SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin.

Christèle Maison, Delphine Bailly, Danièle Roche, Rocio Montes de Oca, Aline V Probst, Isabelle Vassias, Florent Dingli, Bérengère Lombard, Damarys Loew, Jean-Pierre Quivy, et al.

SUMOylation promotes *de novo* targeting of HP1α to pericentric heterochromatin

Christèle Maison[1], Delphine Bailly[1], Danièle Roche[1], Rocío Montes de Oca[1], Aline V. Probst[1#], Isabelle Vassias[1], Florent Dingli[2], Bérengère Lombard[2], Damarys Loew[2], Jean-Pierre Quivy[1]* and Geneviève Almouzni[1]*

[1] Institut Curie UMR 218 CNRS, Laboratory of Nuclear Dynamics and Genome Plasticity, 26 rue d'Ulm, 75248 Paris Cedex 05, France

[2] Institut Curie, Laboratory of Proteomic Mass Spectrometry, 26 rue d'Ulm, 75248 Paris Cedex 05, France

[#] Present address: CNRS, UMR 6247 GreD, Clermont Université, INSERM U931, 24 avenue des Landais, 63177 Aubière, France

* Corresponding authors

For correspondence: G. Almouzni

Tel: + 33 (0) 1 56 24 67 01

Fax: + 33 (0) 1 46 33 30 16

E-mail: genevieve.almouzni@curie.fr

Keywords: HP1α, sumoylation, non-coding centromeric RNAs, pericentric heterochromatin

ABSTRACT

HP1 enrichment at pericentric heterochromatin is considered important for centromere function. While HP1 binding to H3K9me3 can explain its accumulation at pericentric heterochromatin, how it is initially targeted there remains unclear. Here, in mouse cells, we reveal the presence of long nuclear non-coding transcripts corresponding to major satellite repeats at the periphery of pericentric heterochromatin. Furthermore, we find that major transcripts in the forward orientation specifically associate with SUMO-modified HP1 proteins. We identified this modification as SUMO-1 and mapped it in the hinge domain of HP1α. Importantly, the hinge domain and its sumoylation prove critical to promote the initial targeting of HP1α to pericentric domains using *de novo* localization assays whereas they are dispensable for maintenance of HP1 domains. We propose that SUMO-HP1, through a specific association with major forward transcript, is guided at pericentric heterochromatin domain to seed further HP1 localization.

Heterochromatin at pericentric domains represents a paradigm for understanding how a functional nuclear domain is established and maintained. Studies in S. *pombe* and Drosophila have advanced our knowledge concerning basic mechanisms and conserved components in the organization of the domain[1]. In particular HP1 proteins which accumulate in these domains are highly conserved from S. *pombe* to mammals. However in mammalian cells, how HP1 is specifically targeted *de novo* to initiate formation of a domain of accumulation as found in pericentric heterochromatin[2,3] remains mysterious. The recognition of H3K9me3 by HP1[4,5] which exemplifies the paradigm of the reader model in the histone code hypothesis[6,7] does not necessarily provide on its own a *de novo* specific targeting mechanism. One should also consider other HP1 binding partners[8,9], potential post-translational modifications[10] and the elusive RNA that has been linked to the presence of HP1 at pericentric heterochromatin[11,12]. While transcripts from major satellite DNA repeats have been identified[13,14], a functional connection between specific RNAs and HP1 has not yet been established. Taken together, these data prompted us to explore further the HP1-RNA connection and particular post-translational modifications or partners that could provide a means for the *de novo* targeting of HP1 to pericentric heterochromatin, thereby helping to define this specific subnuclear compartment.

We first verified that repetitive DNA sequences in mouse centromeres, known as major and minor satellites in pericentric and centric heterochromatin respectively[15], can be transcribed in both orientations. All these transcripts detected by RT-PCR analysis with strand-specific primers for major or minor satellites showed various sizes corresponding to multiple repeats of their basic units (Fig. 1a). We next examined whether some of these transcripts could be stably found in the nucleus. As a reference for comparison, we used the well-defined organization of chromocenters with major satellite repeats surrounded by minor satellite DNA (Fig. 1b, left, DNA FISH)[2]. For this, we exploited fluorescently labelled locked nucleic acid (LNA) probes for RNA FISH that detect specifically major and minor transcripts in either forward or reverse orientation. While minor RNAs were barely detectable, LNA probes for major RNAs in both orientations revealed a significant number of nuclear spots frequently associated or close to chromocenters (Fig. 1b, middle). These signals were not a result of DNA cross-hybridization, since they were undetectable after RNase treatment (Supplementary Fig. 1). With some variations from cell to cell suggesting cell cycle modulation, reverse major RNAs were usually detected as fewer and larger spots (Fig. 1b, middle, red) and forward major RNA signals were smaller in size and present in higher numbers (green). Interestingly, these forward major RNAs were consistently at the periphery of HP1α domains (Fig. 1b, right, immuno-RNA FISH and Supplementary Fig. 2a). Since active transcription occurs frequently in the external part of chromosomal domains by DNA looping out[16], they may represent primary transcripts stably maintained at the site of transcription. Notably, HP1α accumulated at major satellite domains, away from minor satellites

which were juxtaposed (Fig. 1c and Supplementary Fig. 2b, Immuno-DNA FISH)[2]. We then examined further the possible *in vivo* association of major RNAs with major satellite domains by other methods. Under conditions that preserved RNAs, we immunoprecipitated heterochromatin-associated material with a specific HP1α antibody. We found a significantly higher amount of forward major transcripts compared to reverse transcripts (Fig. 1d). Taken together, our data highlight a particular link between major RNA in the forward orientation and pericentric heterochromatin. While HP1α can bind RNA *in vitro*[12], a specific binding to a given transcript has been reported so far only for TERRA RNA[17]. Using GST-tagged HP1α full-length protein or fragments thereof and radioactively labelled centromeric RNA probes (Fig. 2a), we find that the hinge domain (H) of HP1α strongly recognized both forward and reverse centromeric RNA probes, while the chromo-hinge (CD+H) and the hinge-chromoshadow (H+CSD) domains showed a lower binding capacity. Under these conditions, we did not detect binding between full-length HP1α and RNAs (Fig. 2a). These results suggest that the central hinge domain may adopt a different conformation in the full-length protein, possibly constrained by the CD and/or the CSD. This raised the possibility that HP1α could actually bind RNA *in vivo* and that this interaction can be regulated. To investigate this hypothesis, we aimed to identify proteins from nuclear cell extracts that associated with centromeric RNAs, using *in vitro* transcribed biotin-labelled forward major or minor RNAs immobilized on streptavidin beads (Fig. 2b and Supplementary Fig. 3a). Mass spectrometry analysis of the RNA-associated proteins identified mainly hnRNPs and proteins involved in RNA processing (Supplementary Fig. 3b and Supplementary Table 1) including vigilin, RNA helicase A (RHA) and Ras-GAP SH3 domain binding protein (G3bp), which were all validated by Western blotting (Fig. 2c). Vigilin and the *Drosophila* RHA ortholog, maleless (MLE), are thought to play roles in heterochromatin formation and X-chromosome dosage compensation, respectively[18,19], while G3bp may be involved in RNA metabolism[20]. Surprisingly, under these experimental conditions, we did not detect any significant association between HP1α and major or minor RNAs, neither by mass spectrometry (Supplementary Fig. 3b) nor by Western blotting (Fig. 2c). Given our *in vitro* observations (Fig. 2a), we wondered whether post-translational modification of HP1α could promote *in vivo* HP1α binding to RNA. Indeed, HP1 can be phosphorylated[10] and in fission yeast, Swi6/HP1 is sumoylated *in vivo*[21]. Furthermore, defective sumoylation of Swi6 results in significantly reduced heterochromatin stability. This prompted us to repeat our RNA pull down strategy with nuclear extracts prepared in the presence of a cocktail of phosphatase inhibitors and N-ethylmaleimide (NEM), a strong inhibitor of SUMO isopeptidases[22]. Remarkably, Western blotting with an anti-HP1α antibody revealed a slower migrating band in the input for HP1α that was not detected in absence of NEM and that then was specifically enriched in precipitates with forward major RNAs corresponding to either one or two satellite repeats length (Fig. 2d; Maj1 F or Maj2 F). This band, with an approximate additional molecular mass of ~11 kDa compared to HP1α, might represent an ubiquitin or a SUMO moiety. We confirmed that

this slower-migrating form was modified by SUMO-1, but not SUMO-2/3, using specific antibodies (Fig. 2d). We obtained similar results with antibodies specific of HP1β and HP1γ (Supplementary Fig. 3c) indicating that all HP1 isoforms can potentially be sumoylated *in vivo* and are specifically retrieved with forward major RNAs. While the forward major RNA pulls down SUMO-1-HP1α, we estimated that only a small percentage of HP1α was SUMO-1 modified *in vivo* (less than 1% of total protein from nuclear extracts, Fig. 2d). This low representation is in agreement with rapid cycles of sumoylation and de-sumoylation as reported for other proteins weakly detected in a SUMO-modified state *in vivo*[23,24].

To confirm independently that HP1α gets sumoylated *in vivo*, we co-transfected HA-tagged HP1α (e-HP1α) and GFP-SUMO-1 into NIH3T3 cells to prepare total cell extracts and carried out immunoprecipitations with anti-HA beads under conditions that preserved the SUMO modification[25]. Western blot analysis using anti-GFP antibodies clearly revealed a band corresponding to GFP-SUMO-e-HP1α (Supplementary Fig. 4a), demonstrating unambiguously that HP1α can be sumoylated *in vivo*. To determine which region of HP1α was SUMO-modified, we next used an *in vitro* sumoylation assay with either wildtype (W) or mutant (M) SUMO-1 protein, in the presence of E1 activating and E2 conjugating (Ubc9) SUMO enzymes and various GST-HP1α domains (Fig. 3a). Western blot analysis using anti-GST antibodies showed SUMO-modification on the full-length HP1α, the chromo-hinge, the hinge-chromoshadow and the hinge domains, the latter domain showing the highest level of sumoylation (Fig. 3b and Supplementary Fig. 4b, asterisks). This is also true when using SUMO-2 or SUMO-3 (Supplementary Fig. 4c). These results revealed that the hinge, the same domain of HP1α that displayed RNA binding activity (Fig. 2a), is a target for sumoylation. Since the fusion of Ubc9 to a substrate provides a convenient way to increase its sumoylation[26], we also verified with a GST-HP1α-Ubc9 fusion protein that we could enhance SUMO-1-modification of HP1α *in vitro* without adding the E2 enzyme (Fig. 3c). Then, using the GST-HP1α hinge and SUMO-1 proteins as above (Supplementary Fig. 5a), we further identified specific sumoylated residues on HP1α by mass spectrometry. In the hinge domain, among the 13 lysines (K) as potential targets for sumoylation, we found that K84 of HP1α in the peptide, "EKSEGNK", was unequivocally identified as sumoylated by mass (with high accuracy in the orbitrap; Fig. 3d, right spectra, arrow) and sequence (by MS/MS in the Qstar mass spectrometer; Fig. 3d; shown y and b ions). However, we also found other potentially sumoylated lysines in HP1α (e.g., KMoxK, SKK, KYK or YKK; Supplementary Fig. 5b and c) indicating alternative usage of various lysine residues. We thus mutated successively each of the individual 13 lysines to arginine and performed *in vitro* sumoylation assays as above and *in vivo* co-transfections as in Supplementary Figure 4a. We observed in all cases that once we mutated one residue an alternative sumoylation site was used (data not shown), showing the

usage of more than one sumoylation site and corroborating our mass spectrometry data. This suggests a lack of a strict requirement for a ”specific” SUMO modified residue.

Since the fusion of Ubc9 to HP1α increases sumoylation efficiency *in vitro* (Fig. 3c), we thus generated an HP1α-Ubc9-HA fusion construct (e-HP1α-Ubc9) for *in vivo* expression and further analysis in a cellular context. Transfection of e-HP1α-Ubc9 or e-HP1α into NIH3T3 cells lead to comparable levels of proteins being expressed (Supplementary Fig. 6a and b; anti-HA). Interestingly, a significantly higher amount of SUMO-modified e-HP1α-Ubc9 could be detected. In total cell extracts, we found that e-HP1α-Ubc9 was mainly SUMO-2/3-modified reflecting the readily available endogenous SUMO-2/3 proteins compared to the limited amount of free endogenous SUMO-1 protein in cells (compare SUMO-1 *vs* SUMO-2/3 inputs in Fig. 2d)[23]. This could be compensated for by providing exogenous SUMO-1 by transient transfection (Supplementary Fig. 6c). These *in vivo* results underline the fact that HP1α can be modified by SUMO-1 or SUMO-2/3 depending on the available substrate, as shown *in vitro (*Fig. 3b&c and Supplementary Fig. 4b&c*)*. We verified that the catalytic activity of the fused Ubc9 was directly involved in HP1α sumoylation enhancement using a Ubc9 catalytic mutant fused to HP1α (e-HP1α-Ubc9C93S). In comparaison with e-HP1α-Ubc9wt, we detected a strongly reduced sumoylation of e-HP1α-Ubc9C93S (Supplementary Fig. 6c). Next, we examined e-HP1α-Ubc9 and e-HP1α localization in NIH3T3 cells. Both of them could accumulate at pericentric domains where endogenous HP1α is already located (Fig. 3e). Thus under these conditions, promoting HP1 sumoylation did not give a particular advantage for the recruitment and maintenance of exogenous HP1 to preexisting HP1 domains of accumulation. Next, we wondered whether HP1α sumoylation could be required more specifically for a *de novo* targeting of HP1α to heterochromatin domains. To test this hypothesis we used MEFs derived from *Suv39h* double-null (dn) mice, in which the H3K9me3 mark and HP1α are no longer enriched at pericentric heterochromatin[11,27] (Fig. 4). Transfection with exogenous Myc-SUV39H1 can restore the proper localization of these marks (Fig. 4b)[5,28]. Since we postulated that the interaction between major RNAs and sumoylated HP1α targets HP1α to pericentric domains, we first verified that we could detect major RNAs by RNA-FISH in *Suv39h* dn MEFs (Supplementary Fig. 7a). We then transiently transfected *Suv39h* dn cells with e-HP1α or e-HP1α-Ubc9 in the absence of Myc-SUV39H1 and verified that the proteins were expressed at comparable levels (Fig. 4c). We could not detect the typical localization to chromocenters in cells transfected with e-HP1α (Fig. 4c). In contrast, in about 10% of cells transfected with e-HP1α-Ubc9 cells, we found a faint but detectable pericentric localization 6h post-transfection. Remarkably, in the latter case, we could not detect H3K9me3 accumulation at pericentric heterochromatin (Fig. 4c and Supplementary Fig. 7b) while H3K9me1 was clearly visible at these domains in all *Suv39h* dn cells (Supplementary Fig. 7c). Thus, a targeting of e-HP1α-Ubc9 to pericentric heterochromatin could

occur in the absence of SUV39H1-dependent H3K9me3. However, while e-HP1α-Ubc9 was more efficiently targeted than e-HP1α, the low fraction of cells showing this staining suggested that following this initial recruitment, retention at pericentric domains was rather inefficient. We thus modified our assay to monitor the localization of e-HP1α or e-HP1α-Ubc9 to pericentric chromatin in the presence of Myc-SUV39H1 assuming that H3K9me3 could promote stabilization (Fig. 4d). Immunofluorescence analysis revealed that e-HP1α, e-HP1α-Ubc9C93S and e-HP1α-Ubc9wt accumulated at pericentric heterochromatin when co-transfected with Myc-SUV39H1, in contrast to a negative control protein e-hnRNPC (Fig. 4e). Remarkably, e-HP1α-Ubc9wt localized to these domains more efficiently than e-HP1α and e-HP1α-Ubc9C93S (62% *vs* 37% and 39% of positive cells, respectively 6h post-transfection) for comparable levels of expressed proteins (Supplementary Fig. 8a). Also, the e-HP1α and e-HP1α-Ubc9C93S overall staining in the nucleus was rather diffuse compared to e-HP1α-Ubc9wt staining suggesting that a significant fraction of e-HP1α and e-HP1α-Ubc9C93S is not localized at pericentric heterochromatin (Fig. 4e, Supplementary Fig. 8b and supplemental legend for details). Taken together, these results indicate that enhancing sumoylation on HP1α promotes a more efficient accumulation at pericentric heterochromatin. We then performed a time course analysis to compare the appearance of HP1α at pericentric heterochromatin, in *Suv39h* dn cells co-transfected with Myc-SUV39H1 and e-HP1α or e-HP1α-Ubc9 (Fig. 4f). We found that e-HP1α-Ubc9 always localized more efficiently to pericentric heterochromatin compared to e-HP1α (52% vs 31% of positive cells, respectively, 4h post-transfection) for comparable levels of expressed proteins along the time course analysis (Fig. 4f). The different efficiencies with which these proteins localized to pericentric heterochromatin support the hypothesis of sumoylation acting as a limiting step to promote HP1α targeting. In support of this hypothesis, although toxic, the direct fusion of SUMO-1 to HP1α showed an even more efficient accumulation compared to e-HP1α-Ubc9 (Supplementary Fig. 9 and supplemental legend for details). The simplest interpretation is that a higher amount of sumoylated protein provided by HP1α-Ubc9 fusion would allow a rapid targeting to pericentric domains whereas HP1α alone would be delayed by the time needed to undergo an entire cycle of sumoylation. Overall our data underline the importance of the HP1α SUMO modification prior to HP1α targeting to pericentric heterochromatin.

Given that the hinge domain, which shows RNA-binding properties[12] is the target for sumoylation at multiple sites (Fig. 3b), we generated mutants of e-HP1α and e-HP1α-Ubc9 lacking the hinge domain and assayed whether they could specifically localize *de novo* to pericentric heterochromatin in *Suv39h* dn cells. After transfection of e-HP1αΔH and e-HP1αΔH-Ubc9 constructs in NIH3T3 cells, we found that the hinge domain was perfectly dispensable to localize HP1α at pre-existing HP1 domains in a 'maintenance assay' (Fig. 5a and b) as reported in *Drosophila* Kc167 cells[29]. Remarkably, when using *Suv39h* dn cells

for our '*de novo*' localization assay, after co-transfection with Myc-SUV39H1, the e-HP1α mutant lacking the hinge domain did not localize at pericentric heterochromatin, in contrast to the wild type protein which did accumulate at these domains (44% positive cells, Fig. 5c). We verified that all transfected proteins were expressed at comparable levels (Fig. 5d). Even when Ubc9 was fused to the mutant protein, we did not detect localization of e-HP1αΔH-Ubc9 at pericentric domains. These data clearly show that the hinge domain, the SUMO modification and the association with major RNAs are critical for *de novo* localization of HP1α at pericentric heterochromatin.

Based on our data, we propose a model for the *de novo* targeting and local accumulation of HP1α at pericentric heterochromatin by a multistep mechanism involving initial SUMO-dependent targeting as a "seeding" step (Fig. 6; "1"). This SUMO modification imposed in the hinge domain of HP1 would leave the CD and CSD available for other interactions with heterochromatin proteins (as represented in Fig 6). Subsequent "chromatin marking" steps would follow, including SUV39-dependent H3K9me3 to ensure HP1 stabilization (Fig. 6; "2") and accumulation through a self-enforcing loop (Fig. 6; "3")[1]. Our identification of a specific association between SUMO1-HP1α and major RNAs in the forward orientation (Fig. 2) provides a molecular basis for the "seeding" step. Such association may help to guide SUMO-HP1α specifically to pericentric heterochromatin. The specificity of the interaction promoted by sumoylation, either by the sequence, by particular structures formed by these RNAs or by know HP1 partners[9], remains to be determined. Importantly, the HP1-RNA interaction is specific for the forward strand, which is purine rich and may thus adopt a distinct structure. Furthermore, It will be interesting to examine potential SUMO-binding protein candidate[30]. Following the "seeding" event, coordination with SUV39-mediated H3K9 methylation would be key for HP1α stabilization. While this step could formally be independent of sumoylation, it is possible that SUMO-HP1 and RNA interaction may promote additional interactions with other partners (such as SUV39) and enhance SUV39 enzymatic activity. In this way, the seeding event might favor further stabilization steps to establish a robust system for the maintenance of heterochromatin domains. Given that several proteins involved in heterochromatin stability have been shown to bind SUMO conjugates[31], it is tempting to speculate that they could also bind SUMO-HP1. Future work should address how and where HP1 gets sumoylated and which enzymes trigger both its sumoylation and desumoylation as well as the impact on centromere function. Given the conserved importance of HP1 from fission yeast to mammals[3,32], further investigation of these issues in various organisms should help to define some general principles.

We detected forward RNAs as long species (several repeat lengths, Fig. 1a), and showed that they localize at the periphery of pericentric domains (Fig. 1b and Supplementary Fig. 2a), however we did not detect small dsRNA corresponding to major satellites, as described for the maintenance of heterochromatin

in *S. pombe*[1,33]. It is intriguing that even in *S. pombe* primal RNAs have been reported to be important for heterochromatin formation[34]. Although we do not exclude an RNAi pathway in connection with heterochromatin in mammalian cells[35,36], the fact that we detected long non-coding RNA that are stably located in the nucleus just at the periphery of major satellite domains is very compelling. This is reminiscent of other long nuclear non-coding RNA like the lincRNA HOTAIR, proposed to serve as a modular scaffold for histone modification complexes involved in Polycomb function[37] or Xist, which is critical to establish the silent state of the inactive X in mammals. Moreover, Xist RNA is known to be critical in setting up a *de novo* silent domain at specific times during development[38,39]. In this respect, the situation of early development in mice[40,41] is particularly interesting, since a burst of transcription of major RNAs occurs just prior to the formation of HP1 domains of accumulation on the paternal genome[42]. Thus, it is tempting to speculate that sumoylation could be critical also at this time and perhaps during other developmental time windows. This mechanism could apply to certain cell types when major rearrangements of the genome occur as observed during spermatogenesis[43], differentiation[44,45], reprogramming in the mouse germ line[46], in specialized cell types like Rod photoreceptor cells[47], or even in the formation of Senescence Associated Heterochromatin Foci (SAHF)[48].

Beyond HP1, the concept that a modification like SUMO imposed on a chromatin protein could promote localization guided by a particular transcript in the nucleus, should be explored. One may wonder whether similar mechanisms also apply to the formation of other local domains of accumulation or nuclear compartments such as polycomb or insulator bodies.

ACKNOWLEDGMENTS

We thank A. Bird, D. Shapiro, R. Hay, T. Jenuwein, R. Losson and J. Seeler for constructs and reagents. We thank W. Faigle for Mass Spectrometry support, G. Cappello for helpful discussions, A. Cook for critical reading and P. Le Baccon at the Curie Imaging Platform. GA's Laboratory is funded by la Ligue Nationale contre le Cancer (Equipe labellisée la Ligue), Curie PIC Programs, the European Network of Excellence Epigenome (LSHG-CT-2004-503433), ACI-2007-Cancéropôle IdF "Breast cancer and Epigenetics", ANR "ECenS" ANR-09-BLAN-0257-01, INCa "GepiG", ERC Advanced Grant 2009-AdG-20090506 and DL's Laboratory by le Cancéropôle Ile-de-France and l'INCA.

AUTHOR CONTRIBUTIONS

C.M., J.-P.Q and G.A. conceived and designed the experiments. C.M., D.B. and J.-P.Q. performed most of the experiments. D.R. and I.V. performed Immuno-DNA FISH and Immuno-RNA FISH. A.V.P.

performed RNA FISH. F.D., B.L. and D.L. performed and analyzed mass spectrometry data using samples prepared by R.M. and they wrote together the corresponding parts. C.M. generated all figures. C.M., J.-P.Q. and G.A. analyzed the data. C.M. and G.A. wrote the paper. All authors contributed to final editing of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

FIGURE LEGENDS

Figure 1. Strand-specific localization of centromeric RNAs.

a. Transcription from both strands of major and minor satellite repeats. Top: schematic representation of mouse major and minor satellite repeats and strand specific primers (forward [For], reverse [Rev]) used for analysis by RT-PCR. Bottom: corresponding results for RNAs isolated from NIH3T3 mouse cells. We show PCR reactions in the presence (+RT) or absence (-RT) of reverse transcriptase, or primers, as controls.

b. Nuclear localization of centromeric RNA compared to major and minor satellite DNA in NIH3T3 cells. Left: DNA FISH. Scheme of an acrocentric mouse chromosome with telomeres (black), major (red) and minor (green) satellites. Major (red) and minor (green) satellites are shown along with a merged image of major and minor satellites and DAPI stained DNA. Middle: RNA FISH. We localized major (forward in green, reverse in red) and minor (forward in red, reverse in green) RNAs with strand-specific LNA probes and show a merged image of DAPI and centromeric RNA staining. Insets show magnifications of chromocenters. Right: Immuno-RNA FISH. Forward major RNAs (green) and anti-HP1α antibodies (red) staining are shown along with a merged image of DAPI and major RNA staining. Scale bar, 10 μm.

c. HP1α accumulation at major satellite DNA domains. Immuno-DNA FISH with anti-HP1α antibodies (green) and major or minor satellite DNA (red) probes. Insets as in b. Scale bar, 10 μm.

d. RNAs associated with HP1α. ChIP experiments using pre-immune serum (Ctr) or anti-HP1α antibodies (HP1α) analysed by strand-specific RT-PCR as indicated. We show PCR reactions with (+RT) or without (-RT) reverse transcriptase, and without cDNA (mock), as controls. The input corresponds to the soluble chromatin prior to ChIP.

Figure 2. SUMO-1-modified HP1α interacts specifically with forward major RNAs.

a. Northwestern blot using recombinant GST-HP1γ, GST-HP1α or GST-HP1α domain fragments and *in vitro* transcribed radioactively labelled RNAs: forward (F) and reverse (R) major (Maj2) and minor (Min); U1 used as negative control.

b. Experimental scheme.

c. RNA pull down using forward major (Maj2 F) or minor (Min F) RNAs, or no RNA as negative control, in the absence of NEM. We show Western blot analysis with anti-vigilin, anti-RHA, anti-G3bp or anti-HP1α antibodies. Input is 10% of nuclear extracts.

d. RNA pull down using forward (F) and reverse (R) major (Maj2 F, Maj1 F, Maj2 R) or minor (Min F, Min R) RNAs as baits, or no RNA control, in the presence of NEM. Western blot analysis using anti-HP1α, anti-SUMO-1 and anti-SUMO-2/3 antibodies revealed endogenous unmodified HP1α (HP1α), modified-HP1α, SUMO-HP1α (S-HP1α) and free SUMO-2/3. Input is 10% of nuclear extracts. Asterisk (*) marks SUMO-HP1α in the input.

Figure 3. Sumoylation of HP1α occurs at its hinge domain *in vitro*.

a. Experimental scheme.

b. HP1α sumoylation *in vitro*. Left: Schematic representation of full-length HP1α and fragments thereof. Right: Western blot analysis of the sumoylation reaction mixture with anti-GST antibodies revealed the positions of SUMO-1-modified full-length HP1α (S-HP1α) or fragment domains (S-CD+H and S-H) marked by an asterisk (*).

c. HP1α-Ubc9 sumoylation *in vitro*. Western blot analysis of the sumoylation reaction mixture with anti-GST antibodies revealed the positions of SUMO-1-modified HP1α-Ubc9 (S-HP1α-Ubc9) and unmodified HP1α-Ubc9. Arrow indicates a degradation product of GST-HP1α-Ubc9.

d. Mass spectrometry analysis of the *in vitro* SUMO-1-modified HP1α hinge fragment. Shown are the MS (right) and MS/MS (left) fragmentation spectra of the tryptic peptide corresponding to residues 79-97 of SUMO-1 (top right, black) and 83-89 of HP1α (E<u>K</u>SEGNK; red) where K84 is sumoylated. The precursor ion mass was fragmented and acquired in Qstar (m/z 981.8 [3+]; left) and Orbitrap (m/z 981.77563, mass deviation 2 ppm; right, arrow) mass spectrometers. The majority of the fragment ions could be assigned to the y or b ion series, as annotated in the spectra and peptide sequence (top right).

e. Localization of e-HP1α and e-HP1α-Ubc9 in Triton extracted NIH3T3 cells. Left: Experimental scheme. Right: Immunofluorescence using anti-HA (red) and anti-HP1α (green) antibodies. Scale bar, 10 μm.

Figure 4. Sumoylation of HP1α promotes its targeting and accumulation at pericentric heterochromatin.

a. Experimental scheme.

b. Endogenous HP1α (red) and H3K9me3 (green) localization in wild-type and in *Suv39h* dn cells by immunofluorescence. Transfection of Myc-SUV39H1 in *Suv39h* dn cells restored HP1α and H3K9me3 localization in DAPI dense domains. Scale bar, 10 μm.

c. *de novo* localization of e-HP1α or e-HP1α-Ubc9 in *Suv39h* dn cells by immunofluorescence. Left: HA (red) and DAPI (green) staining with 3X magnification of selected chromocenters (arrows). Middle: HA (red) and H3K9me3 (green) staining. For each condition, we examined 300 transfected cells and calculated the percentage of cells with HA signal enriched (positive) or not (negative) at pericentric domains. Scale bar, 10 μm. Right: comparison of protein expression by Western blot revealing HA and β actin. Arrow indicates degradation of e-HP1α-Ubc9.

d. Experimental scheme.

e. *de novo* localization of e-hnRNPC, e-HP1α, e-HP1α-Ubc9C93S and e-HP1α-Ubc9wt in *Suv39h* dn cells co-transfected with Myc-SUV39H1 by immunofluorescence to reveal HA (red) and Myc (green). The percentage of positive cells was calculated as in c. Scale bar, 10 μm.

f. Time course analysis of the *de novo* localization of e-hnRNPC, e-HP1α and e-HP1α-Ubc9 in *Suv39h* dn cells co-transfected with Myc-SUV39H1. Top: the percentage of positive cells as a function of the time after transfection is represented. Symbols indicate the mean and error bars indicate the standard deviation of three independent experiments (300 co-transfected cells counted in each condition). Bottom: comparison of protein expression as in c. Arrow indicates degradation of e-HP1α-Ubc9.

Figure 5. The hinge domain is required for *de novo* localization of HP1α at pericentric heterochromatin.

a. Experimental scheme.

b. Localization of wild type (WT) or mutant (ΔH) e-HP1α and e-HP1α-Ubc9 in NIH3T3 cells by immunofluorescence using anti-HA antibodies (red) 24 h after transfection. Scale bar, 10 μm.

c. *de novo* localization of wild type (WT) or mutant (ΔH) e-HP1α and e-HP1α-Ubc9 in *Suv39h* dn cells co-transfected with Myc-SUV39H1 by immunofluorescence using anti-HA (red) and anti-Myc (green) antibodies 6 h after transfection. For each condition, we calculated the percentage of cells with HA signal enriched at pericentric domains (positive cells). Scale bar, 10 μm.

d. Comparison of protein expression corresponding to experiment in c by Western blot using anti-HA and anti-β actin antibodies. Arrows indicate degradation products of e-HP1α-Ubc9 and e-HP1αΔH-Ubc9 .

Figure 6. Model for a *de novo* HP1α targeting to pericentric heterochromatin.

A schematic representation of a nucleus with pericentric domains enriched in HP1 (red) is depicted showing the nuclear non-coding forward major RNA (green) at the periphery. HP1 (red) most likely as part of a complex becomes sumoylated. This SUMO-modified form of HP1α recognizes and binds to major RNAs (green) at pericentric heterochromatin, providing specificity to the initial targeting of HP1α to these domains (1). HP1α stabilization is then ensured by the recognition of H3K9me3 (blue) introduced by SUV39 (light brown) (2). Further HP1α accumulation involves a "self-enforcing" loop in which new HP1α directly binds to chromatin by multimerizing with other HP1α molecules, or by associating with other proteins and/or newly methylated H3K9 (3).

ONLINE METHODS

Mouse cell lines. We cultured wild type and *Suv39h* double-null MEFs (provided by T. Jenuwein)[27] and NIH3T3 cells (ATCC #CRL-1658) as described[2]. We transfected MEFs and NIH3T3 cells with Nucleofector Kit 2 (Amaxa) and Lipofectamine 2000 (Invitrogen) respectively, according to manufacturer instructions.

Plasmids. We obtained plasmids encoding GFP-SUMO-1 from R. Hay, Myc-SUV39H1 from T. Jenuwein[22] and GST-HP1α full-length protein and fragments thereof (CD, CD+H, H and CSD) from R. Losson[49]. We carried out cloning using standard PCR-based techniques. We made e-HP1α, e-Ubc9 and e-SUMO-1Δ6 constructs by inserting HP1α (gift of R. Losson), Ubc9 (gift of J. Seeler) and SUMO-1Δ6 (from pEGFP-SUMO-1) fused to a HA-tag in C-terminus into pcDNA5 vector (Invitrogen). We generated e-HP1α-Ubc9 and e-HP1α-SUMO-1 constructs by inserting the HP1α cDNA at the N-terminus of the e-Ubc9 and the e-SUMO-1Δ6 plasmids, respectively. To avoid conjugation of e-HP1α-SUMO-1 into other proteins, we removed the last six C-terminal amino acids of SUMO-1, which contain the diglycine motif required for isopeptide bond formation. We generated HP1α and e-HP1α-Ubc9C93S point mutants using the QuickChange site-directed mutagenesis kit (Stratagene). To generate e-HP1αΔH and e-HP1αΔ-Ubc9 mutants, we made truncation contructs of e-HP1α and e-HP1α-Ubc9 in which amino acids M67 to R117 forming the hinge domain were deleted and replaced by a linker of 2 amino acids (ID). These truncation contructs led to the fusion of the chromo domain to the chromoshadow domain. We generated GST-HP1α-hinge-chromoshadow domain and GST-HP1α-Ubc9 by subcloning from GST-HP1α full-length and e-HP1α-Ubc9, respectively. Each mutation and truncation was verified by sequencing.

Antibodies. We used: mouse monoclonal anti-HP1α (2HP-1H5-AS for immunofluorescence and 2HP-2G9-AS for Western blot; 1:1,000), anti-HP1β (1MOD-1A9-AS; 1:1,000), and anti-HP1γ (2MOD-1G6-

AS; 1:1,000) all from Euromedex, rabbit polyclonal anti-vigilin (from D. Shapiro[50]; 1:3,000), rabbit polyclonal anti-RHA (PA-001, Vaxron; 1:10,000), mouse monoclonal anti-SUMO-1 (#33-2400; 1:500) and rabbit polyclonal anti-SUMO-2/3 (#51-9100; 1:250) both from Zymed, rat monoclonal anti-HA (#1867423, Roche; 1:2,000 for Western blot and 1:250 for immunofluorescence), mouse monoclonal anti-G3bp (#611126, BD Biosciences; 1:1,000), mouse monoclonal anti-Myc (ab32, Abcam; 1:500), rabbit polyclonal anti-GST (ab9085 Abcam; 1:1,000), rabbit polyclonal anti-GFP (sc-8334, Santa-Cruz; 1:500), rabbit polyclonal anti-H3K9me3 (#07-442, Upstate; 1:500), mouse monoclonal anti-β actin (#A5441 Sigma; 1:20,000). For chromatin immunoprecipitations we used a rabbit polyclonal anti-HP1α antibody generated against the full length GST-HP1α protein (Agro-Bio).

Nuclear extracts. After incubation in hypotonic buffer A (20 mM Hepes-KOH pH 7.8, 5 mM potassium acetate, 0.5 mM $MgCl_2$, 0.5 mM DTT) for 10 min at 4°C, we disrupted NIH3T3 cells by 25 strokes with a dounce homogenizer and separated nuclei from the soluble proteins by centrifugation at 1600 g. After incubation of nuclear pellets in buffer A containing 615 mM NaCl for 1.5 h at 4°C, followed by centrifugation at 14 000g for 20 min, we collected supernatant as the nuclear extract that we aliquoted and flash froze in liquid nitrogen. All buffers contained protease and phosphatase inhibitors (10 µg/ml pepstatin, 10 µg/ml leupeptin, 100 µM PMSF, 5 mM sodium fluoride, 10 mM β-glycerophosphate, plus or minus 20 mM N-ethylmaleimide [NEM; Sigma] when indicated).

Centromeric RNA pull down. The Maj9-2 and Min5-1 pCR4 plasmids contain 542-bp of the mouse major satellite DNA and 162-bp of minor satellite DNA, respectively (provided by T. Jenuwein)[51]. We subcloned a cDNA encoding a 234-bp repeat unit of the mouse major satellite DNA from pUC19-Sat15 (provided by A. Bird) into the pBS vector (Stratagene). We obtained biotinylated major (Maj1 from Sat15-pBS and Maj2 from Maj9-2 pCR4) and minor (Min from Min5-1 pCR4) RNAs by *in vitro* transcription with T7 or T3 RNA polymerases (Promega) in the presence of biotin RNA labelling mix (Roche), at 37°C for 2 h. After DNA digestion by RNase-free DNase I (Promega), we removed un-incorporated rNTPs by a Sephadex G-50 quick spin column (Roche). For RNA pull down, we incubated 2 µg of biotinylated RNA with nuclear extracts in binding buffer (20 mM Hepes pH 7.6, 100 mM KCl, 2 mM EDTA, 0.01% Nonidet P-40, 1% gelatin) with 200 µg/ml tRNA, 4 mg/ml heparin, 80 U RNasin, plus or minus 20 mM NEM as indicated, for 30 min at room temperature. For each binding reaction we used 100 µl of streptavidin-coated magnetic beads (Dynabeads; Invitrogen) for 1 h at room temperature, on a rotating wheel. After 4 washes with binding buffer containing 20 µg/ml tRNA and 0.1% Tween-20, followed by one more wash with 20 mM Hepes pH 7.6, plus 50 mM KCl, we eluted bound proteins with

SDS-PAGE loading buffer and ran samples on NuPAGE 4-12% Bis-Tris gels (Invitrogen) with MOPS running buffer (Invitrogen). We stained gels with Coomassie brilliant blue.

Immunoprecipitations and Western blotting. We lysed NIH3T3 cells 48 h post-transfection, with lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 15 mM $MgCl_2$, 1% Nonidet P-40, and 0.75% sodium deoxycholate) supplemented with protease and phosphatase inhibitors, and 20 mM NEM (Sigma). We incubated cell lysates corresponding to $4x10^6$ cells with 40 µl of monoclonal anti-HA agarose-conjugated beads (Roche) for 2 h at 4°C. After washing the beads with lysis buffer, we eluted the immunocomplexes with SDS-PAGE loading buffer, and resolved proteins by 4-12% Bis-Tris NuPAGE gels (Invitrogen), and transferred to nitrocellulose membranes (Protran). For visualization of proteins after Western blots with the indicated antibodies we used the Super Signal detection kit (Pierce).

Northwestern blotting. We resolved recombinant full-length GST-HP1γ and GST-HP1α full-length and fragments thereof by SDS-PAGE and transferred them to nitrocellulose. We incubated the membranes with *in vitro* transcribed radioactively labelled RNAs in buffer containing 20 mM Hepes pH 7.6, 100 mM KCl, 2 mM EDTA, 0.01% Nonidet P-40, overnight at 23°C. After 3 washes with the same buffer, we visualized bound radioactively labelled RNAs by autoradiography.

Immunofluorescence microscopy. We processed cells for immunostaining as described[11]. We used an epifluorescence photo-microscope (DM6000B, Leica) piloted with Metamorph software, an x63 (NA 1.32) objective lens and an HQ2 CoolSnap camera (Photometrics) for image acquisition. For all time course studies, we performed three independent experiments. For spot formation analysis, after image acquisition, we draw a scan line across the nucleus and measured the relative intensity of fluorescence across this line for signals corresponding to e-HP1α and Myc-SUV39H1 (ImageJ software). For Immuno-RNA FISH, we acquired 50 optical sections separated by 0.2 µm with an Imager.Z1 microscope (Zeiss) piloted with Metamorph software and made Z projections. We quantified the localization of forward major transcripts at the periphery of pericentric heterochromatin domains from 3D image series (z-step 0.2 µm) of major RNA FISH/DAPI staining acquired on a Delta Vision system (Applied Precision, 100x objective). For each nucleus, we manually scored the total number of RNA foci (f_{tot}) and the number of RNA foci at the periphery of DAPI dense pericentric domains (f_{per}) using ImageJ software and the Image 5D plugin to allow co-visualization in 3D of the RNA FISH and DAPI signals. We determined the volume of the nuclei (V_{nuc}) and the volume of the individual pericentric domains (V_{dom}, DAPI dense) using the ImageJ software and the 3D Object Counter plugin. We calculated the volume of peripheral pericentric domain (V_{per}) as the difference between the volume corresponding to the pericentric domain with an

increase of 20% of the radius of the domain ($V_{dom20\%}$, assuming a spherical form) and the volume of the domain ($V_{per}=V_{dom20\%}-V_{dom}$). We obtained the concentration of RNA foci (foci/ μm^3) at the periphery from the ratio f_{per}/V_{per} and in the nucleus from the ratio $f_{tot}/(V_{nuc}-V_{dom})$. We removed the volume of the pericentric domains from total volume given that RNA foci are never found within domains.

RNA and DNA FISH. For RNA FISH, after extraction with 0.5% Triton X-100 in CSK buffer (10 mM Pipes pH7, 100 mM NaCl, 300 mM sucrose, 3 mM MgCl2, supplemented with 10 mM Vanadyl ribonucleoside complex (VRC)) for 5min on ice, we fixed cells in 3% paraformaldehyde in PBS for 12 min and stored them in 70% EtOH at -20°C overnight. Following dehydration in 80%, 95% and 100% EtOH, we carried out hybridization with 0.4 µM locked nucleic acid (LNA) fluorescent probes (Exiqon) in 50% formamide (Sigma), 2x SSC (Sigma), 10% dextran sulfate (Fluka), 10 mM VRC (NEB) and 2 mg/mL BSA (NEB) in a humid chamber for 35 min at 37°C. After 3 washes in 0.1x SSC for 5 min at 60°C, we stained DNA with DAPI staining and mounted the cells in Vectashield (Vector Laboratories). For Immuno-RNA FISH, we performed RNA FISH as described above. After post-hybridization washes, we post-fixed cells in 3% paraformaldehyde in PBS for 12 min and processed them for immunostaining. DNA FISH and Immuno-DNA FISH were performed as described[2], except that the hybridization mix contained LNA fluorescent probes (0.1 µM) and we performed post-hybridization washes in 0.1X SSC (3 times for 5 min) at 60°C. The sequences of the LNA fluorescent probes are listed in Supplementary Table 2.

RT-PCR analysis. We performed RT-PCR analysis on HP1α-associated RNAs and on total RNA extracted from 3T3 cells with Trizol (Invitrogen). We digested genomic DNA by incubation with DNase 1 (Sigma). We synthesized first strand cDNA from 1 µg RNA in 20 µl buffer containing 1 µM forward (For) or reverse (Rev) specific primers for major or minor satellites, 0.5 mM dNTPs, 40 U/µl RNasin and 10 U/µl of SuperScript II reverse transcriptase (Qiagen). We then amplified the generated cDNA by PCR using 1/2000 and 1/250 dilutions of major and minor cDNA respectively, and a PCR Master Kit (Roche) supplemented with 0.5 µM specific primers during 45 cycles. Primer sequences are listed in Supplementary Table 2.

Chromatin Immunoprecipitation (ChIP). We extracted NIH3T3 cells with Triton X-100 to remove soluble proteins as previously described[52] and cross-linked them with 1.5% formaldehyde in PBS for 20 min at room temperature. We then added 125 mM glycine for 20 min to quench any residual formaldehyde. After one wash with PBS, we collected the cells by scraping and resuspended them in ChIP buffer (20 mM Tris-HCl at pH 7.5, 200 mM NaCl, 0.2% Nonidet P-40, 1% gelatin and 80 U RNasin). We

then sonicated the cells seven times for 30 seconds on high intensity (Bioruptor sonicator, CosmoBio Inc.) and centrifuged at 10 000g for 5 min. We incubated approximately 200-300 µg of supernatant containing the soluble chromatin with 20 µl of sera against HP1α raised in the laboratory for 3 h at 4°C. We used pre-immune sera as negative control. Then, we added 100µl of Protein-A sepharose slurry (50% W/V, Amersham Biosciences) and incubated for 3 h at 4°C on a rotating wheel. We recovered the immunoprecipitated chromatin by centrifugation, washed it five times with 1 ml of ChIP buffer and resuspended it in 10 mM Tris-HCl at pH 7.5, 1 mM EDTA, 0.2% SDS. We reversed the crosslink by incubation at 65°C overnight. We then extracted the HP1α-associated RNAs with Trizol (Invitrogen) and we performed RT-PCR analysis.

***In vitro* sumoylation assays.** We produced recombinant proteins in *E. coli* by expressing constructs corresponding to the GST-HP1α full-length protein and fragments thereof and GST-HP1α-Ubc9. We utilized the recombinant proteins in *in vitro* sumoylation reactions, using the SUMOlink kits (Active Motif), using manufacturer instructions. We validated sumoylation of HP1α and HP1α-Ubc9 by Western blot with anti-GST antibodies. For analysis of the sumoylated GST-HP1α hinge fragment by mass spectrometry, we used a total of 5 µg of GST-HP1α hinge and 10 µg of SUMO-1 protein. We ran 90% of the reaction mixture on a 4-12% Bis-Tris NuPAGE gel (Invitrogen) and stained the gel overnight with LabSafe GEL Blue stain (Biosciences) for later mass spectrometry analysis. We used the remaining 10% of the sumoylation mixture for Western blotting with rabbit anti-GST and rabbit anti-SUMO-1 (Active Motif; 1:4,000 dilution) antibodies.

Mass Spectrometry. We reduced, alkylated and trypsin digested slices (1 mm wide) cut out from Coomassie blue-stained gels as previously described[53]. We dried extracted peptides, resolubilized them in solvent A (95/5 water/acetonitrile, 0.1% formic acid) prior to liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis. We concentrated and separated them on a LC-Packings system (Dionex S.A.) coupled to the nano-electrospray II ionization interface of a QSTAR Pulsar i (Applied Biosystems/MDS Sciex). HPLC mobile phases contained solvent A and solvent B [20/80: water/acetonitrile, 0.085% formic acid]. We eluted bound peptides with a gradient of 5-50% of solvent B. We used information-dependent acquisition (IDA) to acquire MS/MS data, with experiments designed such that the two most abundant peptides were subject to collision-induced dissociation (CID). We analyzed twice the data from the IDA experiments by using MASCOT software (Matrix Science, London) on an internal server, first without taxonomic restriction to reveal the presence of proteins of interest and mammalian contaminants, then again the "*Mus musculus"* (mouse) database of the National Center for Biotechnology Information nr (National Library of Medicine, Bethesda, 2005 05 02, 2 452 561 and 41 362

protein entries). We converted the RAW files to the Mascot Generic Format (MGF) and submitted them to the Mascot search engine (version 1.0). We used the following parameters in the database search: Full trypsin enzyme specificity; Missed cleavages allowed = 1, Peptide mass tolerance = 0.8 Da; Fragment ion tolerance = 1 Da; Monoisotopic molecular weight for both peptide and fragment ion masses, b/y ion search, and fixed carbamidomethyl cysteine modification. We manually validated all data using myProMS[54].

For SUMO-1 LC-MS/MS experiments, we used two different MS platforms in parallel. We achieved peptide concentration and separation using an actively split capillary HPLC system (Ultimate 3000, Dionex, Germering, Germany) connected to each MS platform. The first platform was a quadrupole time-of-flight (Q-TOF) mass spectrometer (QSTAR Elite, Applied Biosystems/MDS Sciex). We acquired a TOF-MS survey scan for 1 sec over a mass range of 800-1200 m/z. We used an IDA method to acquire product ion scans on the three most intense 3+ ions per cycle over a mass range of 65–2000 m/z, excluding previously gated ions for 60 sec. We used a Smart setting of 2.0. The second platform was an LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nanospray source using a Pico-Tip (10 µm i.d., New Objectives). We set the spray voltage to 2.2 kV and the temperature of the heated capillary to 200°C. The mass spectrometer was operated in data-dependent mode to automatically acquire MS and MS/MS spectra. We acquired full scan survey spectra (*m*/*z* 615-1200) in the Orbitrap with a resolution of 100,000 at *m*/*z* 400 after accumulation of 1,000,000 charges. We sequentially isolated the five most intense ions and fragmented them in the linear ion trap by collision-induced dissociation after accumulation of 30,000 ions (normalized collision energy 35%). Maximum inject times were 500 msec for full scans and 200 msec for MS/MS scans. Dynamic exclusion was enabled with exclusion duration of 120 sec. We calculated and used the monoisotopic m/z values for SUMO-1-GST-HP1α hinge branched precursor peptides as described[55] to search for the corresponding ions (assignment was confirmed by manually interpreting all MS/MS spectra). We manually validated all reported MS/MS spectra. We considered only branched peptides having an extensive coverage of b and/or y ions. It was assumed that modified (sumoylated) lysines cannot be cleaved by trypsin, and one trypsin missed cleavage was allowed.

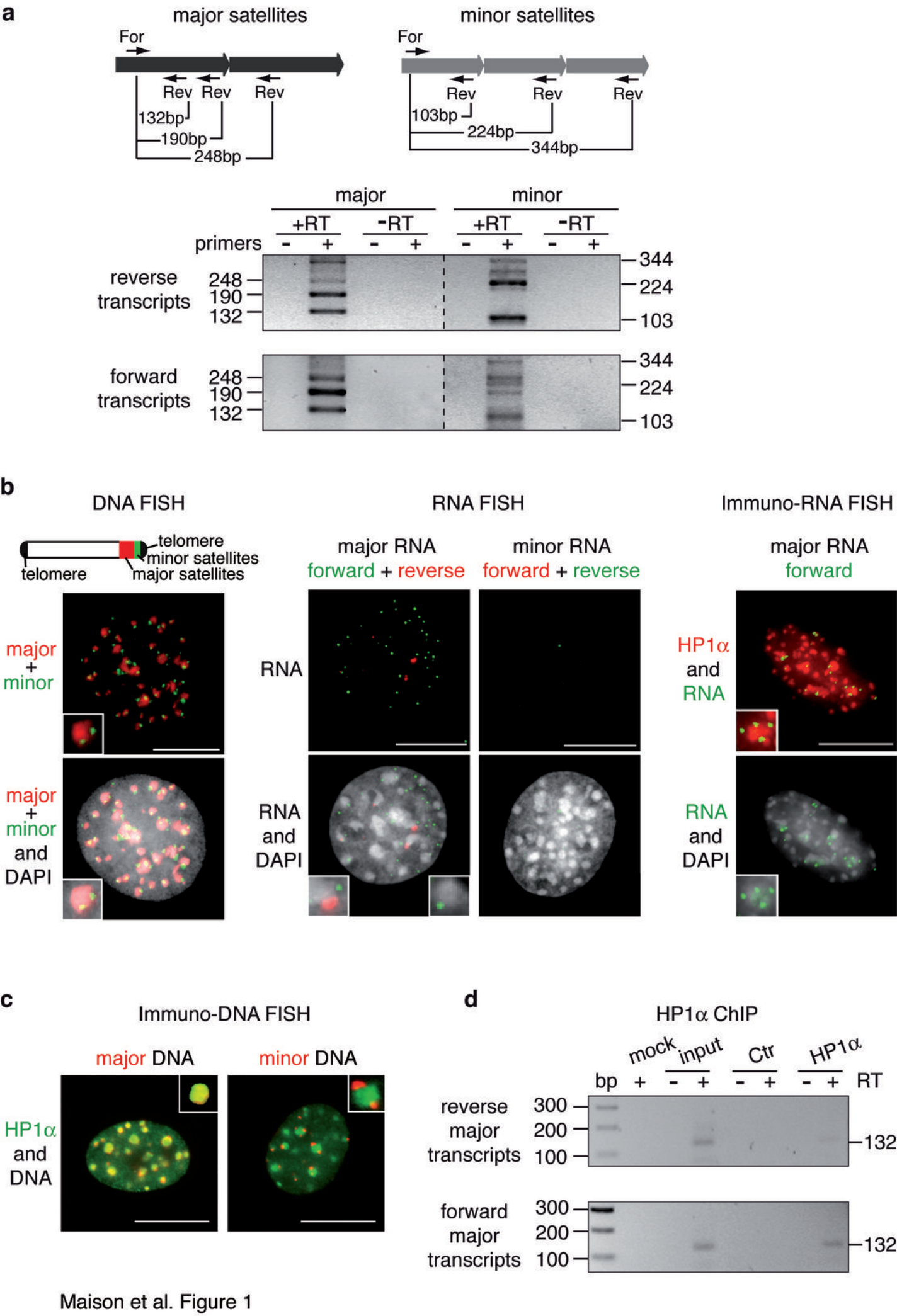
REFERENCES

1. Grewal, S.I. & Jia, S. Heterochromatin revisited. *Nat Rev Genet* **8**, 35-46 (2007).
2. Guenatri, M., Bailly, D., Maison, C. & Almouzni, G. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol* **166**, 493-505 (2004).
3. Probst, A.V., Dunleavy, E. & Almouzni, G. Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol* **10**, 192-206 (2009).

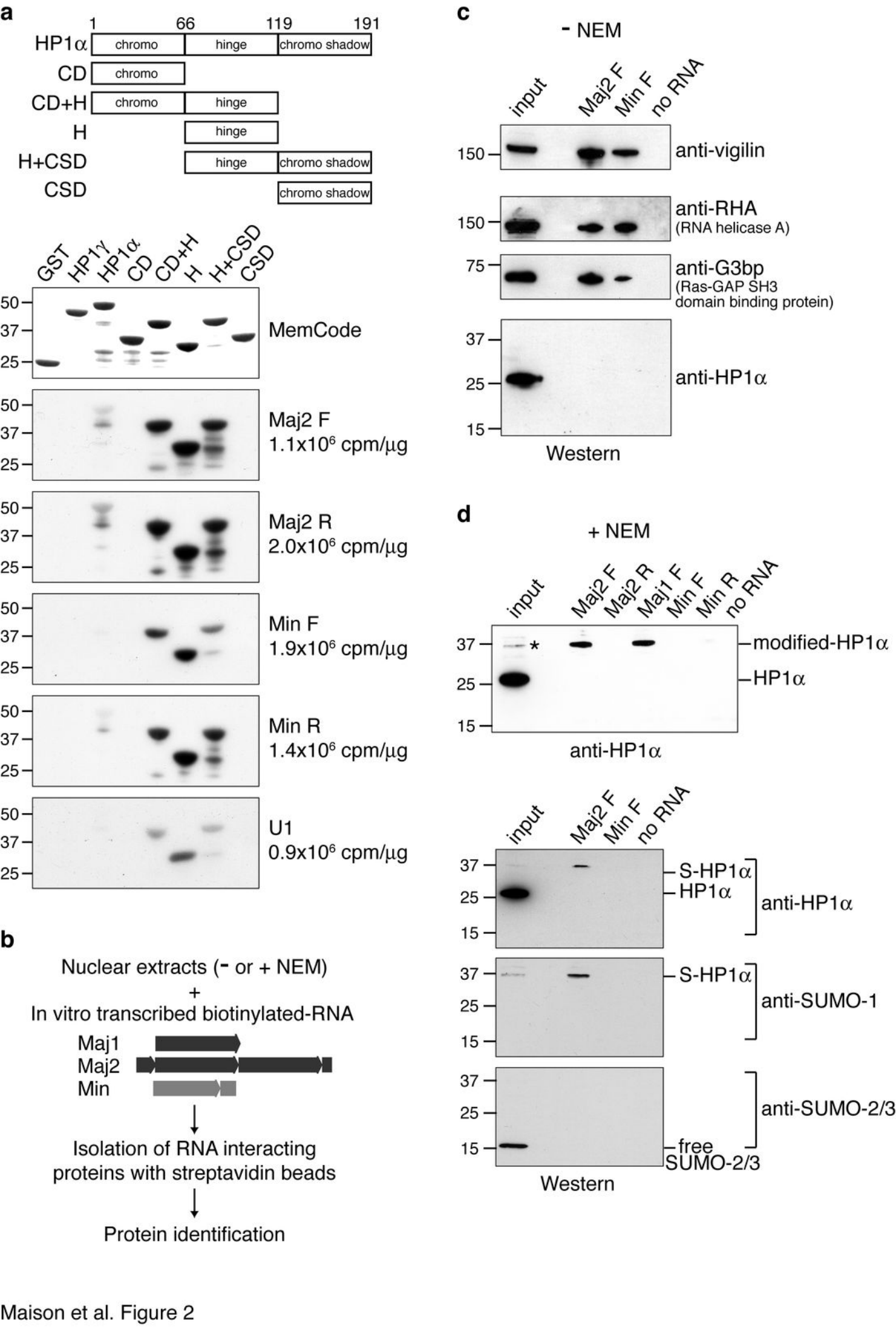
4. Bannister, A.J. et al. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-4 (2001).
5. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-20 (2001).
6. Strahl, B.D. & Allis, C.D. The language of covalent histone modifications. *Nature* **403**, 41-5 (2000).
7. Jenuwein, T. & Allis, C.D. Translating the histone code. *Science* **293**, 1074-80 (2001).
8. Peng, H., Ivanov, A.V., Oh, H.J., Lau, Y.F. & Rauscher, F.J., 3rd. Epigenetic gene silencing by the SRY protein is mediated by a KRAB-O protein which recruits the KAP1 co-repressor machinery. *J Biol Chem* (2009).
9. Quivy, J.P. et al. A CAF-1 dependent pool of HP1 during heterochromatin duplication. *EMBO J* **23**, 3516-26 (2004).
10. Lomberk, G., Bensi, D., Fernandez-Zapico, M.E. & Urrutia, R. Evidence for the existence of an HP1-mediated subcode within the histone code. *Nat Cell Biol* **8**, 407-15 (2006).
11. Maison, C. et al. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat Genet* **30**, 329-34 (2002).
12. Muchardt, C. et al. Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. *EMBO Rep* **3**, 975-81 (2002).
13. Rudert, F., Bronner, S., Garnier, J.M. & Dolle, P. Transcripts from opposite strands of gamma satellite DNA are differentially expressed during mouse development. *Mamm Genome* **6**, 76-83 (1995).
14. Lu, J. & Gilbert, D.M. Proliferation-dependent and cell cycle regulated transcription of mouse pericentric heterochromatin. *J Cell Biol* **179**, 411-21 (2007).
15. Kalitsis, P. & Choo, K.H.A. Centromere DNA of higher eukaryotes. in *The Centromere* (ed. Choo, K.H.A.) 97-140 (Oxford University Press, New York, 1997).
16. Heard, E. & Bickmore, W. The ins and outs of gene regulation and chromosome territory organisation. *Curr Opin Cell Biol* **19**, 311-6 (2007).
17. Deng, Z., Norseen, J., Wiedmer, A., Riethman, H. & Lieberman, P.M. TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol Cell* **35**, 403-13 (2009).
18. Wang, Q., Zhang, Z., Blackwell, K. & Carmichael, G.G. Vigilins bind to promiscuously A-to-I-edited RNAs and are involved in the formation of heterochromatin. *Curr Biol* **15**, 384-91 (2005).
19. Kuroda, M.I., Kernan, M.J., Kreber, R., Ganetzky, B. & Baker, B.S. The maleless protein associates with the X chromosome to regulate dosage compensation in Drosophila. *Cell* **66**, 935-47. (1991).
20. Irvine, K., Stirling, R., Hume, D. & Kennedy, D. Rasputin, more promiscuous than ever: a review of G3BP. *Int J Dev Biol* **48**, 1065-77 (2004).
21. Shin, J.A. et al. SUMO modification is involved in the maintenance of heterochromatin stability in fission yeast. *Mol Cell* **19**, 817-28 (2005).
22. Aagaard, L. et al. Functional mammalian homologues of the Drosophila PEV- modifier Su(Var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. *EMBO J* **18**, 1923-1938 (1999).
23. Hay, R.T. SUMO: a history of modification. *Mol Cell* **18**, 1-12 (2005).
24. Meulmeester, E. & Melchior, F. Cell biology: SUMO. *Nature* **452**, 709-11 (2008).
25. Park-Sarge, O.K. & Sarge, K.D. Detection of sumoylated proteins. *Methods Mol Biol* **464**, 255-65 (2009).
26. Jakobs, A. et al. Ubc9 fusion-directed SUMOylation (UFDS): a method to analyze function of protein SUMOylation. *Nat Methods* **4**, 245-50 (2007).
27. Peters, A. et al. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323-337 (2001).

28. Loyola, A., Bonaldi, T., Roche, D., Imhof, A. & Almouzni, G. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol Cell* **24**, 309-16 (2006).
29. Smothers, J.F. & Henikoff, S. The hinge and chromo shadow domain impart distinct targeting of HP1-like proteins. *Mol Cell Biol* **21**, 2555-69 (2001).
30. Kerscher, O. SUMO junction-what's your function? New insights through SUMO-interacting motifs. *EMBO Rep* **8**, 550-5 (2007).
31. Ouyang, J., Shi, Y., Valin, A., Xuan, Y. & Gill, G. Direct binding of CoREST1 to SUMO-2/3 contributes to gene-specific repression by the LSD1/CoREST1/HDAC complex. *Mol Cell* **34**, 145-54 (2009).
32. Allshire, R.C. & Karpen, G.H. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat Rev Genet* **9**, 923-37 (2008).
33. Moazed, D. Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457**, 413-20 (2009).
34. Blankenship, J.T. & Wieschaus, E. Two new roles for the Drosophila AP patterning system in early morphogenesis. *Development* **128**, 5129-38. (2001).
35. Kanellopoulou, C. et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* **19**, 489-501 (2005).
36. Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S. & Hannon, G.J. Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A* **102**, 12135-40 (2005).
37. Tsai, M.C. et al. Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science* (2010).
38. Savarese, F., Flahndorfer, K., Jaenisch, R., Busslinger, M. & Wutz, A. Hematopoietic precursor cells transiently reestablish permissiveness for X inactivation. *Mol Cell Biol* **26**, 7167-77 (2006).
39. Agrelo, R. et al. SATB1 defines the developmental context for gene silencing by Xist in lymphoma and embryonic cells. *Dev Cell* **16**, 507-16 (2009).
40. Santos, F., Peters, A.H., Otte, A.P., Reik, W. & Dean, W. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev Biol* **280**, 225-36 (2005).
41. Puschendorf, M. et al. PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat Genet* **40**, 411-20 (2008).
42. Probst, A.V. et al. A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev Cell* **19**, 625-38 (2010).
43. Haaf, T. & Ward, D.C. Higher order nuclear structure in mammalian sperm revealed by in situ hybridization and extended chromatin fibers. *Exp Cell Res* **219**, 604-11 (1995).
44. Mayer, R. et al. Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biol* **6**, 44 (2005).
45. Terranova, R., Sauer, S., Merkenschlager, M. & Fisher, A.G. The reorganisation of constitutive heterochromatin in differentiating muscle requires HDAC activity. *Exp Cell Res* **310**, 344-56 (2005).
46. Hajkova, P. et al. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452**, 877-81 (2008).
47. Solovei, I. et al. Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell* **137**, 356-68 (2009).
48. Narita, M. et al. Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* **113**, 703-16 (2003).
49. Nielsen, A.L. et al. Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol Cell* **7**, 729-39 (2001).
50. Dodson, R.E. & Shapiro, D.J. Vigilin, a ubiquitous protein with 14 K homology domains, is the estrogen-inducible vitellogenin mRNA 3'-untranslated region-binding protein. *J Biol Chem* **272**, 12249-52 (1997).
51. Lehnertz, B. et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13**, 1192-200 (2003).

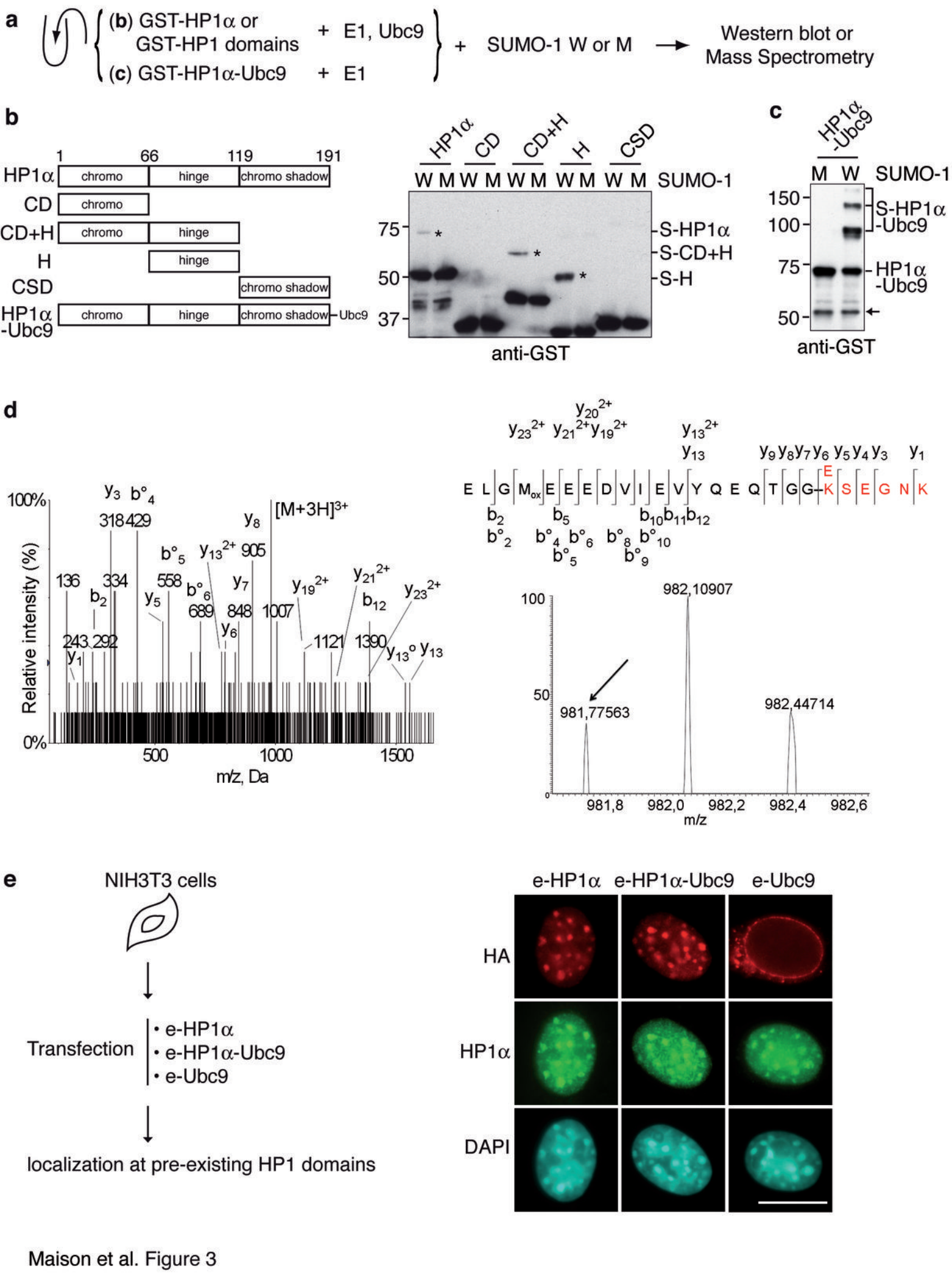
52. Martini, E., Roche, D.M., Marheineke, K., Verreault, A. & Almouzni, G. Recruitment of phosphorylated chromatin assembly factor 1 to chromatin after UV irradiation of human cells. *J Cell Biol* **143**, 563-75 (1998).
53. Fevrier, B. et al. Cells release prions in association with exosomes. *Proc Natl Acad Sci U S A* **101**, 9683-8 (2004).
54. Poullet, P., Carpentier, S. & Barillot, E. myProMS, a web server for management and validation of mass spectrometry-based proteomic data. *Proteomics* **7**, 2553-6 (2007).
55. Matic, I. et al. In vivo identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. *Mol Cell Proteomics* **7**, 132-44 (2008).



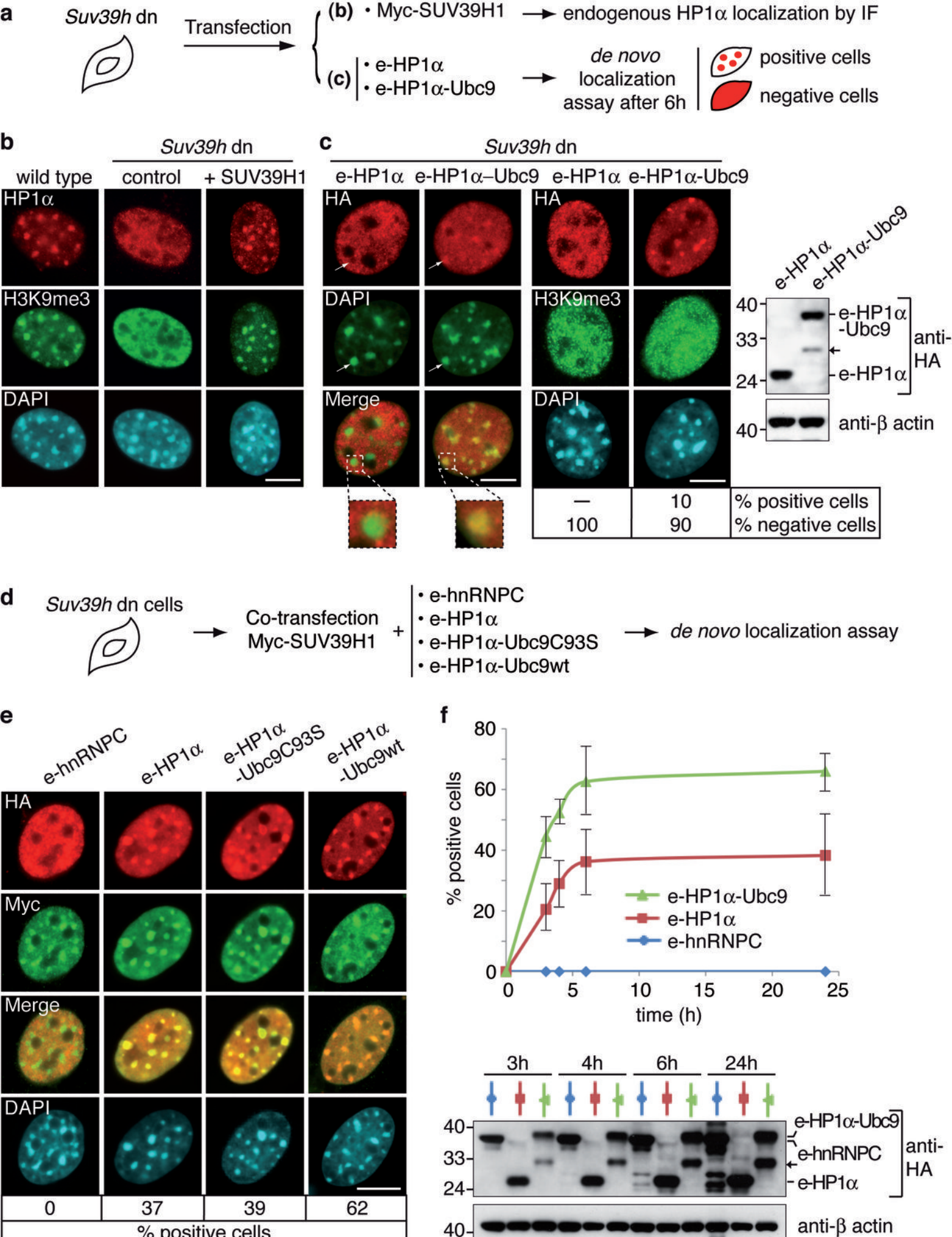
Maison et al. Figure 1



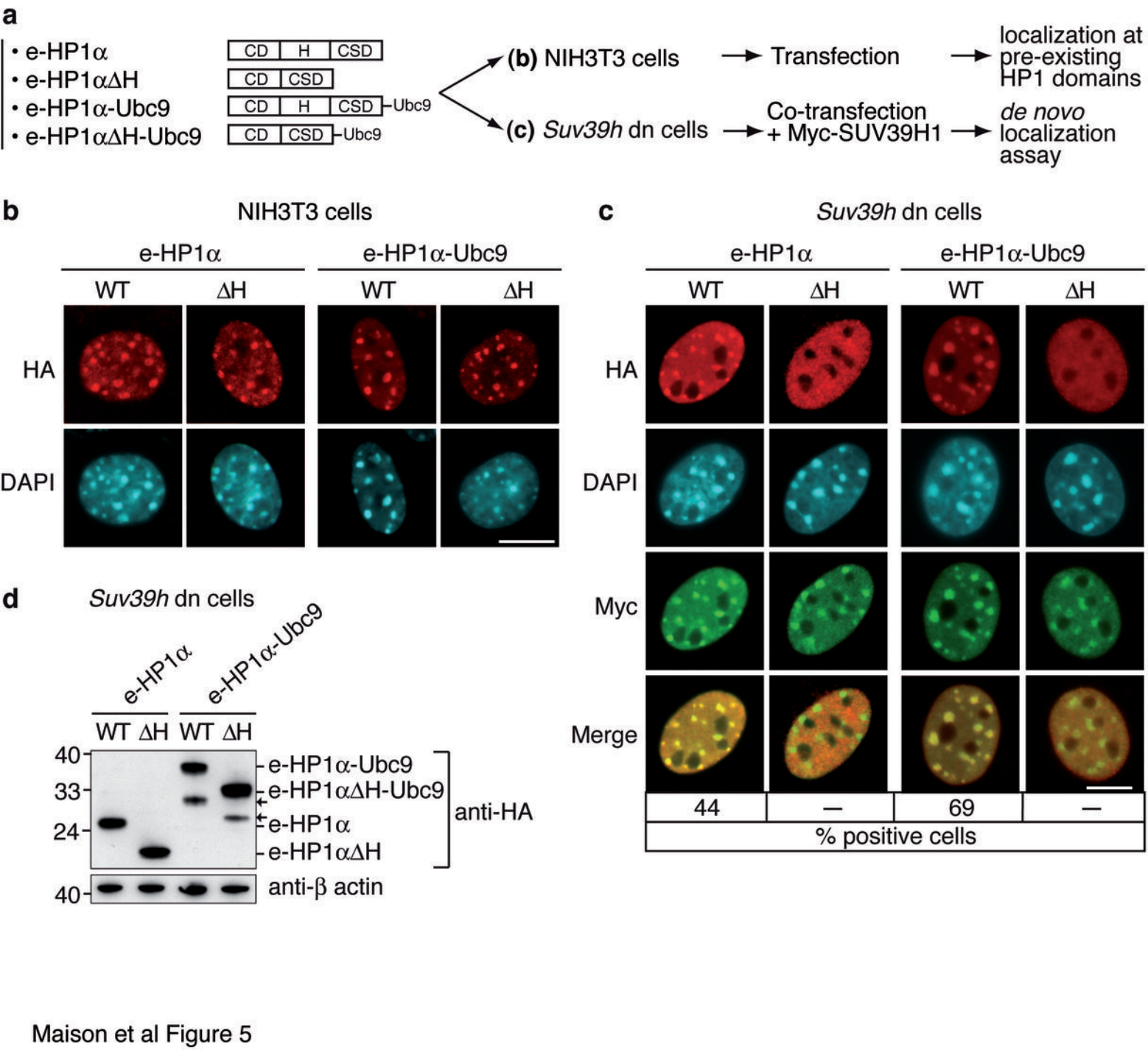
Maison et al. Figure 2



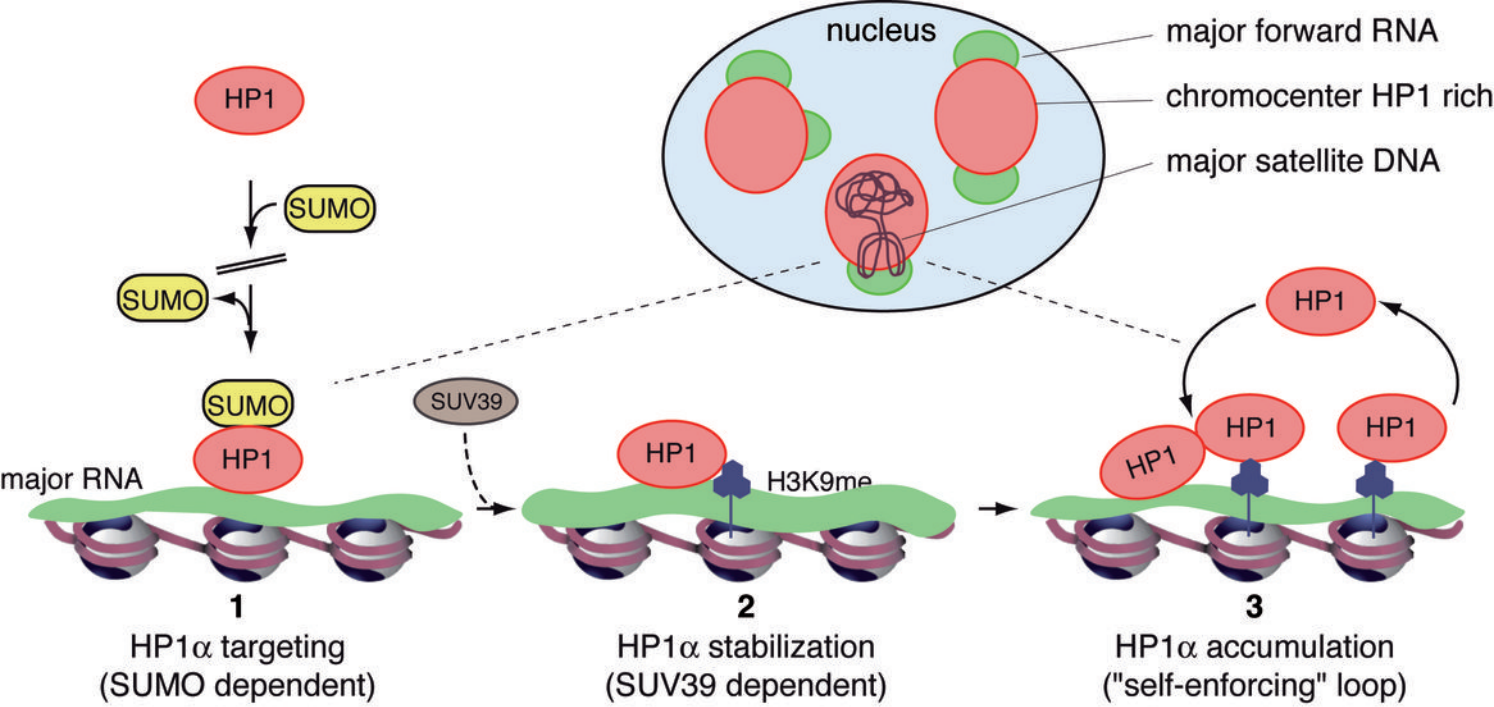
Maison et al. Figure 3



Maison et al Figure 4



Maison et al Figure 5

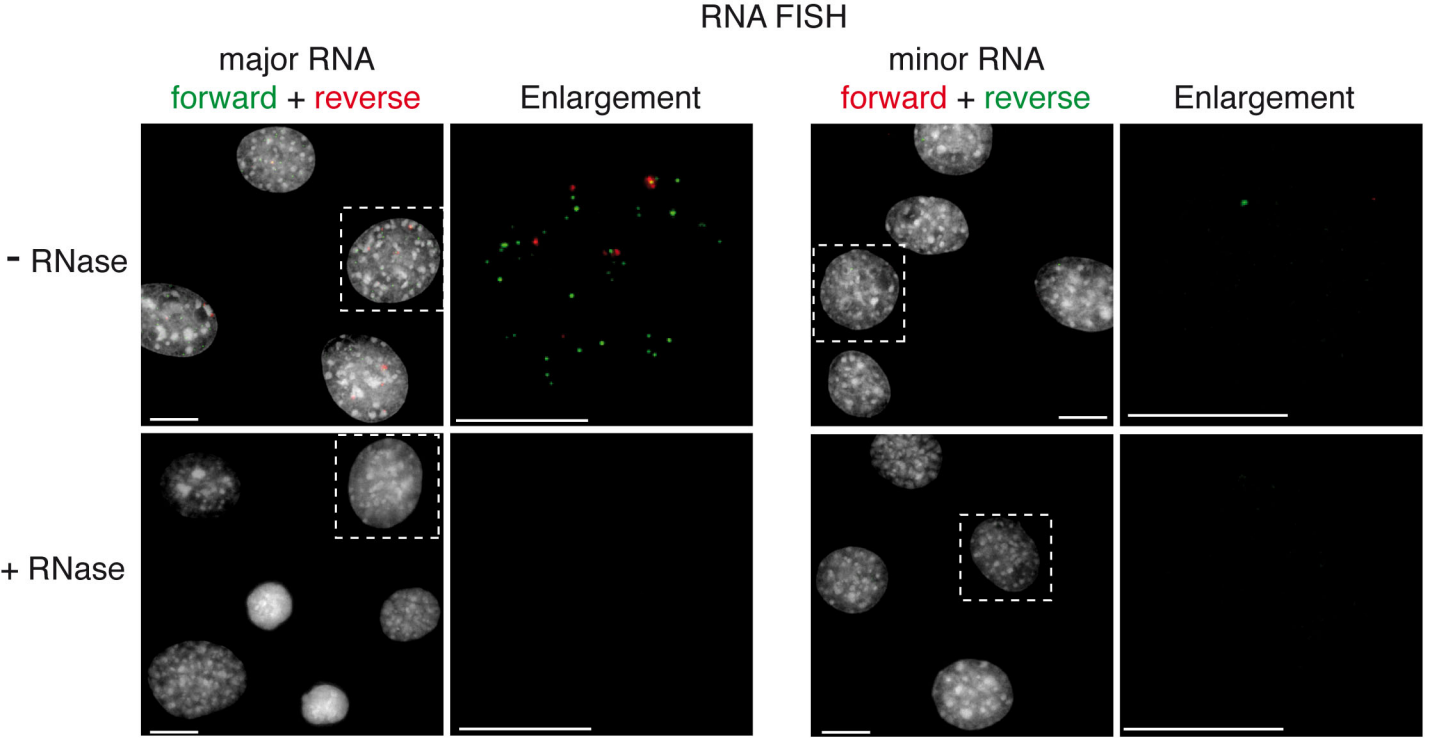


Maison et al Figure 6

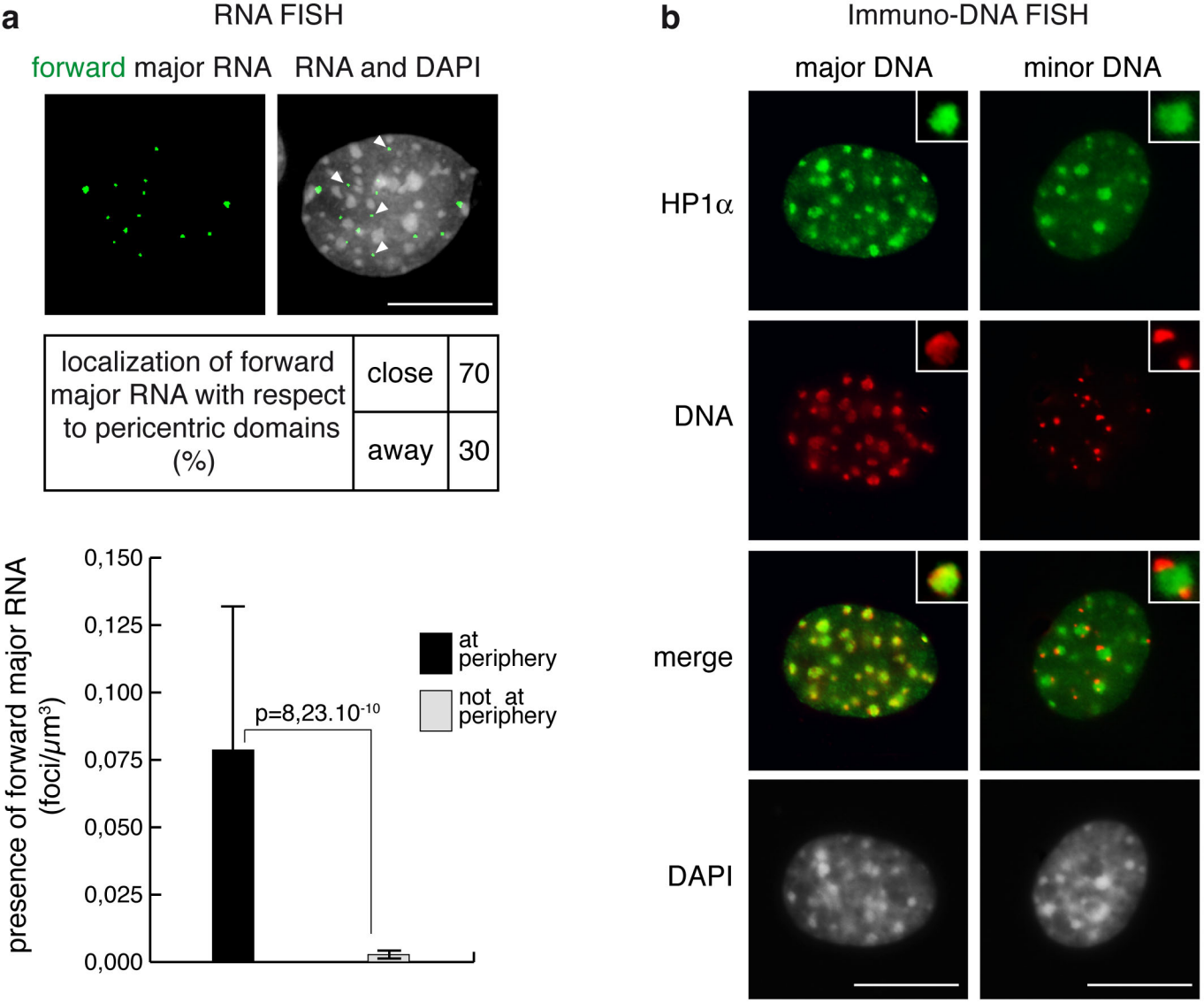
SUPPLEMENTARY DATA

SUMOylation promotes *de novo* targeting of HP1α to pericentric heterochromatin

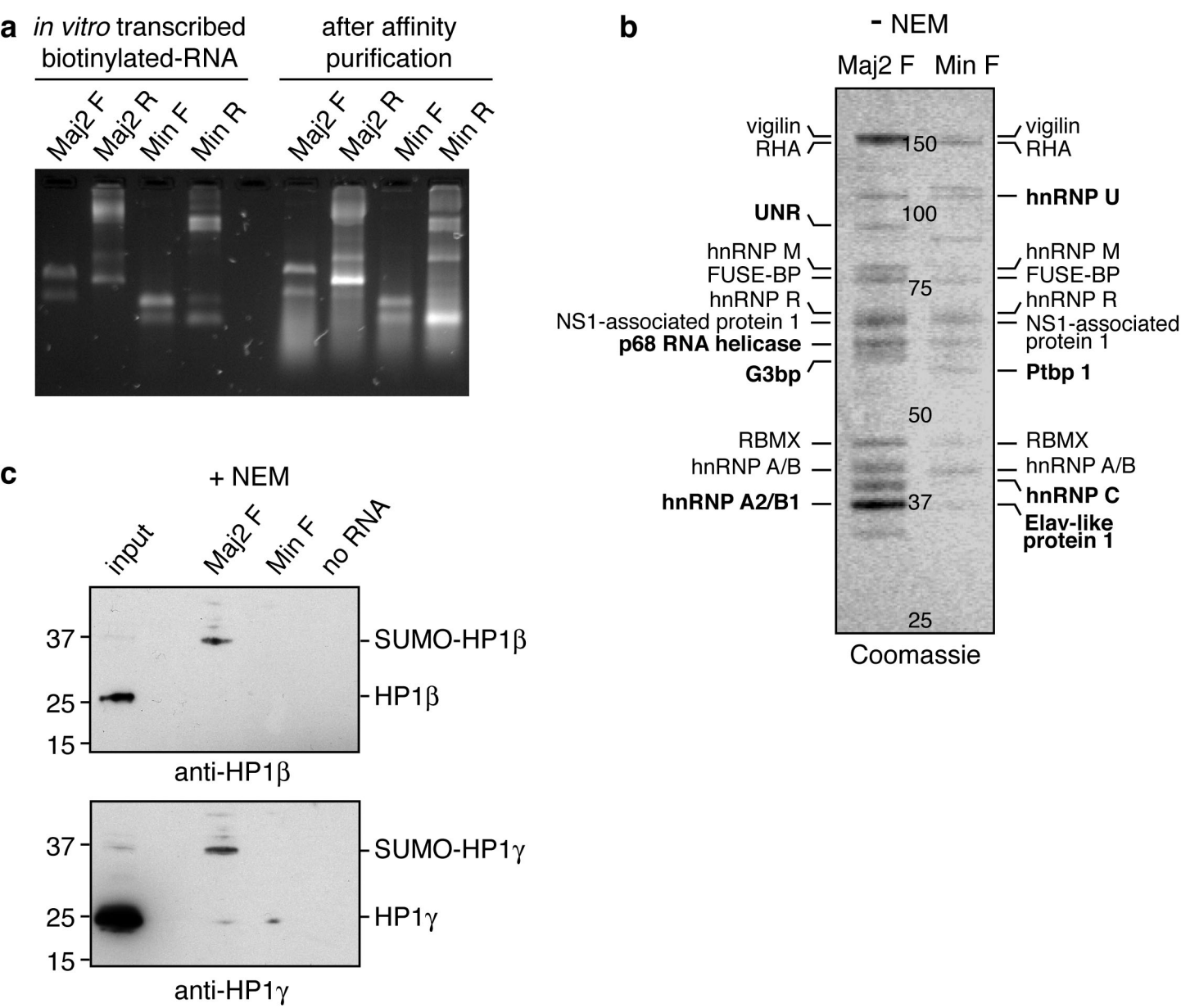
Christèle Maison, Delphine Bailly, Danièle Roche, Rocío Montes de Oca, Aline V. Probst, Isabelle Vassias, Florent Dingli, Bérengère Lombard, Damarys Loew, Jean-Pierre Quivy and Geneviève Almouzni



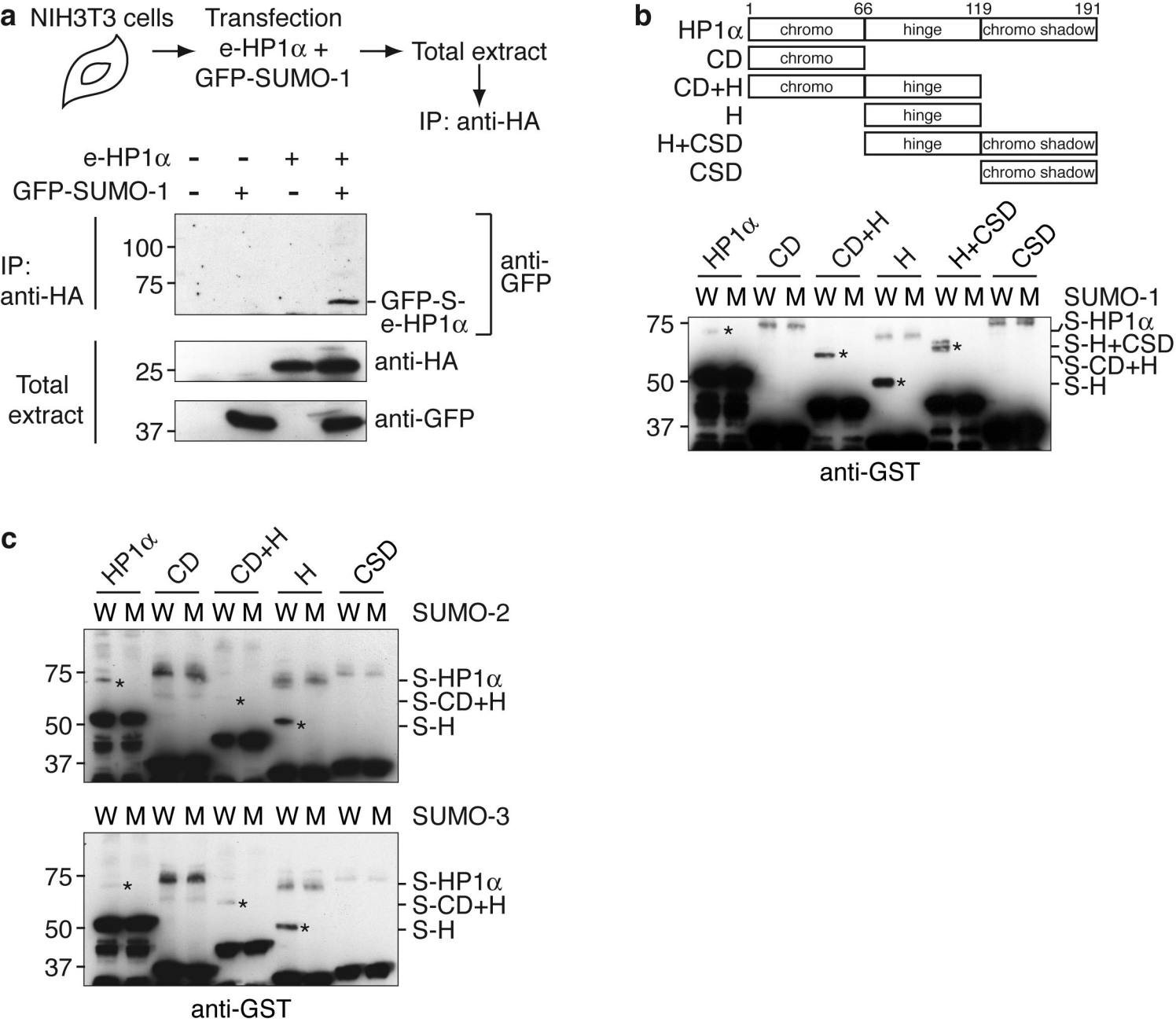
Supplementary Figure 1. Localization of major and minor transcripts. RNA FISH analysis of NIH3T3 cells mock (- RNase) or RNase (+ RNase) treated. Major (forward in green, reverse in red) and minor (forward in red, reverse in green) RNAs are detected with strand-specific LNA probes. Enlargement: magnification of the nucleus inside the white box. Scale bar, 10 μM.



Supplementary Figure 2. Localization of major transcripts. a. Quantification of forward major RNA localization at the periphery of pericentric domains. Top : RNA FISH. We show forward major RNA (green) along with a merged image of DAPI and RNA staining. Arrowheads indicate RNA foci away from pericentic domains. Middle: we calculated from 35 nuclei the percentage of forward major RNA foci (arrowheads) localized close to and away from pericentric domains. Bottom : graph comparing the presence of forward major RNA (foci/ µm3) at the periphery (black) or not at the periphery (grey) of pericentric domains from 35 nuclei which represent 408 pericentric domains and 196 RNA foci in 3 independent experiments. The error barr (SD) and the p_{value} are indicated. **b**. HP1α accumulation at major satellite DNA domains. Immuno-DNA FISH with anti-HP1α antibodies (green) and major or minor satellite DNA probes (red). Insets show magnification of a chromocenter. Scale bar, 10 μM.

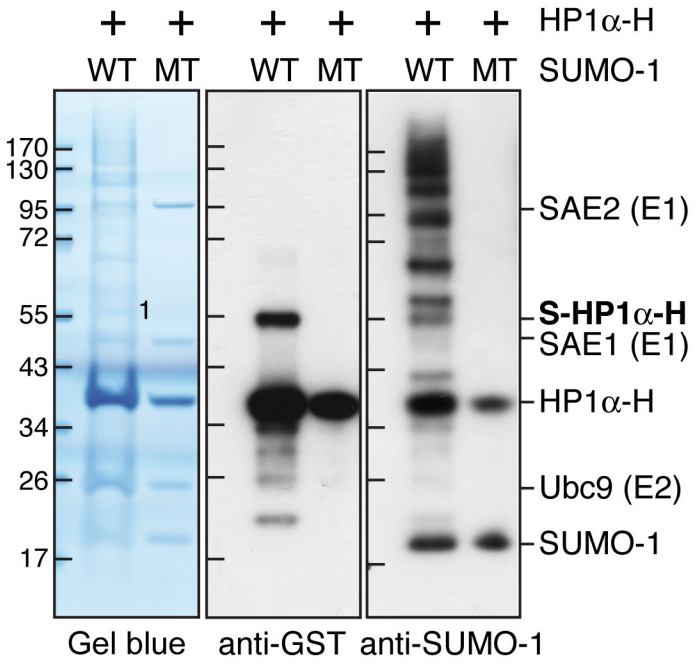


Supplementary Figure 3. Identification of proteins associated with centromeric RNAs. a. Ethidium bromide stained agarose gel shows *in vitro* transcribed biotinylated-RNAs before (10% of starting material) and after affinity purification using forward (F) and reverse (R) major (Maj2 F, Maj2 R) or minor (Min F, Min R) RNAs as baits. **b**. RNA pull down using forward major (Maj2 F) or minor (Min F) RNAs as baits, in the absence of NEM. A coomassie blue stained gel is shown and selected proteins identified by mass spectrometry are indicated. Bold proteins indicate proteins specifically associated with forward major or minor RNAs. **c**. RNA pull down using forward major (Maj2 F) or minor (Min F) RNAs, or no RNA as negative control, in the presence of NEM. Western blot analysis using anti-HP1β and anti-HP1γ specific antibodies revealed endogenous SUMO-HP1β, HP1β, SUMO-HP1γ and HP1γ. Input is 10% of nuclear extracts.



Supplementary Figure 4. **The hinge domain of HP1α is sumoylated by SUMO-1 or SUMO-2/3. a**. HP1 α sumoylation *in vivo*. Top: Experimental scheme. Bottom: We mock transfected or transiently transfected NIH3T3 cells with either e-HP1α, GFP-SUMO-1 or e-HP1α with GFP-SUMO-1, as indicated. 48 h after transfection, we lysed cells in the presence of NEM and immunoprecipitated total extracts using anti-HA beads. Western blotting using anti-GFP antibodies revealed sumoylated e-HP1α (GFP-S-e-HP1α). We verified the expression of e-HP1α and GFP-SUMO-1 in total cell extracts using anti-HA and anti-GFP antibodies. **b**. HP1α sumoylation *in vitro*. Top: Schematic representation of full-length HP1α and fragments thereof. Bottom: Western blot analysis of the SUMO-1 reaction mixtures by anti-GST antibodies revealed the positions of SUMO-1-modified HP1α full-length (S-HP1α) or fragment domains (S-H+CSD, S-CD+H and S-H), indicated by asterisks (*). **c**. HP1α sumoylation *in vitro*. Western blot analysis of the SUMO-2 (top) and SUMO-3 (bottom) reaction mixtures by anti-GST antibodies revealed the positions of sumoylated HP1α full-length (S-HP1α) or fragment domains (S-CD+H and S-H), indicated by asterisks (*).

a



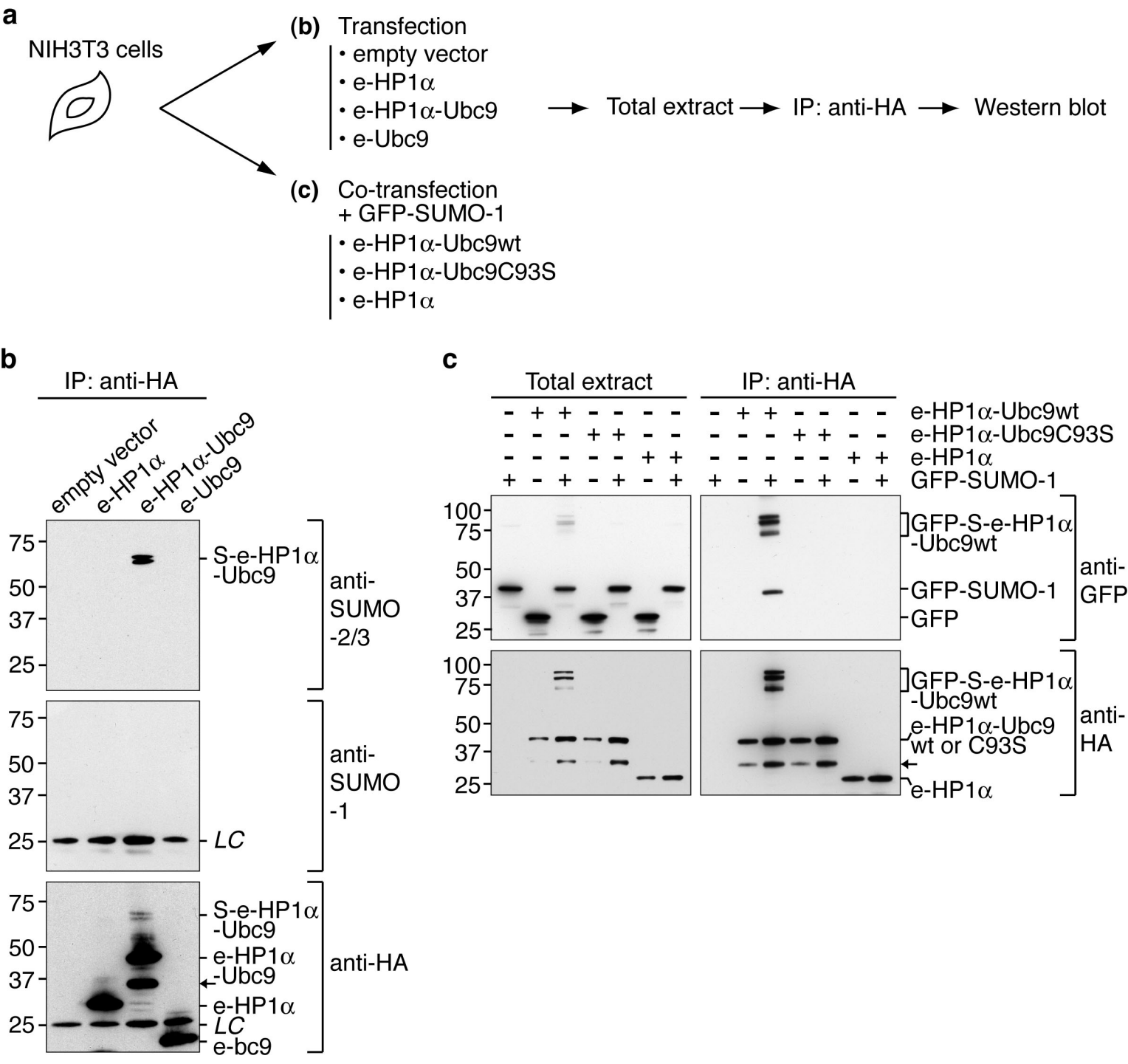
b

HP1 peptide sequence	**Qstar** precursor ion mass (m/z; 3+ ions)*	**Orbitrap** precursor ion mass (m/z; 3+ ions)*	Expected mass (M)
KK	809.7	809.71277 (+/- 4 ppm)	2426.124865
$KM_{ox}K$	**858.7**	**858.72485 (+/- 0.09 ppm)**	**2573.151309**
KYK or YKK	**864.1**	**nd**	**2589.179239**
KR	819.1	819.04779 (+/- 0.88 ppm)	2454.122058
KKR	861.8	nd	2582.217021
EKSEGNK	**981.8**	**981.77563 (+/- 2 ppm)**	**2942.297519**
SKK	**nd**	**838.72467 (+/-1.04 ppm)**	**2513.147939**

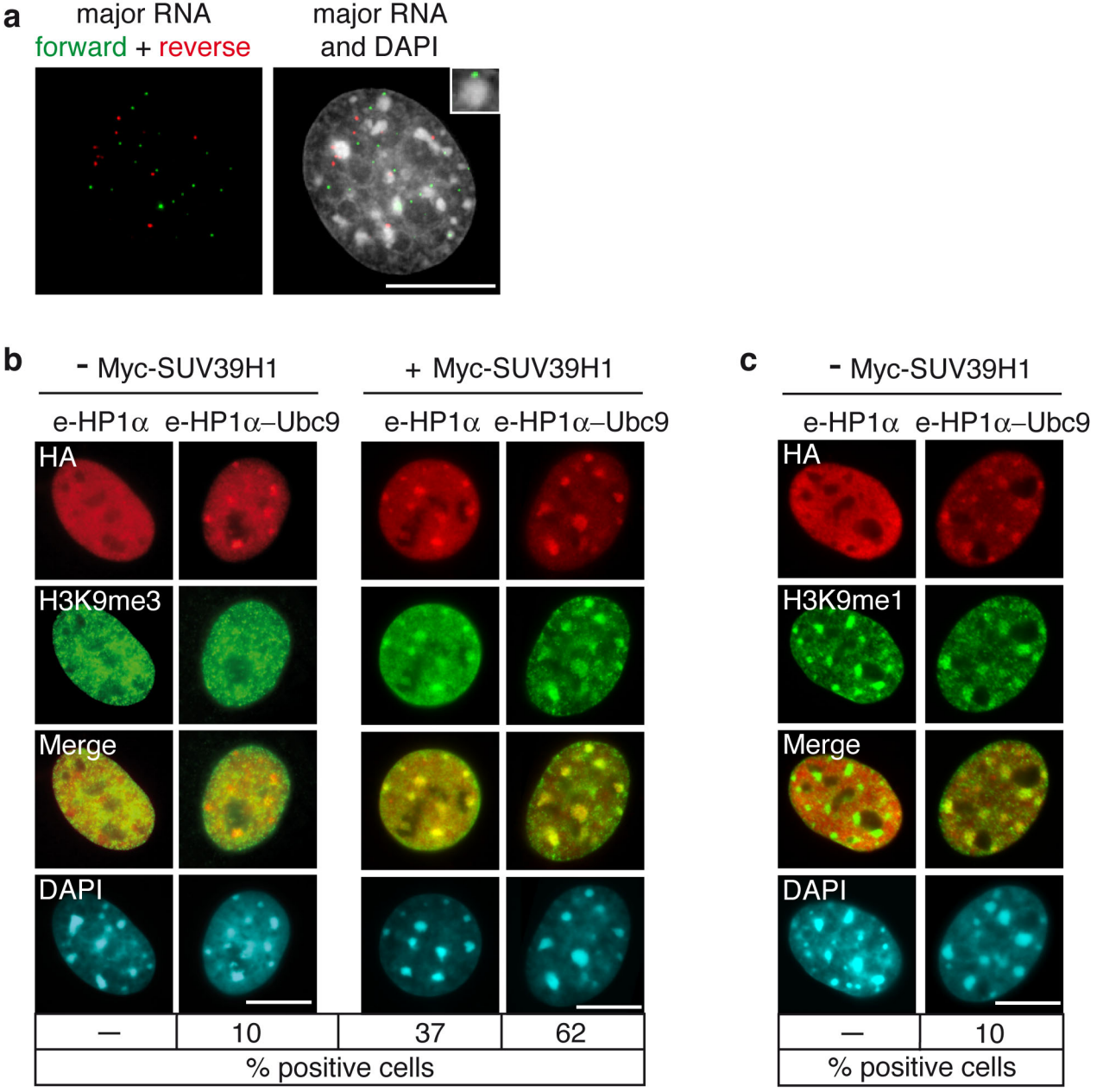
c

1 10 20 30 40 50 60
MGKKTKRTADSSSSEDEEEYVVEKVLDRRMVKGQVEYLLKWKGFSEEHNTWEPEKNLDCP
70 80 90 100 110 120
ELISEFMKKYKKMKEGENNKPREKSEGNKRKSSFSNSADDIKSKKKREQSNDIARGFERG
130 140 150 160 170 180
LEPEKIIGATDSCGDLMFLMKWKDTDEADLVLAKEANVKCPQIVIAFYEERLTWHAYPED
190
AENKEKESAKS

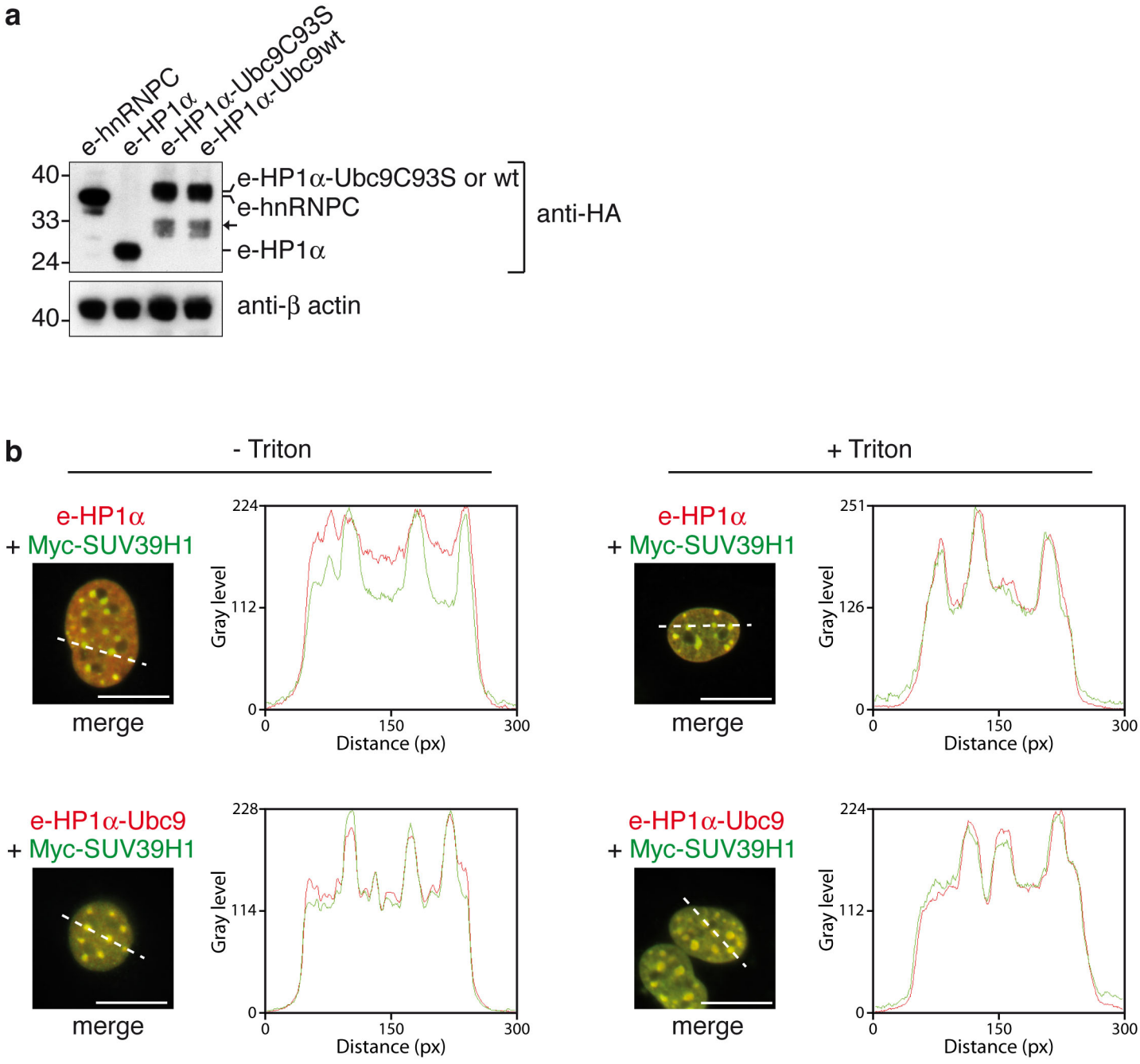
Supplementary Figure 5. Mass spectrometric analysis of *in vitro* sumoylated HP1α hinge domain. a. We divided sumoylation reactions performed as in Fig. 3a into two; we ran 90% of the mixture on a gel and Gel blue stained, then we analysed the remaining 10% by Western blot with rabbit anti-GST and rabbit anti-SUMO-1 antibodies (without stripping) to evidence SUMO-1-HP1α-H specific band (S-HP1α-H in bold). We excised and analyzed by mass spectrometry the corresponding band (labelled as 1 in the Gel blue stained). **b**. List of identified sumoylated peptides; indicated are the sequences from HP1α-H peptides linked to the SUMO-1 peptide (79-97 aa). We show the observed monoisotopic mass of the modified branched peptide (m/z [3+ ions]) in both the Qstar and Orbitrap instruments, as well as the expected mass (M). Accuracy between expected and observed masses is shown in brackets for the Orbitrap instrument. Indicated in bold are peptide sequences unique to HP1α-H. Asterisk (*), all spectra available upon request; “nd”, not determined. **c**. Amino acid sequence of HP1α. The hinge domain is denoted by a black box and the sites of sumoylation identified by mass spectrometry are indicated in red.



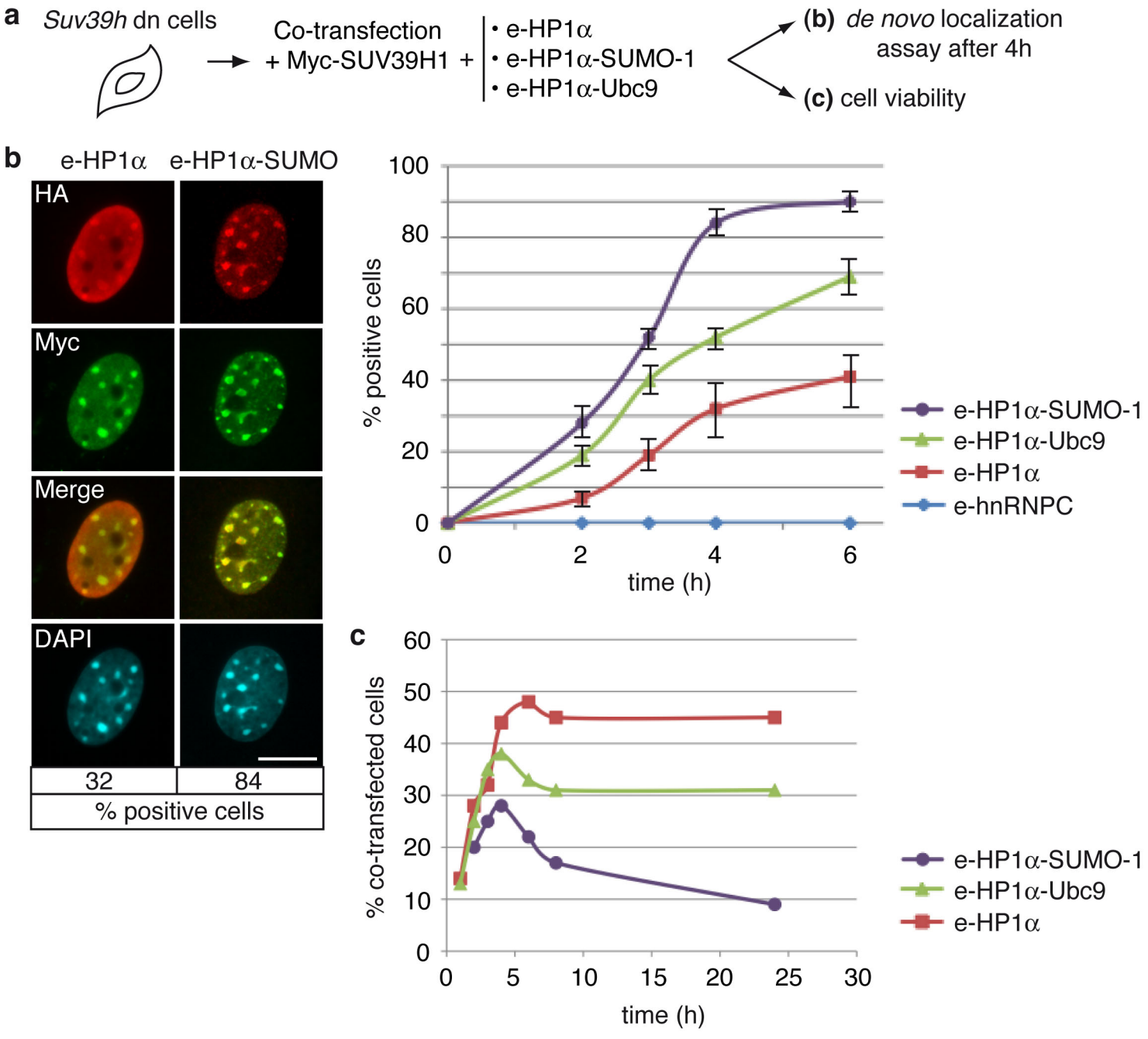
Supplementary Figure 6. **HP1α-Ubc9 is modified by SUMO-1 *in vivo*. a**. Experimental scheme. **b**. HP1α-Ubc9 sumoylation *in vivo*. 48 h after transfection of NIH3T3 cells with either an empty vector, e-HP1α, e-HP1α-Ubc9 or e-Ubc9, we lysed cells in the presence of NEM and immunoprecipitated total extracts using anti-HA beads. Western blotting using anti-SUMO-2/3, anti-SUMO-1 and anti-HA antibodies revealed the positions of SUMO-modified e-HP1α-Ubc9 (S-e-HP1α-Ubc9), unmodified e-HP1α-Ubc9, e-HP1α or e-Ubc9. LC corresponds to the immunoglobulin light chain. Arrow indicates a degradation product of e-HP1α-Ubc9. **c**. HP1α-Ubc9 sumoylation *in vivo*. We transfected NIH3T3 cells with e-HP1α-Ubc9wt, e-HP1α-Ubc9C93S, e-HP1α alone or with GFP-SUMO-1, as indicated. We verified the expression of the proteins in total cell extracts using anti-GFP and anti-HA antibodies (left). In immunoprecipitates, Western blotting using anti-GFP revealed sumoylated e-HP1α-Ubc9wt (GFP-S-e-HP1α-Ubc9wt) and using anti-HA antibodies revealed sumoylated e-HP1α-Ubc9wt (GFP-S-e-HP1α-Ubc9wt) and unmodified e-HP1α-Ubc9wt, e-HP1α-Ubc9C93S, e-HP1α (right). Arrow indicates a degradation product of e-HP1α-Ubc9.



Supplementary Figure 7. **HP1α-Ubc9 is targeted to pericentric heterochromatin in the absence of H3K9me3. a**. Localization of major RNA in Suv39h double-null cells by RNA FISH. Left: forward (green) and reverse (red) major RNAs detected with strand-specific LNA probes. Right: merged image of DAPI and RNA staining. The inset shows a magnification of a chromocenter. Scale bar, 10 μM. **b**. Localisation of e-HP1α and e-HP1α-Ubc9 in the absence (left) or in the presence (right) of Myc-SUV39H1 in Suv39h dn cells by immunofluorescence using anti-HA (red) and anti-H3K9me3 (green) antibodies 6h after transfection. For each condition, we examined 300 transfected cells and calculated the percentage of cells with HA signal enriched at pericentric domains (positive cells). Scale bar, 10 μM. **c**. Localization of e-HP1α and e-HP1α-Ubc9 in Suv39h dn cells by immunoflorescence using anti-HA (red) and anti-H3K9me1 (green) antibodies 6h after transfection. The percentage of positive cells was calculated as in b. The Suv39h dn cells showed H3K9me1 staining at pericentric domains. Scale bar, 10 μM.



Supplementary Figure 8. Sumoylation of HP1α promotes its accumulation at pericentric heterochromatin. a. Comparison of protein expression corresponding to the experiment showed in Fig. 4e by Western blot using anti-HA and anti-β actin antibodies. Arrow indicates a degradation product of e-HP1α-Ubc9. **b**. We co-transfected Suv39h dn cells with Myc-SUV39H1 and either e-HP1α or e-HP1α-Ubc9, and performed immunofluorescence with anti-HA (red) and anti-Myc (green) antibodies in untreated (left, - Triton) and Triton extracted (right, + Triton) cells. For each condition, a merge image of Myc and HA staining is shown with a line scan presenting the relative intensities of the Myc and HA signals across the nucleus. Representation of intensity (gray level) vs distance (pixels) is shown. Scale bar, 10 μM. Using this line scan analysis, we confirmed that while e-HP1α-Ubc9 and Myc-SUV39H1 overlapped in the absence of Triton treatment (left), e-HP1α showed a diffuse staining that spread throughout the nucleus. To determine if this diffuse staining was not due to ectopic localization at many sites within chromatin, we examined HP1α localization after Triton extraction (right). Under these conditions, the diffuse staining was no longer observed and the HP1α chromatin-bound dots showed a perfect overlap with SUV39H1 domains.



Supplementary Figure 9. **HP1α-SUMO-1 accumulates at pericentric domains. a**. Experimental scheme. We tested whether fusing directly SUMO-1 to HP1α could suffice to promote its efficient localization at pericentric domains. After co-transfection of an e-HP1α-SUMO-1 fusion construct together with Myc-SUV39H1 in Suv39h dn cells, we found that e-HP1α-SUMO-1 accumulated ~2.5-fold more efficiently to pericentric domains compared to e-HP1α (b). However, only ~9% of cells co-transfected with e-HP1α-SUMO-1 were viable 24h after transfection compared to 45% and 31% of e-HP1α and e-HP1α-Ubc9, respectively (c). Given the high mortality in this assay, we decided to carry out further experiments with HP1α-Ubc9 which proved less toxic. **b**. Left: *de novo* localization of e-HP1α and e-HP1α-SUMO-1 in Suv39h dn cells co-transfected with Myc-SUV39H1 by immunofluorescence using anti-HA (red) and anti-Myc (green) antibodies 4 h after transfection. For each condition, we examined 300 co-transfected cells and calculated the percentage of cells with HA signal enriched at pericentric domains (positive cells). Scale bar, 10 μM. Right: Time course analysis of the *de novo* localization of e-HP1α , e-HP1α-Ubc9, e-HP1α-SUMO-1 and e-hnRNPC in Suv39h dn cells co-transfected with Myc-SUV39H1. The percentage of positive cells as a function of the time after transfection is represented. Symbols indicate the mean and error bars indicate the standard deviation of three independent experiments (300 co-transfected cells counted in each condition). **c**. Cell viability after co-transfection of Suv39h dn cells with Myc-SUV39H1 and either e-HP1α, e-HP1α-Ubc9 or e-HP1α-SUMO-1. After immunofluorescence with anti-HA and anti-Myc antibodies, we examined 300 cells per condition and represented the number of co-transfected cells as a percentage of the number of total cells, and as a function of the time after transfection.

Major RNA interacting proteins						
Identifier	Protein name	Description	MW (Da)	Mascot score	Peptides	Coverage (%)
gil19527028	Vigilin	high density lipoprotein binding protein	141743	675.79	14	14.1
gil24429590	RHA	DEAH (Asp-Glu-Ala-His) box polypeptide 9	149583	293.56	6	4.7
gil21450287	**UNR**	DNA segment, Chr 3, MJeffers 1	88790.6	180.36	4	5.1
gil21313308	hnRNP M	heterogeneous nuclear ribonucleoprotein M	77648.7	254.84	5	8.0
gil33859724	hnRNP R	heterogeneous nuclear ribonucleoprotein R	70888.1	342.89	7	10.3
gil16975504	FUSE-BP	far upstream element (FUSE) binding protein 1	68539.6	181.07	4	6.6
gil13242328	NS1-associated protein 1	NS1-associated protein 1	69769.8	419.34	8	10.7
gil51263	**p68 RNA helicase**	p68 RNA helicase	69320.3	118.47	3	5.9
gil7305075	**G3bp**	ras-GTPase-activating protein SH3-domain binding protein	51828.8	128.79	3	9.9
gil6755296	RBMX	RNA binding motif protein, X-linked	42300.9	308.13	5	14.8
gil6754222	hnRNP A/B	heterogeneous nuclear ribonucleoprotein A/B	30831.3	268.67	5	14.7
gil7949053	**hnRNP A2/B1**	heterogeneous nuclear ribonucleoprotein A2/B1 isoform 1	35993	542.95	8	37.5

Minor RNA interacting proteins						
Identifier	Protein name	Description	MW (Da)	Mascot score	Peptides	Coverage (%)
gil19527028	Vigilin	high density lipoprotein binding protein	141743	89.67	2	1.2
gil24429590	RHA	DEAH (Asp-Glu-Ala-His) box polypeptide 9	149583	229.21	5	4.1
gil17390825	**hnRNP U**	Hnrpu protein	87917.7	136.22	2	4.1
gil21313308	hnRNP M	heterogeneous nuclear ribonucleoprotein M	77648.7	39.34	1	1.4
gil33859724	hnRNP R	heterogeneous nuclear ribonucleoprotein R	70888.1	132.22	3	4.4
gil16975504	FUSE-BP	far upstream element (FUSE) binding protein 1	68539.6	277.06	5	9.4
gil13242328	NS1-associated protein 1	NS1-associated protein 1	69769.8	132.22	3	4.5
gil13938631	**Ptbp 1**	Ptbp1 protein	59321.6	218.72	6	18.4
gil6755296	RBMX	RNA binding motif protein, X-linked	42300.9	116.85	2	6.9
gil6754222	hnRNP A/B	heterogeneous nuclear ribonucleoprotein A/B	30831.3	81.89	3	9.1
gil8393544	**hnRNP C**	heterogeneous nuclear ribonucleoprotein C	34384.8	173.06	4	12.8
gil31542602	**Elav-like protein 1**	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 1 (Hu antigen R)	36169	638.76	12	43.6

Supplementary Table 1. List of selected major and minor RNA associated proteins identified by mass spectrometry. Table shows identifier (NCBI protein accession number), description (NCBI protein definition), MW (molecular weight in Daltons), mascot score (sum of the unique mascot ions scores), peptides (number of peptides identified per experiment), coverage (% of sequence coverage identified from MS/MS data). Protein name in bold indicates proteins that associated with only major or minor RNAs. We performed the experiment twice and showed the best experiment.

Supplementary Table 2. Probe and primer sequences

LNA fluorescent probes for RNA and DNA FISH

Name	Fluorophore	Sequence
major 1	FITC	TCTTGCCATATTCCACGTCC
major 2	Cy3	GCGAGGAAAACTGAAAAAGG
minor 1	Cy3	GTTCTACAATGCCGGTTTCC
minor 2	FITC	TACACTGAAAAACACATTCG

Primers for RT-PCR

Name	Sequence (5'- 3')
major satellites (For)	AAATACACACTTTAGGACG
major satellites (Rev)	TCAAGTGGATGTTTCTCATT
minor satellites (For)	GAAAATGATAAAAACCACAC
minor satellites (Rev)	ACTCATTGATATACACTGTT