins

To generate pGEX-4T-SUUR(1-370) and pGEX-4T-SUUR(579-962), the corresponding fragment of cDNA was amplified by PCR from a *Drosophila* 0-12 h embryonic library (Canton-S) with the following primers:

BamHI-1-370 forward, AAAAGGATCCATGTATCACTTTGTATCCGAG;

SalI-1-370 reverse, AAAAGTCGACTCATGATTCATTTAATGAGGGCG;

579-962 forward, GAAGCACTTGCCACCTACTC and;

XhoI-579-962 reverse, AAAACTCGAGTCACTTGAACAGTTCCAATC.

PCR products were digested with BamHI and SalI or XhoI (restriction sites are underlined) and cloned into BamHI-SalI- or SmaI-XhoI-digested pGEX-4T-1 vector. All constructs were verified by DNA sequencing. The expression of GST fusion proteins was induced in *E. coli*, BL21(DE3)pLysS (Promega), with 0.2 mM IPTG for 2 h at 25°C. The cells were sonicated in lysis buffer that contained HEMG (25 mM HEPES-KOH, pH 7.6, 0.1 mM EDTA, 12.5 mM $MgCl_2$, 10% glycerol), 800 mM NaCl, 1 mM DTT, 1 mM $Na_2S_2O_5$, 2 μM PMSF and 0.1% NP-40. The lysis buffer was supplemented with 150 μg/ml lyzozyme. The lysates were centrifuged at 15,000 *g* for 20 min, loaded on glutathione Sepharose 4FF (GE Healthcare) and washed three times with the lysis buffer and three times with the storage buffer (HEMG with 150 mM NaCl, 100 mM KCl, 0.01% Triton X-100 and 2 μM PMSF). Aliquots of the beads were boiled in Laemli buffer, and protein loading was analyzed by SDS-PAGE and Coomassie staining. The beads were resuspended in the storage buffer to an approximately equimolar concentration for all GST fusion polypeptides.

To prepare pGEX-4T-PCNA plasmid encoding GST-tagged full-length *Drosophila* PCNA, the corresponding fragment of cDNA was amplified by PCR with the following primers:

EcoRI-PCNA forward, AAAGAATTCATGTTCGAGGCACGCCTG and;

XhoI-PCNA reverse, AAACTCGAGTTATGTCTCGTTGTCCTCGATCTTGG.

PCR products were digested with EcoRI and XhoI (restriction sites are underlined) and cloned into EcoRI-XhoI-digested pGEX-4T-1 vector. All constructs were verified by DNA sequencing. The GST fusion was expressed as described above and eluted with 10 mM reduced

glutathione in 50 mM Tris-HCl, pH 8.0 and dialyzed into PBS. The purified protein was analyzed and quantified by SDS-PAGE along with BSA mass standard (Pierce).

Antibodies and immunoblot analyses

For immunoblots, salivary glands from L3 larvae of a particular genotype were boiled in SDS-PAGE sample buffer, and material equivalent to eight pairs of glands was loaded in each lane. Rabbit polyclonal anti-SUUR (E-45) (Makunin et al. 2002) antibody was used at 1:500, rabbit polyclonal anti-H1 (Lu et al. 2009) was used at 1:50,000, and mouse monoclonal anti-ß-tubulin Bx69 (a generous gift of Harold Saumweber, Humbolt University) was used at 1:800. The following secondary antibodies were used: goat anti-rabbit HRP, goat anti-mouse HRP (both Invitrogen, 1:3,500), and goat anti-Guinea pig HRP (Sigma, 1:5,000).

For IF, Guinea pig polyclonal anti-SUUR #457 (Nordman et al. 2014) antibody was used at 1:80, mouse monoclonal anti-PCNA (PC10, Abcam) was used at 1:500, mouse monoclonal anti-H3K9me2 (Abcam) was used at 1:800, and rabbit polyclonal anti-H1 (Lu et al. 2009) was used at 1:5,000. For co-staining of H3K9me2. SUUR and PCNA (Supplemental Fig. S6A) a polyclonal antibody to PCNA was raised in rat with GST-PCNA fusion protein (see above) as an antigen and used at 1:500. The following secondary antibodies were used: highly cross-absorbed goat anti-rabbit Alexa 488, highly cross-absorbed goat anti-rabbit Alexa 568, highly cross-absorbed goat anti-Guinea pig Alexa 488, highly cross-absorbed goat anti-Guinea pig Alexa 568, highly cross-absorbed goat anti-mouse Alexa 568 and cross-absorbed goat anti-rat Alexa 488 (all Invitrogen, 1:600); donkey anti-rabbit Alexa 647, goat anti-rabbit FITC and goat anti-mouse Texas Red, 1:200 (all Abcam, 1:200).

GST pulldown assays

Extracts from adult *Drosophila* ovaries were prepared in Buffer O (25 mM HEPES-KOH, pH 7.6, 1 mM EDTA, 100 mM KCl, 12.5 mM $MgCl_2$, 0.01% Triton X-100 and Halt™ Protease and

Phosphatase Inhibitor Cocktail, ThermoFisher Scientific) + 450 mM NaCl. The extracts were centrifuged at 15,000 *g* for 15 min and diluted to Buffer O + 150 mM NaCl. Protein-loaded glutathione beads (equivalent to ~20 pmoles of GST fusions) were incubated with ovarian extracts (equivalent to 5 ovaries per reaction) for 3 h at 4°C on a rotating wheel. The beads were washed three times for 15 min with Buffer O, twice for 15 min with Buffer O + 200 mM NaCl, and once for 15 min in Buffer O. (Each wash step was followed by centrifugation at 1,000 g for 2 min.) The wash buffer was removed, the beads were boiled in Laemli buffer for 5 min, the samples were divided in two halves, and the proteins were analyzed by SDS-PAGE and Coomassie staining or immunoblotting.

For GST pulldowns with recombinant H1 polypeptides, binding reactions were performed with ~20 pmoles GST fusion proteins pre-bound to the beads and ~20 pmoles H1 polypeptides in 50 µl of Buffer O + 150 mM NaCl for 3 h at 4°C on a rotating wheel. Unbound proteins were removed in a series of washes, and resin-bound proteins were analyzed as described above.

Supplemental References

Siriaco G, Deuring R, Chioda M, Becker PB, Tamkun JW. 2009. *Drosophila* ISWI regulates the association of histone H1 with interphase chromosomes in vivo. *Genetics* **182**: 661-669.

Supplemental Tables

Supplemental Table S1. DNA underreplication in salivary glands is suppressed across UR domains upon depletion of H1.

Single-end reads were aligned to the *Drosophila* genome using Bowtie (Langmead et al. 2009). Coordinates refer to Berkeley *Drosophila* Genome Project (BDGP) R5/dm3 genome assembly. While scanning in 5-kb windows, UR regions were identified in the control knockdown as regions with 20 consecutive windows below average per-chromosome read count. Fold change in normalized read count between H1 and control knockdowns is shown for each of the derived UR regions (and the central 25 kilobases of each UR region, as indicated) using HOMER annotatePeaks.pl (Heinz et al. 2010). ≥1.20-fold increase in copy number is highlighted in bold. Approximate corresponding cytological region from UCSC Genome Browser is shown for each identified UR region. "–", no polytene structure.

Intercalary heterochromatin (IH)

Genomic coordinates of UR region	~Cytological band	Fold change, copy number H1 KD / NAU KD	
		[Entire UR region]	[Central 25 kb]
X: 2830001-2965000	3C	**1.23**	**1.21**
X: 8675001-8785000	8B	0.99	0.96
X: 11925001-12275000	11A	**1.57**	**2.03**
X: 14170001-14475000	12E	**1.24**	**1.48**
X: 20405001-20535000	19E	1.08	1.03
X: 20540001-20885000	19E	**1.30**	1.18
X: 21995001-22190000	20DE	**1.31**	1.15
2L: 3920001-4040000	24D	0.98	1.16
2L: 4620001-4745000	25A	**1.24**	1.15
2L: 11330001-11470000	33A	1.08	**1.30**
2L: 11575001-11775000	33A	**1.38**	**1.66**

2L: 12770001-12890000	34A	**1.20**	**1.38**
2L: 14710001-14965000	35B	**1.64**	**2.34**
2L; 15295001-15755000	35D	1.17	**1.23**
2L: 15945001-16220000	35E	**1.87**	**2.76**
2L: 16935001-17365000	36C	**1.87**	**2.04**
2L: 17530001-17970000	36D	**1.91**	**2.52**
2L: 21400001-21555000	39DE	**1.92**	**2.38**
2R: 2145001-2390000	42B	**1.20**	**1.23**
2R: 16255001-16400000	57A	1.16	**1.24**
2R: 19035001-19205000	59D	**1.32**	**1.22**
3L: 4865001-5075000	64C	**1.48**	**1.58**
3L: 5390001-5505000	64D	1.18	**1.36**
3L: 6285001-6460000	65B	**1.23**	**1.46**
3L: 6785001-6905000	65D	1.07	1.13
3L: 10000001-10180000	67D	**1.24**	**1.38**
3L: 13555001-13835000	70C	**1.76**	**1.98**
3L: 15190001-15465000	71C	**1.49**	**1.95**
3L: 18175001-18460000	75C	**1.51**	**2.20**
3R: 1930001-2135000	83E	**1.26**	**1.28**
3R: 2340001-2455000	84A	1.16	1.15
3R: 3375001-3550000	84D	**1.47**	**1.71**
3R: 6275001-6475000	86C	1.15	1.17
3R:6735001-6980000	86D	**1.45**	**1.74**
3R:12590001-12775000	89E	**1.62**	**1.72**
3R: 12995001-13160000	90A	**1.25**	**1.43**
3R: 15965001-16085000	92D	1.03	1.05
3R: 17940001-18095000	94A	1.14	1.15
3R: 23835001-24125000	98C	**1.45**	**1.84**

Pericentric heterochromatin (PH)

Genomic coordinates of UR region	~Cytological band	Fold change, H1 KD / NAU KD	
		[Entire UR region]	[Central 25 kb]
2L: 21850001-22075000	40A-E	**1.46**	**1.91**
2L: 22285001-22420000	40F	**2.09**	**2.56**
2L: 22485001-22685000	–	**1.31**	**1.61**
2R: 720001-965000	41CD	**1.25**	**1.34**
2R: 1250001-1490000	41EF	0.96	0.94
3L: 23345001-23705000	80F	**1.21**	**1.36**
3L: 23745001-23935000	–	**1.65**	**1.92**

Supplemental Table S2. DNA underreplication in salivary glands is suppressed at transposable element sequences across the genome upon depletion of H1.

Names and sequences of TE's were obtained from the canonical transposon sequence set, BDGP version 9.4.1. Single-end reads were aligned to the dm3 release with repeat sequences masked by RepeatMasker, combined with the sequences of the canonical transposon sequence set. Shown is the fold change in copy number for each TE sequence between *SuUR*^{ES} and wild-type genotypes (Yarosh and Spradling 2014) and between H1 and control knockdown. ≥1.20-fold increases in copy number are highlighted in bold.

Transposable element	Fold change, copy number	
	SuUR^{ES} / wild type	H1 KD / NAU KD
R2-element	**5.42**	**1.73**
ZAM	**5.31**	**1.67**
Helitron	**4.96**	0.94
gypsy11	**4.35**	**1.38**
3S18	**4.24**	**1.36**
TART-C	**4.24**	**2.08**
invader6	**4.19**	**1.55**
gypsy9	**4.18**	**1.87**
GATE	**4.13**	**1.68**
G6	**3.66**	**1.24**
Osvaldo	**3.61**	1.16
gypsy6	**3.39**	**1.30**
Porto1	**3.14**	**1.31**
Circe	**3.10**	**1.63**
aurora-element	**3.05**	**2.27**
Transpac	**2.95**	**1.20**
frogger	**2.93**	**1.28**
gypsy	**2.80**	**1.36**
McClintock	**2.71**	1.01
opus	**2.67**	**1.20**
TAHRE	**2.60**	0.59
HeT-A	**2.52**	0.64

G2	**2.39**	1.15
Stalker3T	**2.33**	**1.57**
G5A	**2.25**	1.19
rover	**2.18**	0.92
gypsy7	**2.17**	**1.31**
G7	**2.14**	**1.44**
Quasimodo	**2.07**	**1.26**
gypsy3	**2.04**	**1.22**
Doc3-element	**2.03**	**1.46**
invader1	**2.01**	**1.23**
Max-element	**1.96**	**1.52**
R1A1-element	**1.90**	**1.87**
Tabor	**1.88**	1.11
BS3	**1.84**	**1.26**
1731	**1.83**	**1.20**
accord2	**1.82**	**1.35**
G5	**1.82**	1.13
Juan	**1.80**	1.05
baggins	**1.79**	**1.37**
HMS-Beagle2	**1.79**	1.03
gypsy4	**1.78**	**1.23**
hopper	**1.77**	1.02
Ivk	**1.77**	1.00
looper1	**1.77**	**1.34**
invader5	**1.73**	**1.44**
Rt1b	**1.72**	**1.31**
gypsy2	**1.71**	**1.22**
G3	**1.68**	**1.25**
Stalker4	**1.66**	**1.79**
invader2	**1.65**	**1.35**
Stalker	**1.65**	**1.64**
Stalker2	**1.65**	1.08
gypsy10	**1.62**	**1.25**
Cr1a	**1.61**	**1.22**
FB	**1.61**	0.69
Fw2	**1.61**	**1.36**
gypsy12	**1.61**	**1.31**
Rt1a	**1.61**	**1.46**

transib2	**1.61**	1.10
I-element	**1.60**	1.13
mdg1	**1.60**	0.97
hopper2	**1.58**	0.92
S-element	**1.57**	1.05
transib1	**1.57**	**1.41**
diver2	**1.54**	**1.20**
NOF	**1.54**	0.15
G-element	**1.53**	**1.98**
Tom1	**1.52**	0.96
Fw3	**1.51**	1.06
gypsy8	**1.51**	**1.30**
Helena	**1.50**	**1.40**
Q-element	**1.50**	1.06
gypsy5	**1.45**	**1.28**
gtwin	**1.43**	0.81
R1-2	**1.41**	**1.58**
mdg3	**1.38**	0.92
blood	**1.37**	**1.21**
diver	**1.36**	0.86
transib3	**1.34**	**1.29**
Dm88	**1.33**	1.11
Rt1c	**1.32**	**1.56**
jockey	**1.31**	0.92
1360	**1.30**	1.15
Doc4-element	**1.30**	**1.30**
flea	**1.28**	1.15
pogo	**1.28**	**1.23**
G4	**1.27**	1.15
invader4	**1.26**	0.99
412	**1.25**	0.92
INE-1	**1.25**	1.19
Doc2-element	**1.22**	**1.20**
accord	**1.21**	**2.11**
Tc1-2	**1.21**	0.99
rooA	**1.20**	0.89
Tc1	1.19	1.04
jockey2	1.18	1.09

297	1.17	0.88
Burdock	1.14	1.13
mariner2	1.14	1.09
X-element	1.14	**1.24**
Bari1	1.13	**1.49**
BS	1.13	**1.20**
S2	1.10	1.15
BS4	1.09	0.95
17.6	1.07	**1.52**
roo	1.06	0.93
Bari2	1.05	1.01
HB	1.05	0.95
HMS-Beagle	1.00	0.95
transib4	1.00	0.98
Tc3	0.95	**1.48**
micropia	0.82	**1.31**
Doc	0.76	1.12
springer	0.72	0.87
F-element	0.71	1.11
TART-B	0.67	0.97
copia	0.63	0.91
invader3	0.61	**1.27**
Idefix	0.44	**1.31**
TART-A	0.10	0.64
Tirant	0.10	**2.13**
hobo	0.08	0.68

Supplemental Table S3. Primer sequences used for qPCR validation.

Genomic coordinates indicate the full amplicon, including the length of each primer. Coordinates refer to the BDGP R5/dm3 assembly.

75C (Chromosome 3L)

Genomic coordinates, 3L:	Primer Sequences
18055206–18055330	ATTTGGACTGGGGCAGTTTC
	CTGAAACACGGAAGTTGAGTCC
18080166–18080275	AAAACACAAGCACATAGGCAAC
	AGTTTCTGGCGTTGTATCCG
18132092–18132212	TGGCAATAATTCCCCAACCG
	TGATTAAGGCGAACACAGCAC
18157527–18157599	GTGCACGGACGCGTATAATC
	AAGTTAGCTCACGTGAGATGATG
18181945–18182048	ACTATTATTTCTGGCTGGCTACG
	GCCGGCTGCTACTTATGGC
18207203-18207310	ATACAGATACAGCTCGCACTGG
	AGTGGTGCCGATGGAAAAAC
18232035–18232121	ACCACGCCCCTAAGCAAATAG
	ATCTCGCCAGCTAAAGATCTCG
18257054–18257143	TGGGGCATTTTTGACGGTAG
	GCTTTTAGCCTCGAGAAACCG
18307014–1830712	CTTGGCTCAGGTTTCCCTTC
	AAAGGACGCCACAACAATGC
18332011–18332146	ATCTCTCTGGGGCATCCAAG
	CGCCAGCGCAGTTAAAAGTAAC
18357190–18357332	TGCACCAAGCTACACAATGG
	CACAGGACTCCAAATTCTGCAC
18407206–18407287	AGTGATAGCGGAGTAACAGTGG
	GTGGCGTGGATCCAACTTTATG
18432517–18432589	TGCGCTAGTTCTCACCAACG
	ACCAACTTAAGCACCAACTAAGG
18457396–18457476	ACGGGTGCCCTTAATGTTTAC
	GGTCGTTGCCCATGTCTTTG
18485078–18485157	CAACCCTATCCATCCATCCATG
	CAATCGGCCTAATTCACCCATG
18513643–18513780	ACATATTCGCCGACCAAGTG
	ACACTAACACGTGCCCCTAAC

Intergenic86D9 (Chromosome 3R)

Genomic coordinates, 3R:	Primer Sequences
7087055–7087172	TGGCGCCGCTTTCTTATTAG
	AGAACAGGTTTGTGCGCTTG

89E (Chromosome 3R)

Genomic coordinates, 3R:	Primer Sequences
12420110–12420186	CAAACCCGAGCATAGCACAG
	GTGCTTTGGCGCCAGATATG
12469885–12469990	ACCTGGACTTGTCGCTGATTAC
	TTGACTCCGCCAAAGATCAGG
12502544–12502639	TTGTTTTGGCCGGTTGTAGC
	CGTTTTACTCACCCGGCAAAC
12527301–12527377	CTGCAGAGTTCTCACATGTTGG
	AAGGAGACACGACCACGAC
12552419–12552553	AGATCTCCGGGAAAATGCTCTC
	TCCATGCCACATGCTCTTGG
12577376–12577453	ACAGCACGAGCCTATTCCTATC
	TTATTCGACCCCACTGCCATC
12627469–12627554	TGCGTGGAATTTGGTGTGAC
	ATGCATAGTGGTCGGATAGCTC
12652387–12652466	CGAGAGGAAAATGGACAGAGTC
	ACTGCAAAACCGCAAATCCG
12677606–12677752	GCCCGGATAGACAGATTTTACG
	AGCGAACGACCTCTGAATTC
12702428–12702497	GCAAATTGCAAGCGAAGTCG
	CCAGCAGGTGAATCGAGAATC
12727405–12727531	ATGCGCTTCAGAGCTTTTCG
	TGCTAGTCCCGAGCCAAATTAG
12777316–12777412	TATTGCCCTCCGTTGTCGATG
	TAAGGGCGAAAAGGGTTCTCC
12802543–12802640	ACCCATAAGAGTTTCCCGTTCC
	TGCTAGTAGGGCGGAGAAAAG
12835126–12835239	TCTCAGGATGCCACAAAACG
	TTAAAGGTGGCCAAGTGCTC
12859859–12859957	GCGAAAAACCTCCTTCGTTCTC
	TCGGACCAACTTTGTTGTGTG
12884868–12884947	ATACTCATGCCCACCACTGTAG
	CCGAATACATGCTCACCACAC

RpL32 (99D3, Chromosome 3R)

Genomic coordinates, 3R:	Primer Sequences
25871311–25871427	TCTGTTGTCGATACCCTTGG
	TCGATATGCTAAGCTGTCGG

Supplemental Figure Legends

Supplemental Figure S1. Genome-wide analyses of DNA underreplication in H1 knockdown and *SuUR*^{ES} salivary glands.

(*A*) DNA copy numbers (normalized to chromosome arm average) for control (NAU KD, gray) are shown across the entire mapped *Drosophila* genome (cf. **Fig. 1A**). Blue boxes, UR regions identified in this study; red boxes, UR regions identified by microarray analyses (Sher et al. 2012); green boxes, UR regions identified by high throughput sequencing (Yarosh and Spradling 2014); asterisks, UR regions examined cytologically for the frequency of chromosome break points (**Table 1**); double asterisks, UR regions studied cytologically and by qPCR (**Fig. 1H**,**I** and *not shown*); X, 2L, 2R, 3L and 3R, chromosome arms; genomic coordinates (Mbp) are indicated at the bottom.

(*B*) Visual representation of calculations performed for this study. A UR region (black horizontal bar) is called when at least twenty consecutive 5-kb windows fell below the average per-chromosome read count. Normalized reads for each UR region were calculated as the total aligned reads per million divided by the length of the UR domain (kb). Normalized reads for H1 (black, in the background) and control (semi-transparent light gray, in the foreground) knockdowns are shown in the genomic locus flanking cytological region 75C in 3L. Genomic coordinates (Mbp) are indicated at the bottom.

(*C*) Suppression of underreplication in *SuUR*^{ES} salivary gland cells (cf. **Fig. 1C**). For each identified UR region, reads (normalized by the region length and total read count) in homozygous *SuUR*^{ES} mutants (*y*-axis) are plotted against reads in wild type (*y; cn bw sp*, *x*-axis). The dotted line represents equal DNA copy numbers for both conditions.

(*D*) Average extent of underreplication and its suppression by *SuUR*^{ES} mutation across UR regions (cf. **Fig. 1D**). Average read counts (normalized to total read count) were calculated

across all identified UR regions for wild-type and *SuUR*^{ES} mutants as indicated. Distance from the UR region center (kb) is indicated on the *x*-axis.

(*E*) Dependence of the extent of underreplication on the UR region length (cf. **Fig. 1E**). For each UR region, normalized read counts in wild type (Yarosh and Spradling 2014) (*y*-axis) are plotted against the length of the region (*x*-axis).

(*F*) Dependence of the extent of *SuUR*-dependent suppression of underreplication on the UR region length. The $\log_2$ fold change of read counts in *SuUR*^{ES} mutants relative to wild type (*y*-axis) is plotted against the length of each UR region (*x*-axis).

(*G*) *SuUR*-dependent suppression of underreplication of transposable element (TE) loci. Reads were aligned to TE sequences. For each TE, normalized read counts in *SuUR*^{ES} mutant (*y*-axis) are plotted against normalized read counts in wild type (*x*-axis). The dotted line represents equal DNA copy number in both conditions. The axes are shown in $\log_{10}$ scale.

(*H*) H1 depletion-dependent suppression of underreplication of transposable element (TE) loci. Same as (*G*), except normalized read counts are plotted for H1 knockdown (H1 KD, *y*-axis) against control (NAU KD, *y*-axis). See *Materials and methods* for description of analyses and data accession.

Supplemental Figure S2. H1 is an upstream effector of polytene chromosome distribution of SUUR.

(*A*) Decreased abundance of SUUR protein in H1-depleted polytene chromosomes. Polytene chromosomes were prepared from control and H1-depleted salivary glands and stained with PCNA and SUUR antibodies as in **Fig. 2D**. The polytene spread corresponds to a cell in late endo-S phase. PhC (top panel), phase contrast image.

(*B*) Normal abundance and distribution of H1 in *SuUR* mutant polytene chromosomes. Polytene chromosomes were prepared from control and homozygous $SuUR^{ES}$ salivary glands as in (*A*) and stained with H1 and SUUR antibodies. No specific IF signal for SUUR is detectable, whereas H1 signal appears normal (cf. **Fig. 2A**,**B**). The polytene spread corresponds to a cell in late endo-S phase. DAPI staining shows the overall chromosome morphology.

Supplemental Figure S3. GST pulldown analyses of direct physical interactions between SUUR polypeptides and full-length recombinant *Drosophila* H1.

(*A*) Recombinant GST-SUUR fusion polypeptides. GST and GST fusion proteins (as in **Fig. 3A**,**B**) were expressed and purified from *E. coli*, incubated with purified recombinant H1 and analyzed by SDS-PAGE and Coomassie staining. Arrowheads, full length polypeptide products; molecular mass marker sizes (kDa) are shown on the left.

(*B*) SUUR-dependent GST pulldowns of purified recombinant H1. Full-length *Drosophila* H1 protein was expressed in *E. coli* and purified in two chromatographic steps (see *Materials and Methods*), incubated with GST fusion proteins (*A*), and pulldown products were analyzed by H1-specific immunoblot along with the 5% input control as in **Fig. 3C**. Recombinant H1 (arrowhead) strongly interacts with the full-length and the middle fragment (amino acids 371-578) of recombinant SUUR.

Supplemental Figure S4. H1 chromatin loading dynamics in larval polytene chromosomes during endo-S phase.

(*A*) Stage-dependent distribution of H1 during endo-S phase in wild-type distal polytene chromosome X. Detailed view of fragments of polytene chromosomes from (or similar to) those presented in **Fig. 4A**. DAPI staining shows the overall chromosome morphology and was used for an alignment of cytological positions. Red numbers at the top and corresponding cytological bands in grayscale panels are connected with red lines. ER, early replication (early endo-S phase); E-MR, early to mid-replication; MR, mid- replication; M-LR, mid- to late replication; LR, late replication; VLR, very late replication; NR2, no replication (after the completion of the endo-S phase). During early endo-S, H1 is strongly loaded into late-replicating cytological band 6A, but the signal intensity is decreased as the endo-S phase progresses. In contrast, the abundance of H1 is minimal during early endo-S at the early-replicating chromosome segment that spans from the telomere of X to the cytological band 3C (red brackets). However, it is increased by mid- to late endo-S. The pattern of H1 deposition also gradually changes during the course of each stage of the endo-S phase, as H1 is evicted during DNA replication and re-deposited after its completion in individual loci.

(*B*) Stage-dependent distribution of H1 during endo-S phase in wild-type proximal polytene chromosome arm 2R. Detailed view of fragments of polytene chromosomes from (or similar to) those presented in **Fig. 4A**. The rapid, massive loading of H1 during early endo-S is observed in late-replicating cytological bands 44F and 50A of the 2L arm but not in the early-replicating 44CD, where H1 is loaded only during late endo-S.

(*C*) Stage-dependent and mutually exclusive distribution of H1 and PCNA during late endo-S phase in *SuUR* mutant polytene chromosome arms 2L and 2R. Detailed view of a fragment of polytene spread from **Fig. 4C** (row 5 from the top). A split image is presented in the bottom panel; cytological regions 33A (2L) and 56F (2R) are indicated by white lines; white rectangle

indicates a fragment used for a blow-up image in (*D*). Cytological band 56F corresponds to a large (~500 kb) genomic DNA fragment, which is underreplicated in wild type (Yarosh and Spradling 2014). In *SuUR*^{ES} mutant salivary glands, its underreplication is suppressed, it achieves normal DNA copy numbers and assumes an extended polytene morphology; however, it continues to be replicated late in endo-S. PCNA and H1 costaining during late endo-S phase reveals that within 56F, the maximal abundance of replication factors (*e.g.*, PCNA) corresponds to the minimal occupancy of H1, and *vice versa*.

(*D*) Mutually exclusive distribution of H1 and PCNA in late-replicating regions during late endo-S phase in *SuUR* mutant proximal polytene chromosome arms 2L and 2R. A fragment of polytene spread from (*C*, white rectangle) is presented as high-resolution merged and split images; numbers and connected white lines indicate corresponding cytological regions.

Supplemental Figure S5. Biochemical model for the dynamics of chromosome loading of H1 and SUUR during DNA endoreplication.

Linker histone H1 (green circles), SUUR protein (blue circles) and replication factors, including clamp loader PCNA (red circles) exhibit distinct, highly dynamic patterns of association with chromatin during the endocycle in *Drosophila* salivary gland cells. Rectangles represent a schematic view of a polytene chromosome fragment: light gray rectangles, early-replicating chromosome segments corresponding to polytene interbands and light (thin) bands; black rectangles, chromosome segments corresponding to dark (thick) polytene bands that replicate from mid- to late endo-S. Letters at the top indicate chromosome segments replicating in the middle of endo-S phase (M), in late endo-S (L) and intercalary heterochromatin domains (IH) replicating very late (VL). Letters on the left designate stages of the endocycle: NR, non-replicating, equivalent to the G1 phase of dividing cells; ER, early endo-S phase; MR, mid-endo-S phase; LR, late endo-S phase; VLR, very late endo-S phase.

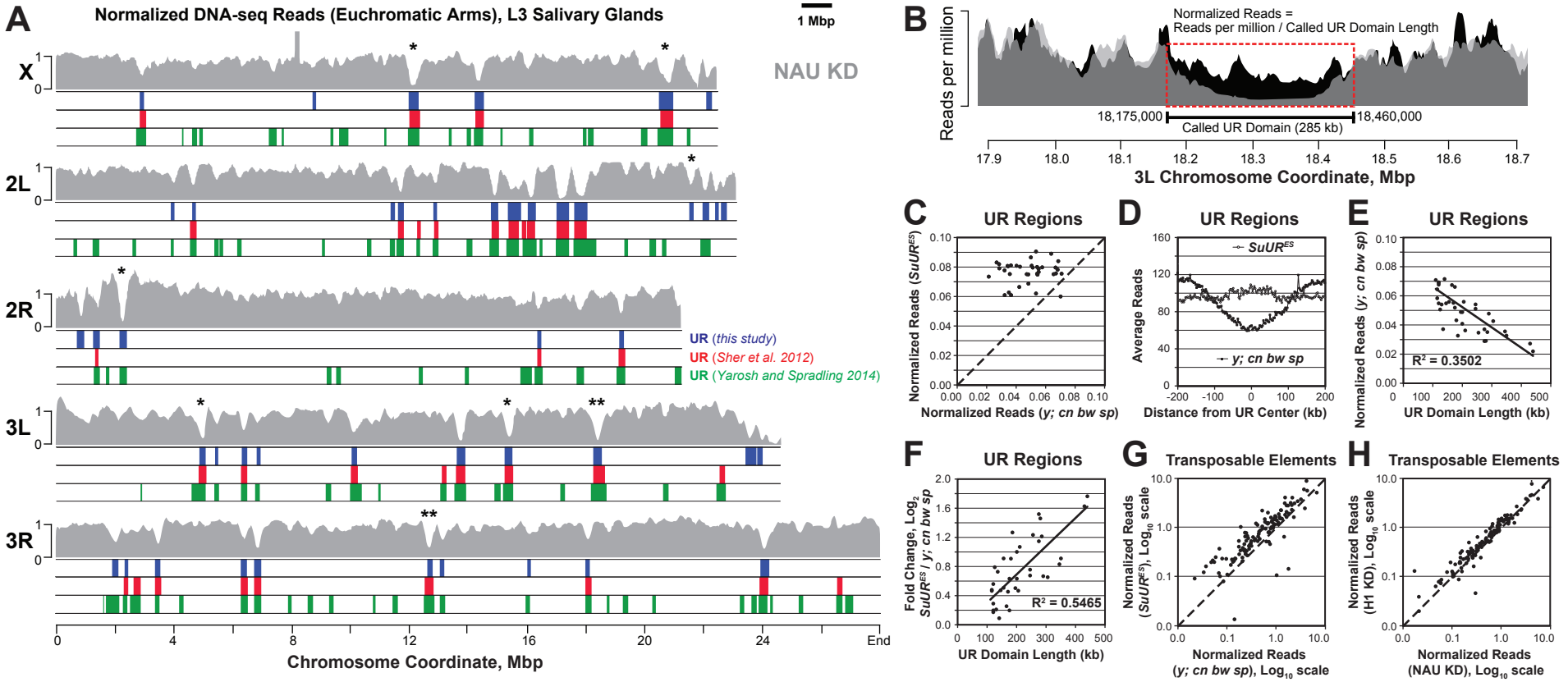
H1 is weakly present in most polytene bands before the onset of endo-S phase. Upon initiation of replication (manifested by loading of PCNA into early-replicating regions), H1 is rapidly and strongly loaded in VL/IH polytene bands through a RI, locus-specific deposition pathway (black open oval and black arrows). As replication progresses into L and VL polytene bands, H1 is removed from replicating loci via RD chromatin disassembly but it is slowly reloaded via RD chromatin assembly upon completion of endoreplication (red open ovals and arrows). The dynamic loading of SUUR into chromatin during the endocycle is nucleated by deposition into highly H1-enriched loci (VL/IH) during MR and LR (green open ovals and arrows). During LR and VLR, SUUR gradually increases its abundance in VL/IH chromosome segments (blue open ovals and arrows) through self-association (Kolesnikova et al. 2005) and by interactions with the factors present in replication forks (Kolesnikova et al. 2013).

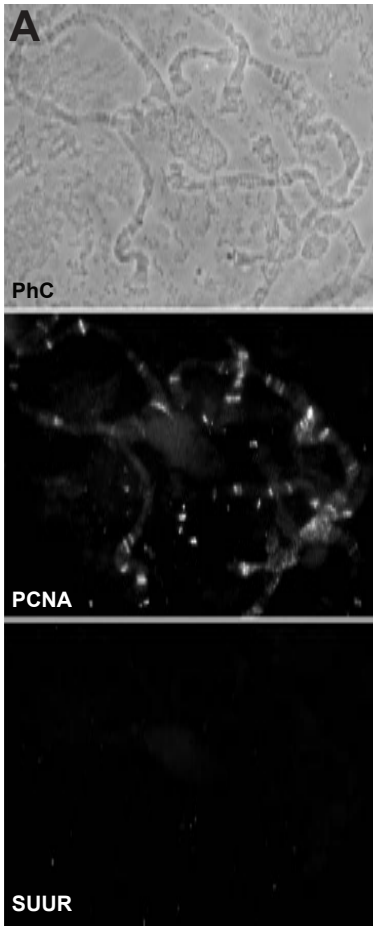
Supplemental Figure S6. H1-dependent SUUR localization in chromatin and possible SUUR-independent functions of H1 in regulation of endoreplication.

(*A*) Decreased abundance of SUUR protein in moderately H1-depleted polytene chromosomes is not paralleled by a comparable decrease of H3K9 dimethylation. Polytene chromosomes were prepared from control (*Oregon R*) and moderately H1-depleted (H1 KD) salivary glands as in (**Fig. 2C**) and stained with DAPI as wells as anti-H3K9me2, SUUR and PCNA antibodies. IF signal for SUUR is not detectable in H1 KD, whereas the signal for H3K9me2 persists in both the chromocenter and polytene arms. Polytene spreads before the onset of endo-S are presented (based on PCNA co-staining, *not shown*). Arroheads, chromocenters.

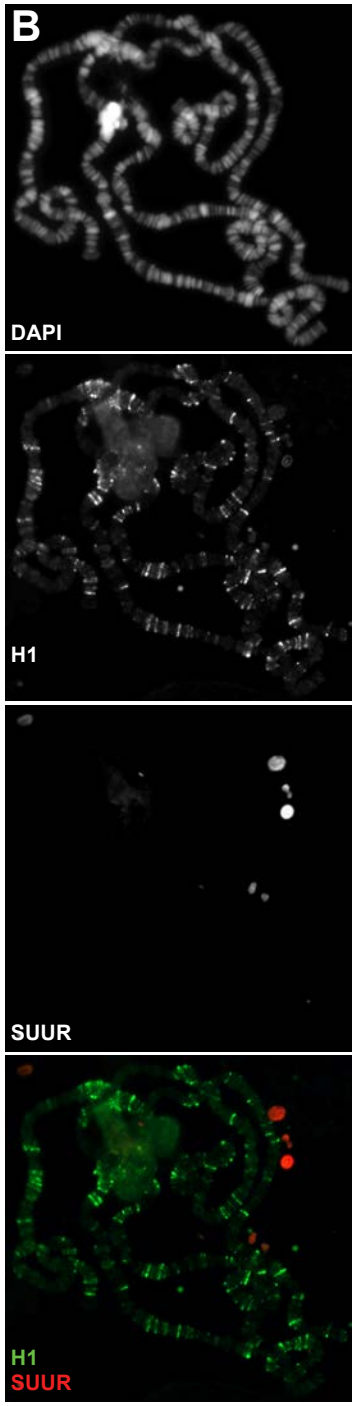
(*B*) Elevated loading of H1 in polytene chromosomes at the onset of endo-S phase in the absence of concomitant loading of SUUR. Polytene chromosome spreads from **Fig. 4B** were additionally stained with SUUR antibody. Two adjacent polytene spreads corresponding to "before" and early endo-S phase exhibit a dramatic difference in the level of H1 loading into chromosomes, whereas the abundance of SUUR remains uniformly low.

(*C*) Close-up views of DNA copy number (from high throughput sequencing, normalized to chromosome arm average) is shown for $SuUR^{ES}$ mutant (black, in the background) and wild-type control (*y; cn bw sp*, semi-transparent gray, in the foreground) (Yarosh and Spradling 2014). The numbers are normalized, and plots are presented exactly as in **Fig. 1G**.

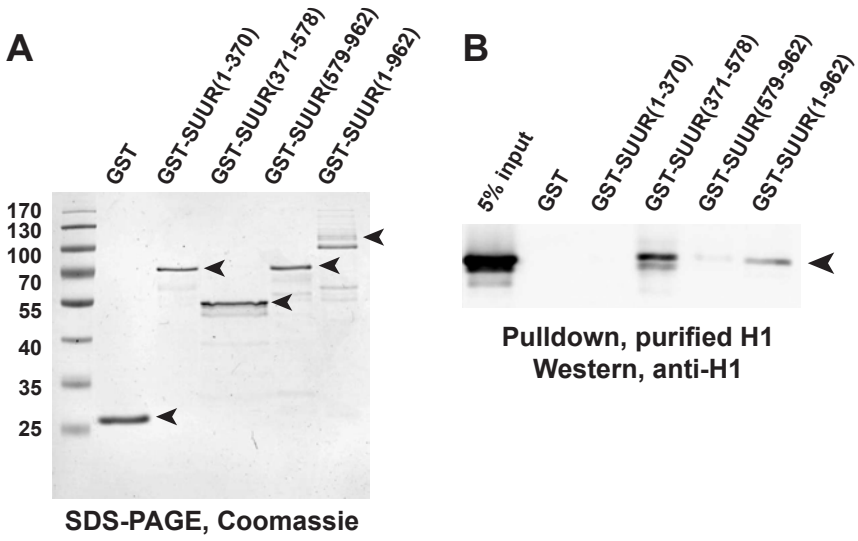

A
Normalized DNA-seq Reads (Euchromatic Arms), L3 Salivary Glands
1 Mbp
NAU KD
X
2L
2R
3L
3R
UR (this study)
UR (Sher et al. 2012)
UR (Yarosh and Spradling 2014)
0
4
8
12
16
20
24
End
Chromosome Coordinate, Mbp
B
Reads per million
Normalized Reads =
Reads per million / Called UR Domain Length
18,175,000
Called UR Domain (285 kb)
18,460,000
17.9
18.0
18.1
18.2
18.3
18.4
18.5
18.6
18.7
3L Chromosome Coordinate, Mbp
C
UR Regions
Normalized Reads (SuURES)
Normalized Reads (y; cn bw sp)
D
UR Regions
SuURES
y; cn bw sp
Average Reads
Distance from UR Center (kb)
E
UR Regions
Normalized Reads (y; cn bw sp)
R² = 0.3502
UR Domain Length (kb)
F
UR Regions
Fold Change, Log2 SuURES / y; cn bw sp
R² = 0.5465
UR Domain Length (kb)
G
Transposable Elements
Normalized Reads (SuURES), Log10 scale
Normalized Reads (y; cn bw sp), Log10 scale
H
Transposable Elements
Normalized Reads (H1 KD), Log10 scale
Normalized Reads (NAU KD), Log10 scale

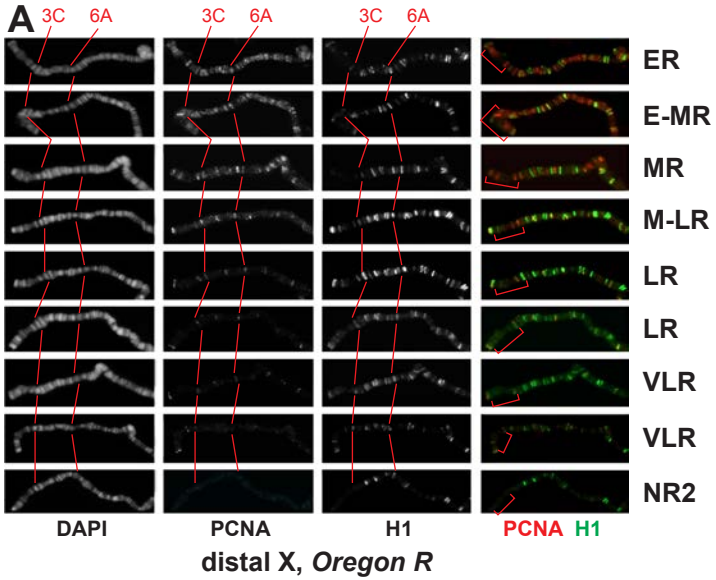


H1 KD, late endo-S

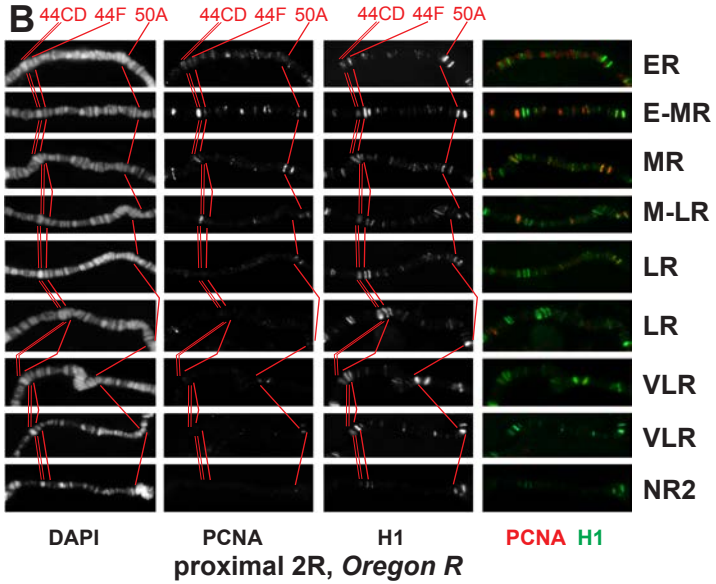


***SuUR*^{ES}, late endo-S**

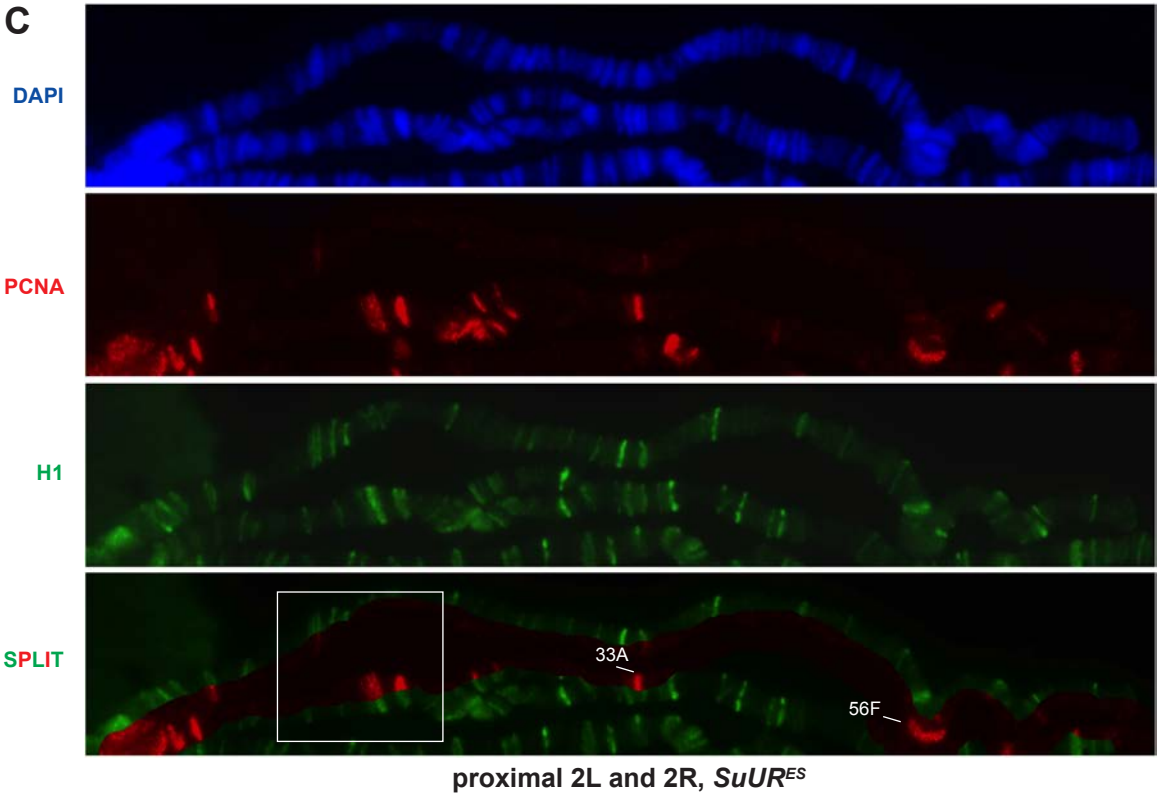

A
GST
GST-SUUR(1-370)
GST-SUUR(371-578)
GST-SUUR(579-962)
GST-SUUR(1-962)
170
130
100
70
55
40
35
25
SDS-PAGE, Coomassie
B
5% input
GST
GST-SUUR(1-370)
GST-SUUR(371-578)
GST-SUUR(579-962)
GST-SUUR(1-962)
Pulldown, purified H1
Western, anti-H1



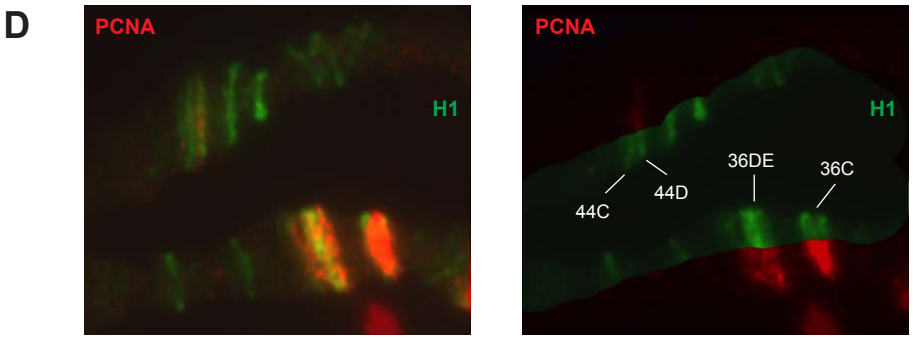
distal X, *Oregon R*



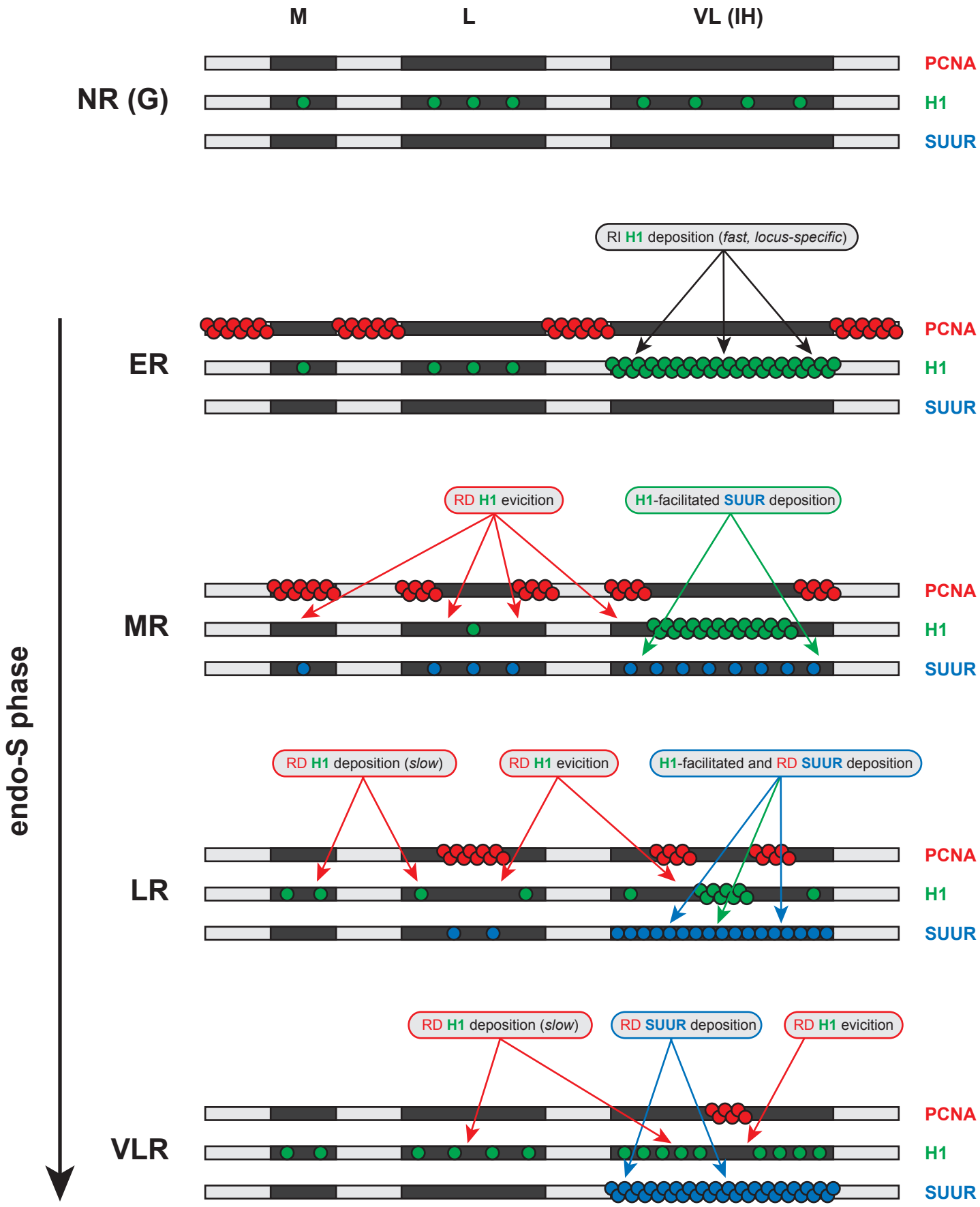
proximal 2R, *Oregon R*

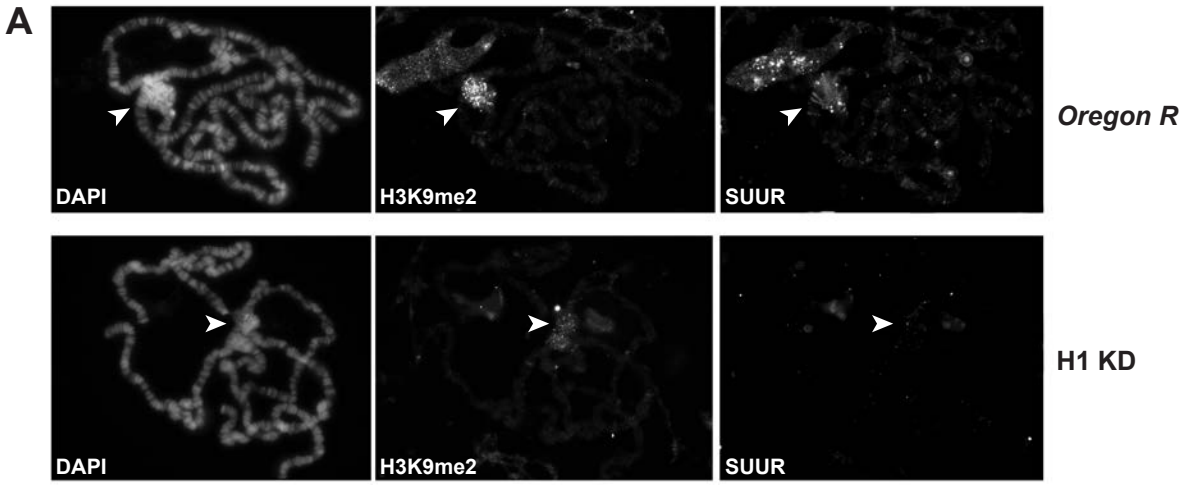


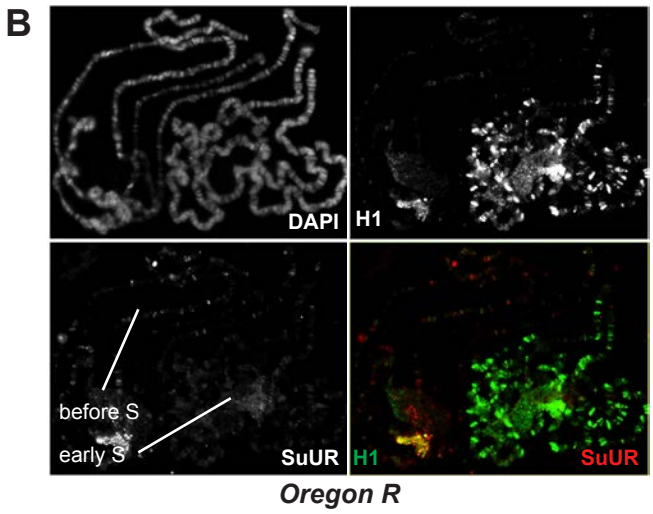
proximal 2L and 2R, *SuUR*^{ES}

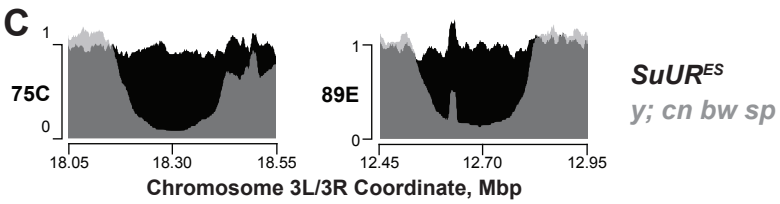


SuUR^{ES}


M
L
VL (IH)
NR (G)
PCNA
H1
SUUR
RI H1 deposition (fast, locus-specific)
ER
RD H1 evicition
H1-facilitated SUUR deposition
MR
RD H1 deposition (slow)
RD H1 evicition
H1-facilitated and RD SUUR deposition
LR
RD H1 deposition (slow)
RD SUUR deposition
RD H1 evicition
VLR
endo-S phase


A
DAPI
H3K9me2
SUUR
Oregon R
DAPI
H3K9me2
SUUR
H1 KD


B
DAPI
H1
before S
early S
SuUR
H1
SuUR
Oregon R


C
1
75C
0
18.05
18.30
18.55
1
89E
0
12.45
12.70
12.95
Chromosome 3L/3R Coordinate, Mbp
SuUR^ES
y; cn bw sp