Systematic identification of fragile sites via genome-wide location analysis of γ-H2AX

Rachel K. Szilard[1,2], Pierre-Étienne Jacques[1,3], Louise Laramée[3], Benjamin Cheng[2], Sarah Galicia[2], Alain R. Bataille[3], ManTek Yeung[2,4], Megan Mendez[2], Maxime Bergeron[3], François Robert[3,5,6] and Daniel Durocher[2,4,6]

[1]These authors contributed equally to this work.

[2]Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario, M5G 1X5, Canada.

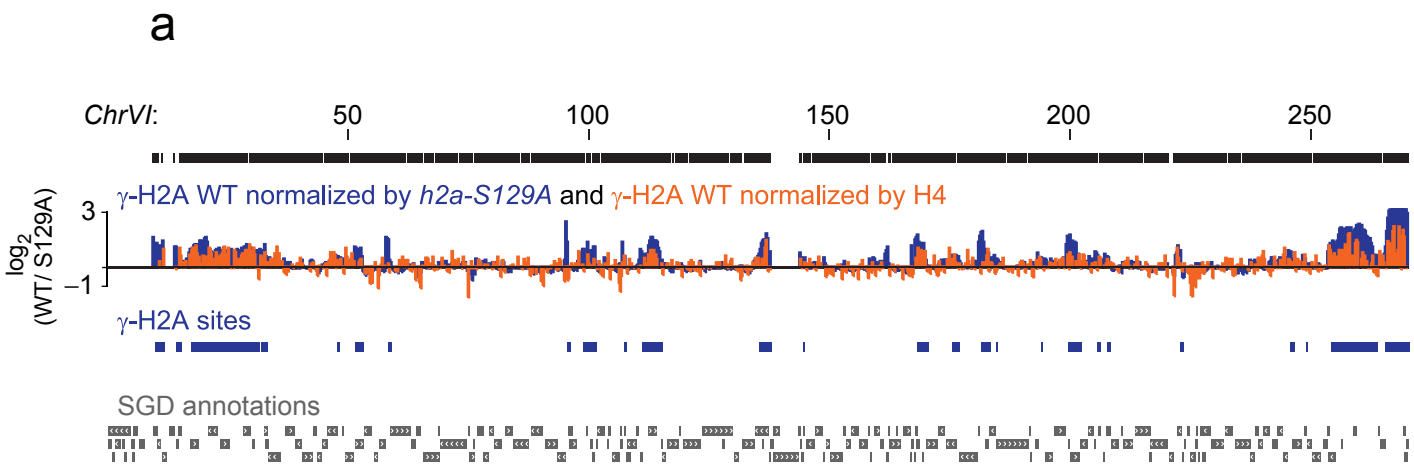
[3]Laboratoire de chromatine et expression du génome, Institut de recherches cliniques de Montréal, 110 Avenue des Pins Ouest, Montréal, Québec, H2W 1R7, Canada.

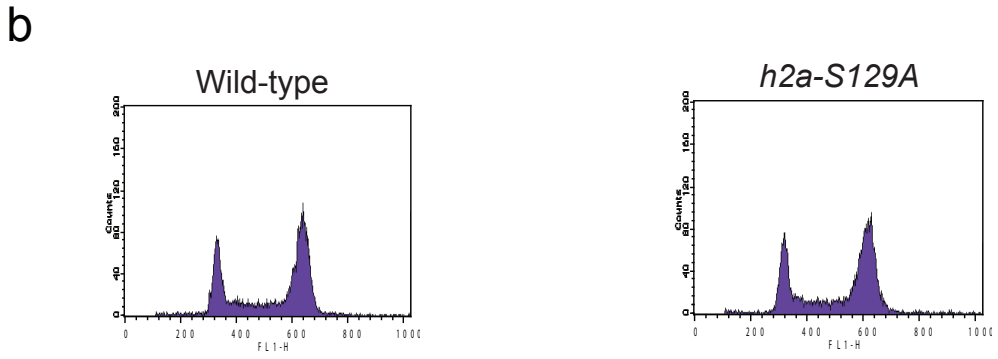
[4]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, M5S 1A8, Canada.

[5]Département de Médecine, Université de Montréal, Montréal, Québec, Canada.

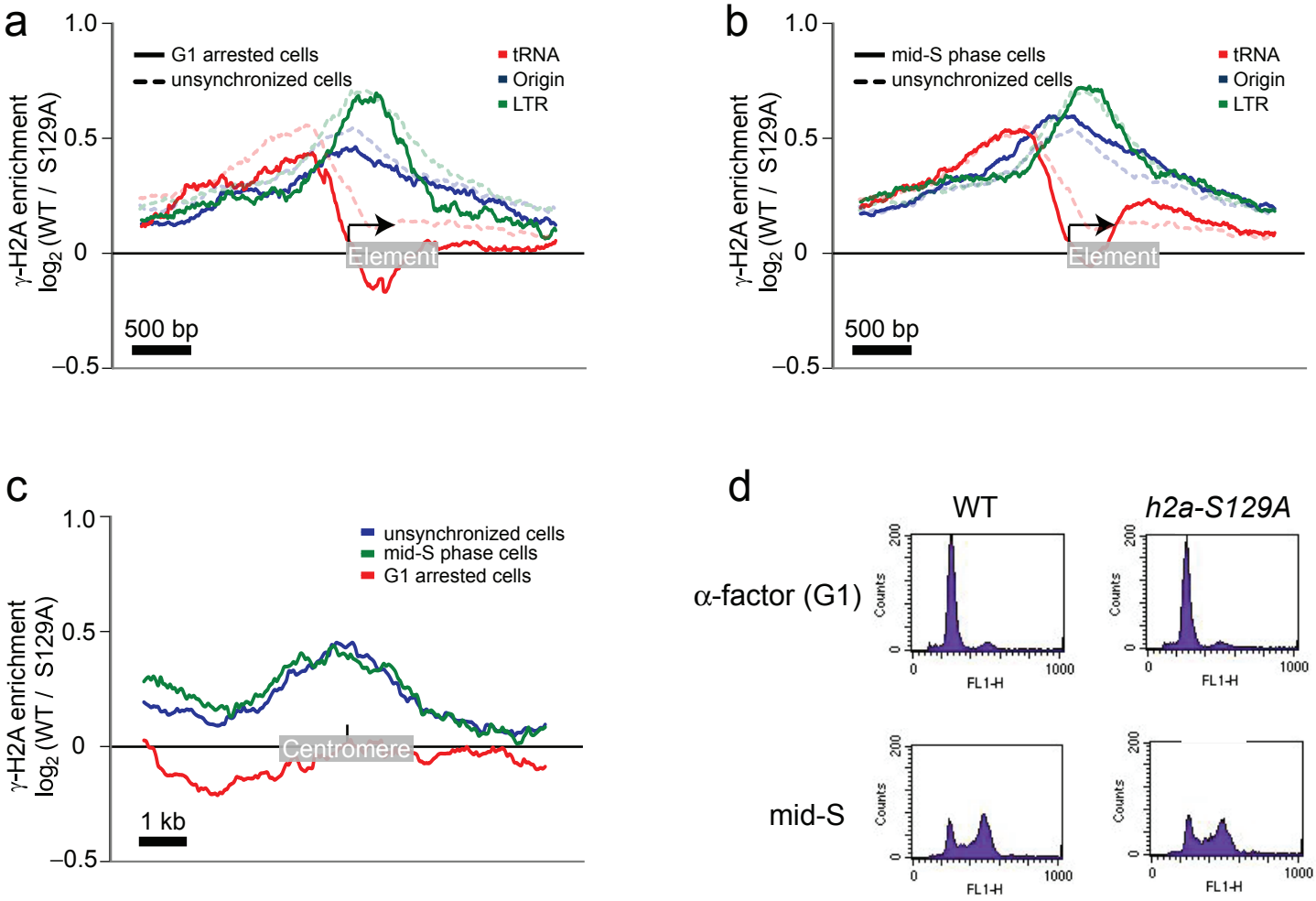
[6]Correspondence to: Daniel Durocher (durocher@lunenfeld.ca) and François Robert (Francois.Robert@ircm.qc.ca).

Contents

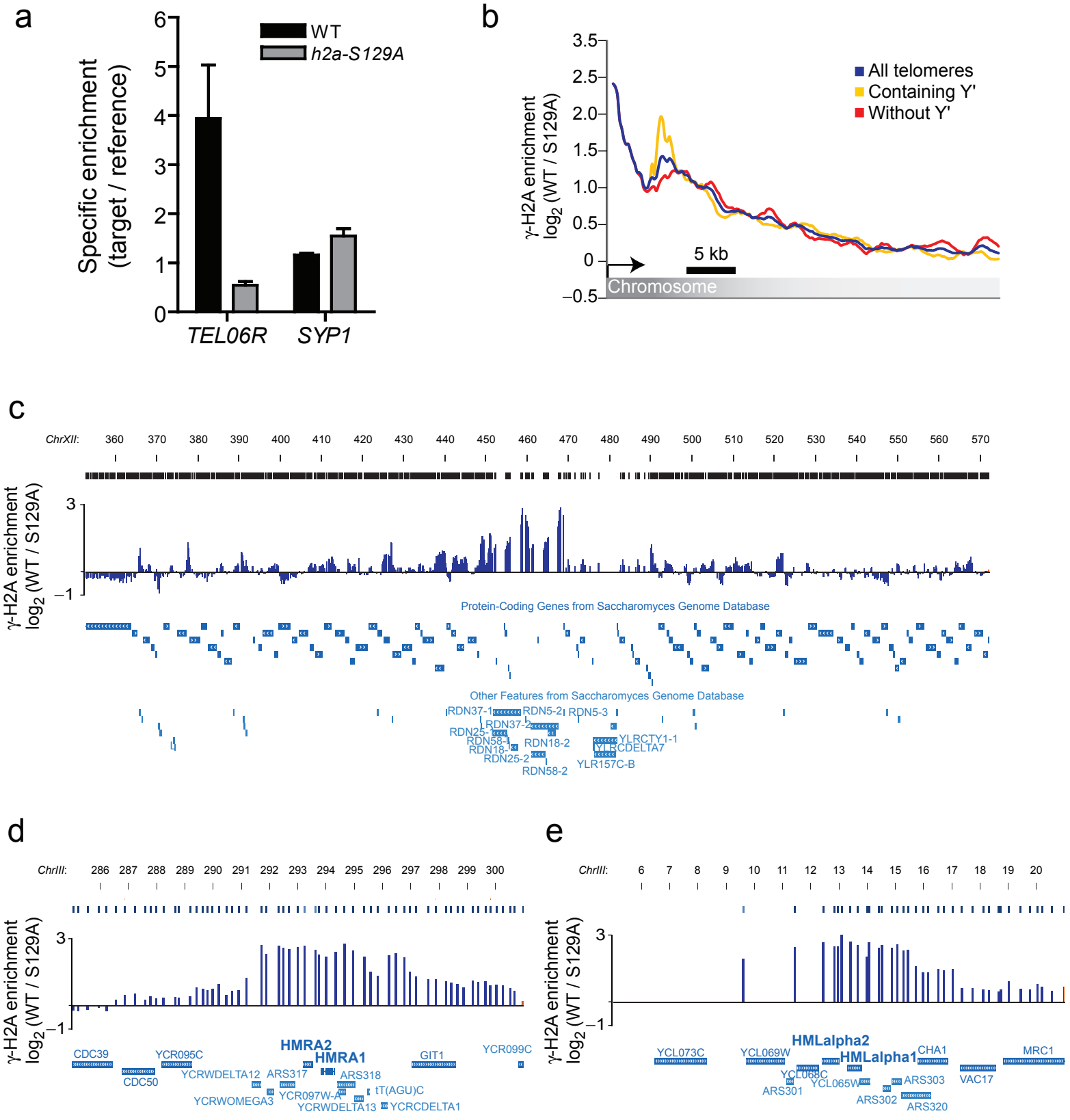




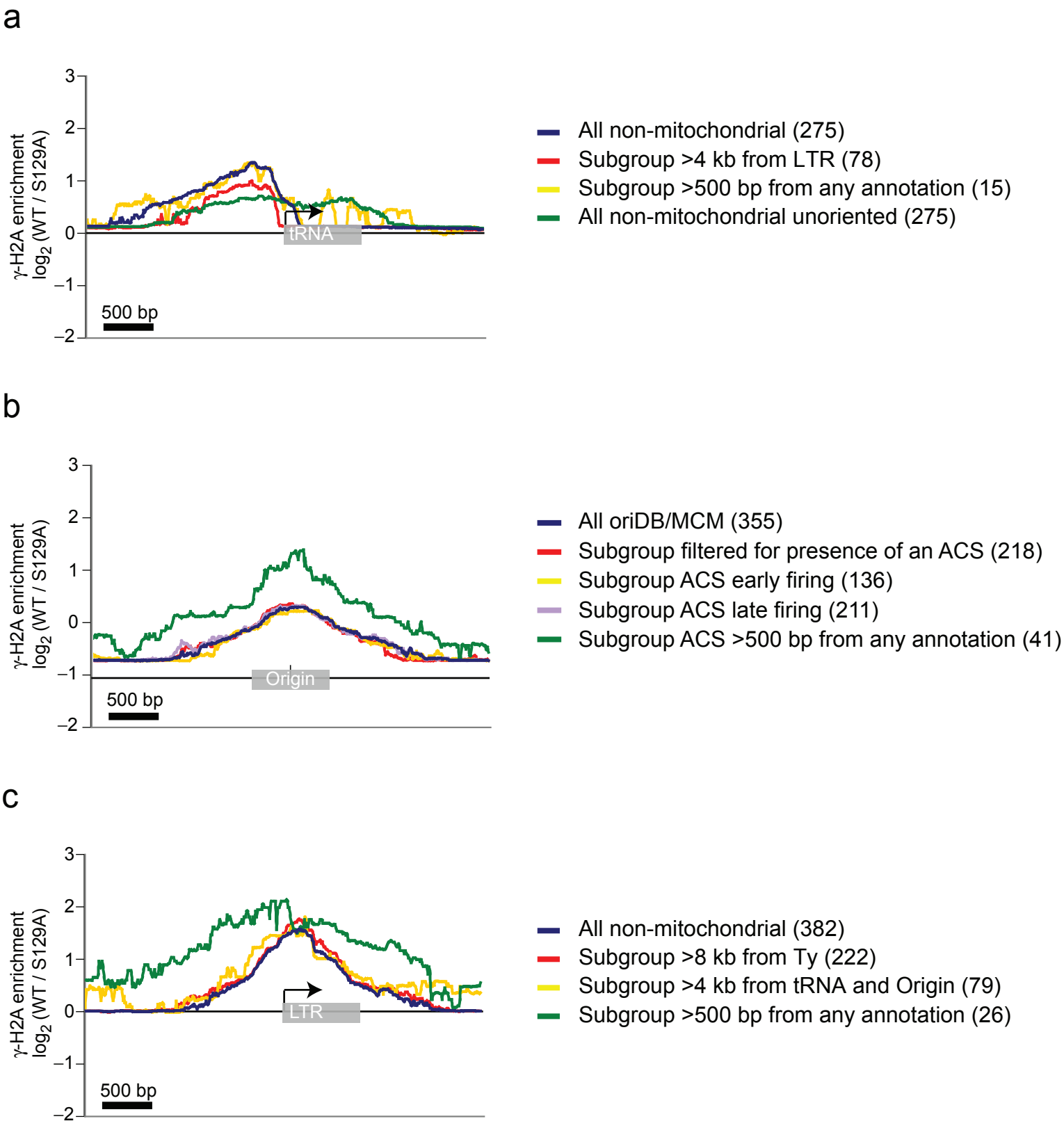
Supplementary Figure 1. Normalization of γ-H2A enrichment and effect of histone H2A phosphorylation on cell cycle profile. (**a**) Genome browser capture of *ChrVI*. The non-smoothed $\log_2$ γ-H2A enrichment ratio in wild-type (WT) is shown when normalized by *h2a-S129A* (blue; as in Fig. 1a) and by histone H4 (orange), along with the γ-H2A sites identified from the *h2a-S129A* normalized data (blue boxes; as in Fig. 1a). The probes are shown as black bars and the SGD annotations are shown in grey. Chromosomal positions are indicated in kilobases. (**b**) The *hta1-S129A hta2-S129A* (*h2a-S129A*) mutation does not affect the cell cycle profile of yeast grown in rich medium. FACS profiles of isogenic wild-type and *h2a-S129A* strains are shown.



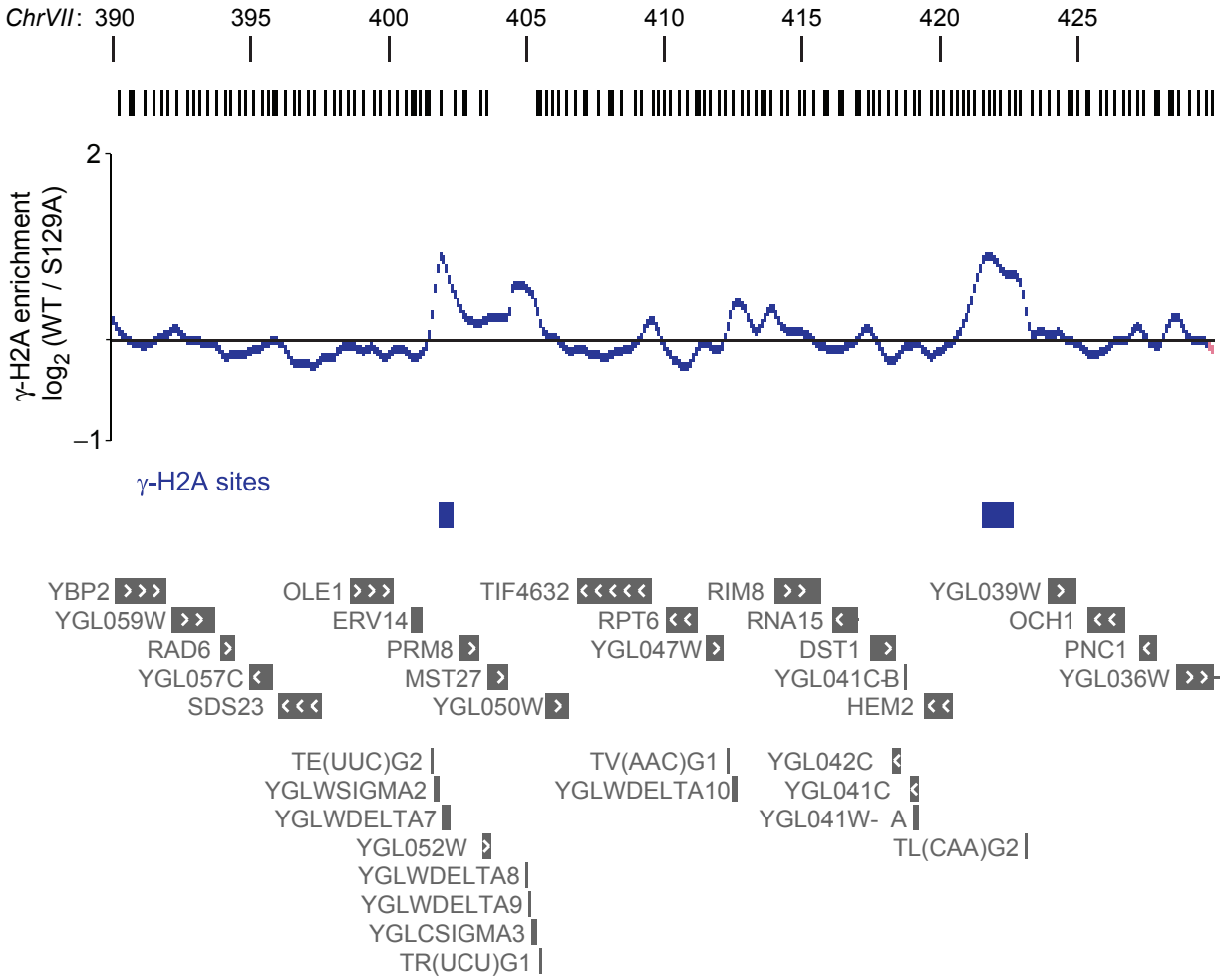
Supplementary Figure 2. γ-H2A enrichment is highly similar in asynchronous cells and those synchronized in G1 or mid-S phase. (**a**) Enrichment of γ-H2A from asynchronous WT cells (blue curves from Fig. 1b-d) are grouped on the same graph (tRNA genes in red, DNA replication origins in blue, and LTRs in green) and presented in dashed transparent curves while the γ-H2A enrichment detected in G1-arrested cells is shown in solid curves. (**b**) As in **a** except that solid curves were obtained from cells synchronized in mid-S phase. (**c**) The γ-H2A enrichment was mapped in the middle of the 16 centromeres in asynchronous (blue), G1-arrested (red), and mid-S phase (green) cells. (**d**) Cell cycle profiles of G1-arrested and mid-S phase cells examined in panels **a-c**. FACS analysis showing the DNA content of WT and *h2a-S129A* cells arrested in G1 with α-factor (upper) and in mid-S phase following release of the α-factor block (lower). A FACS profile of a typical asynchronous culture can be seen in Supplementary Figure 1b.



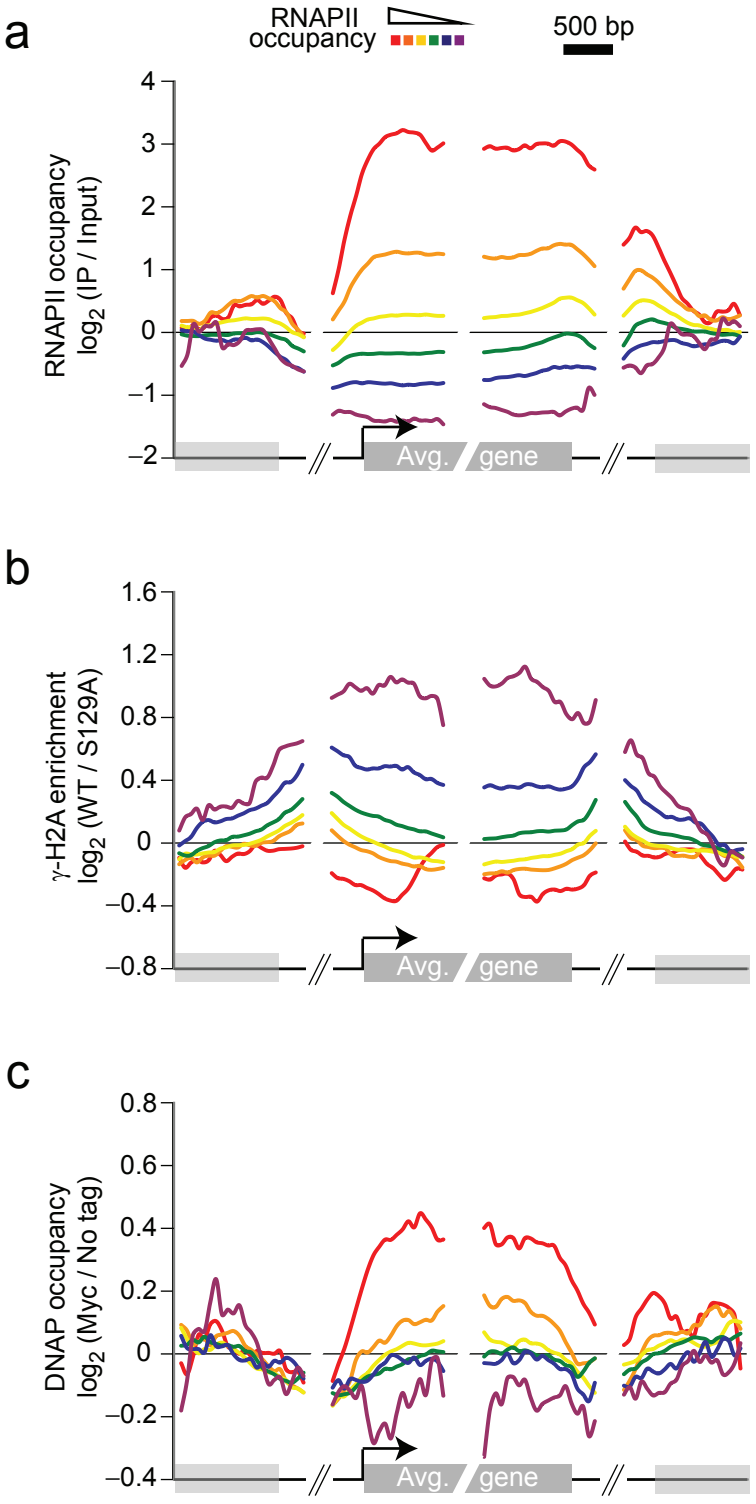
Supplementary Figure 3. Analysis of γ-H2A enrichment at telomeres, rDNA repeats, and the *HMR* and *HML* loci. (**a**) γ-H2A is enriched at the *TEL06R* telomere as determined by qPCR. Chromatin immunoprecipitation using the anti-γ-H2A antibody or normal rabbit serum were carried out with lysates obtained from WT or *h2a-S129A* strains. The *TEL06R* and *SYP1* loci were examined, the latter being a locus that does not show any specific enrichment in γ-H2A by ChIP-chip. Data is expressed as the average and standard deviation of two experiments. (**b**) The average γ-H2A enrichment ratios are mapped relative to the chromosomal end of all telomeres (blue) or sub-groups of telomeres containing (red) or lacking (gold) subtelomeric Y' elements. (**c-e**) Genome browser capture showing the non-smoothed $\log_2$ γ-H2A enrichment ratio in WT around the rDNA (**c**), *HMR* (**d**), and *HML* (**e**) loci similar to Supplementary Figure 1.



Supplementary Figure 4. Mapping of the WT average γ-H2A enrichment ratios on different sub-groups of the genetic elements presented in Figure 1b-d. The number of elements in each group is indicated in parenthesis. In panel **b** the origins were filtered for the presence of an ACS (ARS consensus sequence) based on a study of the Donaldson lab [1].



Supplementary Figure 5. Analysis of γ-H2A enrichment at a known yeast fragile site. Genome browser capture showing the smoothed $\log_2$ γ-H2A enrichment ratio and γ-H2A sites in WT (as in Supplementary Fig. 1) of the 403 locus mapped by Admire *et al.*[2]



Supplementary Figure 6. Analysis of RNAPII, γ-H2A and DNAP levels at protein coding genes. The average RNAPII occupancy (**a**), γ-H2A enrichment (**b**) and DNAP (**c**) occupancy ratios are mapped relative to the 5' and 3' boundaries of protein-coding genes grouped by their level of RNAPII occupancy.

Supplementary Table 1 - Summary of ChIP-chip experiments

Cy5			Cy3				
Antibody	**Description**	**Strain**	**Antibody**	**Description**	**Strain**	**#rep**	**Array used**
pH2A	WT	DDY705	pH2A	h2a-S129A	DDY1455	3	1 x 44k
pH2A	WT	DDY031	H4	WT	DDY031	2	1 x 44k
pH2A	rrm3D	DDY1273	pH2A	rrm3D, h2a-S129A	DDY1712	2	1 x 44k
MYC	POL2-myc	DDY1752	MYC	WT	DDY705	2	1 x 44k
HA	MCM4-HA	YFR215	HA	WT	YFR216	2	1 x 44k
HA	MCM7-HA	YFR214	HA	WT	YFR216	2	1 x 44k
8WG16	WT	YFR116	none	WT	YFR116	2	4 X 44k
pH2A	WT (gal 7h)	DDY705	pH2A	h2a-S129A (gal 7h)	DDY1455	2	4 X 44k
pH2A	hst1D	DDY1797	pH2A	hst1D, h2a-S129A	DDY1801	2	4 X 44k
pH2A	rpd3D	DDY1799	pH2A	rpd3D, h2a-S129A	DDY1803	2	4 X 44k
pH2A	rpd3-H188A	DDY2195	pH2A	rpd3-H188A, h2a-S129A	DDY2197	2	4 X 44k
pH2A	sml1D	DDY2011	pH2A	sml1D, h2a-S129A	DDY2007	2	4 X 44k
pH2A	mec1D, sml1D	DDY005	pH2A	mec1D, sml1D, h2a-S129A	DDY2001	2	4 X 44k
pH2A	tel1D	DDY2148	pH2A	tel1D, h2a-S129A	DDY2152	2	4 X 44k
pH2A	mec1D, sml1D, tel1D	DDY2156	pH2A	mec1D, sml1D, tel1D, h2a-S129A	DDY2158	2	4 X 44k
pH2A	WT (G1 arrested)	DDY705	pH2A	h2a-S129A (G1 arrested)	DDY1793	2	4 X 44k
pH2A	WT (mid-S)	DDY705	pH2A	h2a-S129A (mid-S)	DDY1793	2	4 X 44k

1 x 44k = G4486A

1 x 44k = G4493A

Supplementary Table 2 - Statistical analysis of the association between γ-H2A sites (as described elsewhere[3]) and different annotation categories

γ-H2A sites (param= Probe set p-value <0.1; Filter 1 p-value <0.5; Filter 2 p-value <0.1_0.5); stats were calculated from 10000 permutation tests

Category	Number of members of that category in the genome*	Number of sites associated with that category	Average length of associated sites	SD length of associated sites	% of sites associated with that category	Mean of the expected number of permuted associated sites	SD of the expected number of permuted associated sites	Z-score of the number of associated sites	pValue of the number of associated sites
tRNA	275	83	1441.7	144.6	11.9%	24.6	4.4	13.2	**< E-15**
Origin (MCM)	355	99	2665.7	246.8	14.2%	26.7	4.6	15.7	**< E-15**
LTR	382	74	1916.8	190.1	10.6%	18.6	3.5	15.6	**< E-15**
TEL	32	41	1698.9	129.3	5.9%	19.4	4.4	4.9	**4.6E-07**
CEN	16	2	1325.3	12.4	0.3%	1.4	1.2	0.5	**3.1E-01**
ORF	6576	394	1864.6	363.9	56.5%	485.6	8.0	-11.5	**1.0E+00**
Total	7770	**697**	**1254.8**	422.2	100.0%	697.0	0.0	0.0	---
NPS	1012	210	1940.4	316.5	**30.1%**	73.4	7.2	19.0	**< E-15**
ORF_<0	2837	340	1951.3	350.7	**48.8%**	269.1	11.9	6.0	**1.2E-09**

* excluding mitochondrial chromosome

γ-H2A sites (param= Probe set p-value <0.25; Filter 1 p-value <0.5; Filter 2 p-value <0.25_0.5); stats were calculated from 10000 permutation tests

Category	Number of members of that category in the genome*	Number of sites associated with that category	Average length of associated sites	SD length of associated sites	% of sites associated with that category	Mean of the expected number of permuted associated sites	SD of the expected number of permuted associated sites	Z-score of the number of associated sites	pValue of the number of associated sites
tRNA	275	128	2119.6	256.6	10.8%	43.5	6.6	12.7	**< E-15**
Origin (MCM)	355	158	3228.5	397.3	13.3%	48.3	6.4	17.1	**< E-15**
LTR	382	107	2579.0	285.2	9.0%	33.3	6.2	11.9	**< E-15**
TEL	32	36	1942.8	188.0	3.0%	30.9	5.4	0.9	**1.7E-01**
CEN	16	7	2775.4	53.2	0.6%	2.5	1.6	2.9	**2.0E-03**
ORF	6576	684	2060.2	616.5	57.7%	820.7	10.9	-12.5	**1.0E+00**
Total	7770	**1185**	1368.1	688.3	100.0%	1185.0	0.0	0.0	---
NPS	1012	305	2497.6	483.8	25.7%	126.1	11.0	16.3	**< E-15**
ORF_<0	2837	583	2185.9	593.1	49.2%	459.0	14.7	8.4	**< E-15**

Association criteria between each category member and γ-H2A sites

tRNA => included (>=75%) in a site, or the 5' must be <250nt from the border of a site
Origin => the position of the max MCM value must be included in a site
LTR => must be included (>=75%) in a site
TEL => the middle of the site must be in the first 10kb of the chromosome
Cen => must be included (>=75%) in a site
ORF => must be partly included (>=10%) in a site
NPS => Natural pausing site, including at least one of the often co-localized tRNA, DNA replication origin, LTR, TEL, or Centromere
ORF_<0 => must be partly included (>=10%) in a site AND log2 RNAPII occupancy<0

Supplementary Table 3 - Pearson product-moment correlation coefficient calculated on the complete WIG files*

		1	2	3	4	5	6	7	8	9	10	11	12	13
1	pH2A_WT__vs__pH2A_h2a-S129A	1	0.84	0.84	0.73	0.82	0.79	0.80	0.82	0.90	0.82	0.09	0.60	0.65
2	pH2A_WT__vs__H4_WT	0.84	1	0.81	0.65	0.77	0.72	0.72	0.74	0.76	0.73	0.09	0.55	0.57
3	pH2A_rrm3D__vs__pH2A_rrm3D, h2a-S129A	**0.84**	0.81	1	0.66	0.72	0.74	0.74	0.72	0.78	0.76	0.08	0.53	0.59
4	pH2A_WT (gal 7h)__vs__pH2A_h2a-S129A (gal 7h)	0.73	0.65	0.66	1	0.74	0.70	0.68	0.70	0.77	0.70	0.10	0.58	0.49
5	pH2A_hst1D__vs__pH2A_hst1D, h2a-S129A	0.82	0.77	0.72	0.74	1	0.75	0.72	0.82	0.83	0.80	0.11	0.59	0.62
6	pH2A_rpd3D__vs__pH2A_rpd3D, h2a-S129A	0.79	0.72	0.74	0.70	0.75	1	0.90	0.76	0.80	0.74	0.11	0.57	0.60
7	pH2A_rpd3-H188A__vs__pH2A_rpd3-H188A, h2a-S129A	0.80	0.72	0.74	0.68	0.72	**0.90**	1	0.77	0.80	0.74	0.12	0.57	0.60
8	pH2A_sml1D__vs__pH2A_sml1D, h2a-S129A	**0.82**	0.74	0.72	0.70	0.82	0.76	0.77	1	0.85	0.79	0.16	0.66	0.67
9	pH2A_mec1D, sml1D__vs__pH2A_mec1D, sml1D, h2a-S129A	0.90	0.76	0.78	0.77	0.83	0.80	0.80	**0.85**	1	0.86	0.13	0.66	0.70
10	pH2A_tel1D__vs__pH2A_tel1D, h2a-S129A	**0.82**	0.73	0.76	0.70	0.80	0.74	0.74	0.79	0.86	1	0.12	0.58	0.64
11	pH2A_mec1D, sml1D, tel1D__vs__pH2A_mec1D, sml1D, tel1D, h2a-S129A	0.09	0.09	0.08	0.10	0.11	0.11	0.12	**0.16**	0.13	0.12	1	0.17	0.17
12	pH2A_WT (G1 arrested)__vs__pH2A_h2a-S129A (G1 arrested)	**0.60**	0.55	0.53	0.58	0.59	0.57	0.57	0.66	0.66	0.58	0.17	1	0.54
13	pH2A_WT (mid-S)__vs__pH2A_h2a-S129A (mid-S)	**0.65**	0.57	0.59	0.49	0.62	0.60	0.60	0.67	0.70	0.64	0.17	0.54	1

*** In order to calculate the Pearson correlation between datasets obtained from different array designs, the first 3 datasets (hybridized on 1 X 44k) were smoothed and the average signal was calculated on the same oligo positions as the 4 X 44k (always excluding chrM).**

Supplementary Table 4 - Statistical analysis of the TFBS[4] (MacIsaac p0.001 c3) associated with γ-H2A sites (pVal<0.1) calculated from 10000 permutation tests

Name	Total number of TFBS in the genome	Number of pH2A sites containing this TFBS	Mean of the expected number of permuted pH2A sites containing this TFBS	SD of the expected number of permuted pH2A sites containing this TFBS	Z-score
UME6	72	27	4.33	2.22	**10.21**
SUM1	37	14	2.69	1.60	**7.05**
MBP1	127	11	5.26	2.27	**2.53**
GAL4	8	2	0.43	0.66	**2.37**
YOX1	6	2	0.48	0.68	**2.25**
IME1	3	1	0.19	0.43	**1.89**
MIG1	3	1	0.19	0.44	**1.87**
SKO1	11	2	0.57	0.77	**1.86**
YHP1	2	1	0.21	0.45	**1.77**
PUT3	3	1	0.21	0.46	**1.74**
PHD1	174	8	4.66	2.16	**1.55**
SWI6	165	8	5.18	2.25	**1.25**
STB4	6	1	0.31	0.56	**1.24**
BAS1	37	3	1.64	1.26	**1.08**
…	…	…	…	…	**…**
…	…	…	…	…	**…**
DIG1	163	3	5.32	2.30	**-1.01**
SKN7	131	3	5.29	2.25	**-1.02**
YAP5	34	0	1.26	1.14	**-1.11**
INO4	22	0	1.23	1.10	**-1.11**
ARG80	20	0	1.28	1.11	**-1.15**
INO2	22	0	1.33	1.14	**-1.17**
GCN4	107	3	5.83	2.37	**-1.19**
ROX1	29	0	1.61	1.26	**-1.28**
HAP4	31	0	1.69	1.31	**-1.29**
SPT2	35	0	1.68	1.28	**-1.31**
RCS1	55	0	2.20	1.48	**-1.49**
ABF1	150	5	9.51	3.02	**-1.49**
TEC1	55	0	2.28	1.49	**-1.53**
STE12	182	3	7.25	2.70	**-1.58**
YAP7	53	0	3.33	1.81	**-1.84**
HAP1	60	0	3.29	1.78	**-1.85**
YAP6	69	0	3.35	1.79	**-1.87**
RPN4	47	0	3.75	1.93	**-1.94**
FHL1	76	0	4.05	2.03	**-2.00**
CIN5	82	0	4.72	2.10	**-2.24**

Supplementary Table 5 - GO analysis using GO-TermFinder of SGD with all the ORFs as background

γ-H2A-associated ORFs with an enrichment difference >0.3 in the *hst1*Δ mutant compared to WT

Process

GOID	GO_term	Cluster frequency	Background frequency	P-value
30476	**ascospore wall assembly**	23 out of 201 genes, 11.4%	43 out of 6608 background genes, 0.7%	**5.01E-22**
42244	**spore wall assembly**	23 out of 201 genes, 11.4%	43 out of 6608 background genes, 0.7%	**5.01E-22**
30437	**ascospore formation**	25 out of 201 genes, 12.4%	114 out of 6608 background genes, 1.7%	**6.2E-13**
34293	**sexual sporulation**	25 out of 201 genes, 12.4%	114 out of 6608 background genes, 1.7%	**6.2E-13**
30154	**cell differentiation**	30 out of 201 genes, 14.9%	172 out of 6608 background genes, 2.6%	**6.43E-13**
48869	**cellular developmental process**	30 out of 201 genes, 14.9%	172 out of 6608 background genes, 2.6%	**6.43E-13**
22414	**reproductive process**	31 out of 201 genes, 15.4%	187 out of 6608 background genes, 2.8%	**9.26E-13**
30435	**sporulation**	25 out of 201 genes, 12.4%	126 out of 6608 background genes, 1.9%	**7.39E-12**
48610	**reproductive cellular process**	27 out of 201 genes, 13.4%	154 out of 6608 background genes, 2.3%	**1.57E-11**
3	**reproduction**	37 out of 201 genes, 18.4%	323 out of 6608 background genes, 4.9%	**2.67E-10**
22607	**cellular component assembly**	25 out of 201 genes, 12.4%	170 out of 6608 background genes, 2.6%	**8.3E-09**
45229	**external encapsulating structure organiz**	27 out of 201 genes, 13.4%	203 out of 6608 background genes, 3.1%	**1.39E-08**
7047	**cell wall organization and biogenesis**	27 out of 201 genes, 13.4%	203 out of 6608 background genes, 3.1%	**1.39E-08**
32502	**developmental process**	34 out of 201 genes, 16.9%	352 out of 6608 background genes, 5.3%	**2.75E-07**

γ-H2A-associated ORFs with an enrichment difference >0.3 in the *rpd3*Δ mutant compared to WT

Process

GOID	GO_term	Cluster frequency	Background frequency	P-value
7127	**meiosis I**	20 out of 242 genes, 8.3%	75 out of 6608 background genes, 1.1%	**4.08E-10**
279	**M phase**	31 out of 242 genes, 12.8%	194 out of 6608 background genes, 2.9%	**6.85E-10**
51321	**meiotic cell cycle**	23 out of 242 genes, 9.5%	126 out of 6608 background genes, 1.9%	**3.76E-08**
51327	**M phase of meiotic cell cycle**	23 out of 242 genes, 9.5%	126 out of 6608 background genes, 1.9%	**3.76E-08**
7126	**meiosis**	23 out of 242 genes, 9.5%	126 out of 6608 background genes, 1.9%	**3.76E-08**
22403	**cell cycle phase**	36 out of 242 genes, 14.9%	301 out of 6608 background genes, 4.6%	**6.07E-08**
7131	**meiotic recombination**	15 out of 242 genes, 6.2%	54 out of 6608 background genes, 0.8%	**1.59E-07**
22402	**cell cycle process**	38 out of 242 genes, 15.7%	368 out of 6608 background genes, 5.6%	**1.3E-06**
6259	**DNA metabolic process**	36 out of 242 genes, 14.9%	345 out of 6608 background genes, 5.2%	**2.68E-06**
7049	**cell cycle**	38 out of 242 genes, 15.7%	382 out of 6608 background genes, 5.8%	**3.71E-06**
6311	**meiotic gene conversion**	9 out of 242 genes, 3.7%	24 out of 6608 background genes, 0.4%	**2.93E-05**
6310	**DNA recombination**	18 out of 242 genes, 7.4%	112 out of 6608 background genes, 1.7%	**3.46E-05**

Supplementary Table 6 - Strains used in this study

Strain designation	Genotype	Background	Source
DDY705	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+*	W303	Rodney Rothstein (via David Lydall)
DDY031	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 rad5-535*	W303	Kim Nasmyth (via Angelika Amon)
DDY1455	*MAT* **a** *ade2-1 can1-100 leu2-3,112 trp1-1 ura3-1 rad5-35 hta1-S129A hta2-S129A*	W303	William Bonner (via Alain Verreault)
DDY1793	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ hta1-S129A hta2-S129A*	W303	This study
DDY1273	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ rrm3::NATMX*	W303	This study
DDY1712	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ rrm3::NATMX hta1S129A hta2S129A*	W303	This study
DDY1797	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ hst1::KANMX*	W303	This study
DDY1801	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ hst1::KANMX hta1-S129A hta2-S129A*	W303	This study
DDY1799	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ rpd3::KANMX*	W303	This study
DDY1803	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ rpd3::KANMX hta1-S129A hta2-S129A*	W303	This study
DDY1752	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 rad5-535 POL2::POL2-9Myc::TRP1*	W303	Karim Labib
DDY005	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ mec1::TRP1 sml1::HIS3*	W303	Rodney Rothstein
DDY2001	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ mec1::TRP1 sml1::HIS3 hta1-S129A hta2-S129A*	W303	This study
DDY2011	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ sml1::HIS3*	W303	This study
DDY2007	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ sml1::HIS3 hta1-S129A hta2-S129A*	W303	This study
DDY2148	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ tel1::NATMX*	W303	This study
DDY2152	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ tel1::NATMX hta1-S129A hta2-S129A*	W303	This study
DDY2156	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ mec1::TRP1 sml1::HIS3 tel1::NATMX*	W303	This study
DDY2158	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ mec1::TRP1 sml1::HIS3 tel1::NATMX hta1-S129A hta2-S129A*	W303	This study
DDY2195	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 rad5-535 rpd3-H188A::18myc:TRP1*	W303	This study
DDY2197	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 rad5-535 rpd3-H188A::18myc:TRP1 hta1-S129A hta2-S129A*	W303	This study
DDY821	*MAT* **a** *ADE2 bar1::LEU2 trp1-1 LYS2 RAD5+ RAD52-YFP*	isogenic to W303	Rodney Rothstein
DDY964	*MAT* **a** *ADE2 bar1::LEU2 trp1-1 LYS2 RAD5+ esc2::KANMX RAD52-YFP*	isogenic to W303	This study
DDY1729	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ DDC2-GFP:URA3MX3*	W303	This study
DDY1964	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ hta1-S129A hta2-S129A DDC2-GFP:URA3MX3*	W303	This study
DDY1746	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ <pRS415-RAD52-YFP>*	W303	This study
DDY1967	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ hta1-S129A hta2-S129A <pRS415-RAD52-YFP>*	W303	This study
DDY1775	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ ChrXV::CAN1-URA3*	W303	This study
DDY1771	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 RAD5+ ChrXV::CAN1-URA3 hta1-S129A hta2-S129A*	W303	This study
YFR214	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 rad5-535 bar1::hisG mcm7 Δ::MCM7-HA::TRP1*	W303	John Wyrick
YFR215	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 rad5-535 bar1::hisG mcm4 Δ::MCM4-HA::TRP1*	W303	John Wyrick
YFR216	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 rad5-535 bar1::hisG*	W303	John Wyrick
YFR116	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1*	W303	Richard Young
YFR639	*MAT* **a** *ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3-1 rad5-535 POL2::POL2-9Myc::TRP1 tS(GCU)L* ^{amb}	W303	This study

Supplementary Table 7 - qPCR Oligonucleotides

Locus	Primer Pair (5'- 3')
TEL6-R	TGAGGCCATTTCCGTGTGTA
	CCCAGTCCTCATTTCCATCAA
SYP1	CCCGTGGCTTGAGGATTT
	CGCTACCCCTACTTCTTCTTCC
TSC11	GTTGCCGCTTCAGTTGTTGT
	ATCATTCTACGCTTTCCATTCC
tS(GCU)L-1	TGAGAATGGAAAGGGCAAGATAA
	TTACTTCGTCGTGAGGGTGGAT
tS(GCU)L-2	GAGGTGTACTACCATTCATGAGAGTCG
	GGGACGGTTGAACAACGATCTA
tS(GCU)L-3	AGGATATAGGAATGCGCAAATGG
	TACGCCGAATGCGGGAAT
tS(GCU)L-4	CGAATCCTGTCTGTGACGCTTT
	GCGAATCTTCGAGCAATAAAGTTC
tS(GCU)L-5	TGCTCGAAGATTCGCTATCCAT
	TCCCTGGTATAAGTAATTGACGCAAC
ARN1	GAGAGCTATCGAATGTTTCCTC
	TGCACCCATAAAAGCAGGTGT

Supplementary File Legends

Supplementary File 1. The combined datasets of all the ChIP-chip data used in this study (Supplementary Table 1), along with the γ-H2A sites identified with the two sets of parameters analyzed (Supplementary Table 2). The file, compressed with gzip, is in the BED format compatible with the UCSC Genome Browser (http://genome.ucsc.edu/)[5].

Supplementary File 2. The mean $\log_2$ calculated on the promoter (defined as half of the upstream intergenic regions and up to 250 bp[6]) and on the complete length of each annotation of SGD (version Feb. 02 2008) for each ChIP-chip dataset produced in this study.

Supplementary Methods

Identification of γ-sites. We identified the γ-sites listed in Supplementary Table 2 using an algorithm developed by Boyer *et al.*[3] in the Young laboratory. The main idea is to automatically identify statistically enriched regions in a dataset by incorporating information from neighboring probes. For each oligonucleotide, the algorithm calculates the average X score (from an error model[7]) of the oligonucleotide and its two immediate neighbors taking into account the distance between the probes. This set of averaged values gives a new distribution that is subsequently used to calculate p-values of average X (probe set p-values). If the "probe set p-value" is less than the selected threshold (0.1 and 0.2 in the present study), the three probes are marked as potentially bound. Other filters are also applied, see the Supplemental Data of the original article.

Mapping γ-H2A enrichment at natural replication fork barriers. We used the complete sets (Fig. 1b-d, Fig. 3a,b,d and Supplementary Figure 2a,b) or subsets (Supplementary Fig. 4) of the non-mitochondrial tRNA genes, DNA replication origins and LTR elements to map the γ-H2A enrichment (a total of 275, 355 and 382 elements, respectively). We mapped the non-smoothed data on either the 5' end of the tRNA genes, the middle of the DNA replication origins or the 5' end of the LTR elements into 10 bp windows, and applied a sliding window of 500 bp to the averaged ratios. Similarly, we mapped the non-smoothed data on the middle of the 16 centromeres (Fig. 3c and Supplementary Fig. 2c) into 50 bp windows with a sliding window of 1.5 kb. In each replicate of the *mec1Δ tel1Δ sml1Δ* cells, different genomic regions (or chromosomes) were unstable. We removed those regions from the combined dataset prior to the mapping. We also applied the same mapping procedure to the 32 non-mitochondrial chromosome ends using windows of 250 bp and a sliding window of 2 kb (Supplementary Fig. 3b).

Correlating γ-H2A enrichment and RNAPII occupancy. In order to investigate the anti-correlation between γ-H2A and RNAPII occupancy (Fig. 5a), we used the smoothed data for every non-mitochondrial protein-coding gene (using SGD version Feb. 02 2008; http://www.yeastgenome.org/) to calculate the mean $\log_2$ γ-H2A (WT / *h2a-S129A*) and the mean $\log_2$ RNAPII (IP / Input) ratios on the complete length of the gene and on the promoter (half of the upstream intergenic (IG) regions and up to 250 bp) (Supplementary File 2). We then sorted genes by their mean RNAPII occupancy calculated on the complete length of the gene.

We applied the median of a sliding window of 100 genes (in red) to both the mean γ-H2A enrichment and the mean RNAPII occupancy.

Mapping γ-H2A enrichment along protein-coding genes. First, we grouped protein-coding genes with an average γ-H2A enrichment >0.5 (Fig. 5b, solid curves). As a control, we randomly selected a group of 494 genes from those with an average γ-H2A enrichment <0.5 (Fig. 5b, dashed curves). Similarly, we grouped genes based of their level of RNAPII occupancy on the complete length of the gene (red 301, gold 764, yellow 2516, green 2069, blue 588, and purple 147 genes) (Supplementary Fig. 6). Note that we excluded all mitochondrial genes as well as nuclear genes that are completely embedded within a longer gene from these analyses. As in Rufiange *et al.*[6], we mapped the non-smoothed data onto the 5' and 3' boundaries in 50 bp windows for each half-gene and adjacent half-IG regions. We then applied a sliding window of 300 bp to the ratios. We used the same procedure in Figure 5d to map the difference of γ-H2A enrichment in galactose vs. glucose (calculated probe by probe) on groups of genes based on their level of induction in galactose (from the blue to red curves, respectively 423, 2626, 1774 and 540 genes).

Analysis of genes affected in HDAC mutants. We calculated the average γ-H2A enrichment difference between mutants and WT for each protein-coding gene as in Figure 5a. We submitted genes with a difference above 0.2 (201 genes) and 0.3 (242 genes) respectively in the *HST1* and *RPD3* mutants to the GO-TermFinder tool[8] of SGD (http://db.yeastgenome.org/cgi-bin/GO/goTermFinder.pl) to identify significantly enriched GO terms (Supplementary Table 5), and used these to evaluate the proportion of genes whose promoter is directly bound by the HDACs ("HDAC-dependent γ-H2A enrichment" Fig. 6e,f). We relativised the Hst1 and Rpd3 enrichment levels from Robert *et al.*[9] by their respective maximal value, and then by the number count in each relative bin (based on -0.1, 0.1 and 0.2).

Microscopy. We fixed cells in 4% (w/v) paraformaldehyde/0.1M sucrose for 15 minutes in the dark at room temperature, then stored and imaged the cells in a 1.2M sorbitol, 0.1M sodium phosphate (pH 7.5) buffer. We mounted the cells on Fisherbrand microscope slides and coverslips (12-550A; 12-542B). We captured images at room temperature with a Nikon E600 FN microscope (http://www.nikonusa.com) equipped with an ORCA ER2 camera (http://www.hamamatsu.com), Chroma filters (YFP: JP4 C/YFP, dichroic 86002v2, batch C37370, excitation = 500 ± 20 nm, emission 535 ± 30 nm; GFP: 41025 PSTN GFP, dichroic Q495LP, batch C37368, excitation = 470 ± 40 nm, emission = 515 ± 30 nm; http://www.chroma.com/) and a 100X/1.4 oil emersion Nikon Plan Apo objective. We processed images with Metamorph v5.0r6. We acquired 8-10 z-stacks over 1-2 microns to capture the entire nucleus with a binning of 1. The exposure time for Ddc2-GFP and Rad52-YFP was 1s. The acquired images have a resolution of 1200x1024. DIC images have a bit depth of 8 while GFP and YFP images have a bit depth of 16. Budded cells were scored for foci by visually examining each plane; images were not manipulated with deconvolution or three-dimensional reconstructions.

Supplementary References

1. Nieduszynski, C.A., Knox, Y. & Donaldson, A.D. Genome-wide identification of replication origins in yeast by comparative genomics. *Genes Dev*. **20**, 1874-9 (2006).

2. Admire, A. *et al*. Cycles of chromosome instability are associated with a fragile site and are increased by defects in DNA replication and checkpoint controls in yeast. *Genes Dev.* **20**, 159-73 (2006).

3. Boyer L.A. *et al*. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-56 (2005).

4. MacIsaac, K.D. *et al*. An improved map of conserved regulatory sites for Saccharomyces cerevisiae. *BMC Bioinformatics* **7**, 113 (2006).

5. Kent, W.J. *et al*. The human genome browser at UCSC. *Genome Res.* **12**, 996-1006 (2002).

6. Rufiange, A., Jacques, P.E., Bhat, W., Robert, F. & Nourani, A. Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. *Mol. Cell* **27**, 393-405 (2007).

7. Ren, B. *et al*. Genome-wide location and function of DNA binding proteins. *Science* **290**, 2306-9 (2000).

8. Boyle, E.I. *et al*. GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **20**, 3710-5 (2004).

9. Robert, F. *et al*. Global position and recruitment of HATs and HDACs in the yeast genome. *Mol. Cell* **16**, 199-209 (2004).