

Reviewed Preprint

v1 • July 29, 2026

Not revised

✉ For correspondence:

danica.fujimori@ucsf.edu

bo.huang@ucsf.edu

* These authors contributed to the manuscript equally

Competing interests: No competing interests declared

Funding: See [page 33](#)

Reviewing editor: Xiaobing Shi, Van Andel Institute, United States

© 2026, Berrocal et al. This article is distributed under the terms of the

[Creative Commons Attribution](#)

[License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Preferential formation of NUP98-KDM5A condensates at specific H3K4me3-rich loci drives leukemogenic gene expression

Augusto Berrocal^{1,*}, Jonathan E Sandoval^{2,*}, Neha Khetan^{3,*}, Alexa Ma⁴, Tianhong Wang¹, Camille Moore^{3,5}, Geeta J Narlikar³, Hao Li^{3,6}, Danica Galonić Fujimori^{1,2} ✉, Bo Huang^{1,3,6} ✉

¹Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, United States •

²Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, United States •

³Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, United States •

⁴Department of Molecular and Cellular Biology, University of California, Berkeley, Berkeley, United States •

⁵Tetrad Graduate Program, University of California, San Francisco, San Francisco, United States • ⁶Biohub, San Francisco, San Francisco, United States

eLife Assessment

This is an **important** study establishing the mechanistic principles how NUP98-KDM5A phase separates together with H3K4me3 chromatin in leukemia. Methods are sound and results are **convincing**, providing a framework to understand gene expression changes in leukemia patients.

<https://doi.org/10.7554/eLife.111922.1.sa3>

Abstract

Chromosomal translocations involving NUP98 generate fusion proteins that alter gene expression programs, yet the fundamental principles governing their gene targeting and condensate behavior remain poorly understood. Using NUP98-KDM5A as a model, we integrate cellular imaging, *in vitro* reconstitution, and genomic analyses to dissect how chromatin engagement shapes condensate formation. We find that NUP98-KDM5A forms sub-diffraction-limited, gel-like condensates whose assembly is potentiated by binding to H3K4me3. This interaction creates a quantitative targeting mechanism in which, at the native expression level, condensates preferentially form at genomic loci with high local H3K4me3 density. Such local density-dependent recruitment explains selective enrichment at the leukemogenic HOX gene clusters, despite widespread presence of H3K4me3 across the genome. Analysis of single-cell sequencing data from patients further supports a correlation between local H3K4me3 density and transcriptional activation in NUP98-KDM5A-driven leukemia. Together, our findings reveal how activating chromatin marks and condensate-forming proteins synergize to generate specificity within euchromatin, offering a generalizable framework for understanding how chromatin-associated condensates interpret epigenetic landscapes.

Introduction

Chromosomal translocations that generate NUP98-containing fusion proteins are a hallmark of multiple hematological disorders. These oncofusions are most prevalent in acute myeloid leukemia (AML) and are associated with poor prognosis (Bertrums et al., 2023 [↗](#); Gough et al., 2011 [↗](#); Tian et al., 2024 [↗](#); Umeda et al., 2025 [↗](#)). While all NUP98 oncofusions share an N-terminal intrinsically disordered region (IDR) derived from NUP98, the C-terminal portion is

derived from over 30 functionally diverse partner proteins (Michmerhuizen et al., 2020). The C-terminal fusion partners include those harboring a DNA- or chromatin-binding domain, allowing the oncofusions to localize to chromatin and alter gene expression. Several mechanisms of chromatin association have been described, including sequence-specific DNA-binding, such as NUP98-HOXA9 fusion, and specific histone mark recognition, exemplified by NUP98-KDM5A, NUP98-PHF23, and NUP98-KMT2A, all of which can bind trimethylated Lys4 in histone H3 (H3K4me3) through a PHD domain (Ahn et al., 2025; Calvo et al., 2002; Chandra et al., 2022; Gough et al., 2014; Wang et al., 2007, 2009).

NUP98 fusion oncoproteins have been shown to form condensates (Ahn et al., 2025; Chandra et al., 2022; Jevtic et al., 2022; Oka et al., 2023; Terlecki-Zaniewicz et al., 2021; Tripathi et al., 2023). These condensates may enrich general transcriptional machinery and transcriptional coactivators, including histone acetyltransferases, bromodomains, long noncoding RNAs, and the components of the menin-MLL complex, which promote sustained expression of leukemogenic genes (Ahn et al., 2025; Hamamoto et al., 2025; Heikamp et al., 2022, 2024; Michmerhuizen et al., 2025; Terlecki-Zaniewicz et al., 2021). Critically, the NUP98-derived region, which contains an FG-repeat sequence that mediates phase separation and condensate formation, is essential for leukemogenic transformation, as truncations and mutations of the phenylalanines in the FG repeats abrogate condensate formation and arrest the proliferation of hematopoietic stem and progenitor cells (Ahn et al., 2021, 2025; Chandra et al., 2022; Jevtic et al., 2022; Terlecki-Zaniewicz et al., 2021). In addition, chromatin binding is central to the leukemogenic potential of NUP98 fusion proteins, as mutational disruption of DNA recognition in NUP98-HOXA9 fusion and H3K4me3 binding by the PHD3 chromatin-binding domain in NUP98-KDM5A abolishes leukemogenic transformation (Ahn et al., 2021; Chandra et al., 2022; Wang et al., 2009). These findings indicate that both condensate formation and chromatin association of NUP98 oncofusions are required for leukemogenic transformation.

In normal cellular functions, several chromatin-associated condensates have been described, and they are associated with distinct chromatin states. For instance, the association between heterochromatin marks, like H3K9me3 and H3K27me3, and repressive factors, including HP1 and PRC1, leads to the formation of heterochromatin condensates, which compact chromatin and enrich silencing machinery to enforce a repressive chromatin state (Ackermann & Debelouchina, 2019; Deshpande et al., 2024; Larson et al., 2017; Plys et al., 2019; Sanulli et al., 2019; Strom et al., 2017; Tatavosian et al., 2019; Wang et al., 2019; Zhen et al., 2014). On the other hand, condensates associated with active, decompacted euchromatin are less well understood, with the exception of transcriptional condensates at super-enhancers that involve transcriptional coactivators and condensates at active gene arrays such as histone locus bodies (Sabari et al., 2018; Strom et al., 2024; Zhang et al., 2021). Oncofusions associated with the active-promoter histone mark H3K4me3, such as NUP98-KDM5A, offer disease model systems to probe the function of euchromatin-associated condensates.

Because HOXA9 is a leukemogenic transcription factor in itself, oncofusions with its DNA-binding domain enable targeting the NUP98-HOXA9 fusion to critical proleukemogenic genes. By contrast, H3K4me3 association could bring NUP98 fusions with chromatin readers to all active promoters across the genome. How these oncofusions bias the activation of leukemogenic genes remains an open question. Here, we address this key question using NUP98-KDM5A as our model system. Using a combination of cellular microscopy and *in vitro* reconstitution, we showed that NUP98-KDM5A forms gel-like condensates, and condensate formation is enhanced by H3K4me3 chromatin association. This association leads to a concentration-dependent selectivity of condensate targeting, with condensates preferentially forming at genomic loci with high local H3K4me3 mark density at pathologically relevant expression levels, such as the leukemogenic HOX gene clusters. We confirmed the correlation between gene activation and local H3K4me3 density by analyzing patient sequencing data. Our results not only provide mechanistic insights into the leukemogenesis process of NUP98 oncofusions involving chromatin readers, but also illustrate how functional specificity can be attained through the quantitative interplay between epigenetic landscapes and condensate-forming chromatin factors.

Results

NUP98-KDM5A forms condensates at H3K4me3-marked chromatin in a concentration-dependent manner

We sought to establish an expression system and to dissect the molecular determinants governing the cellular behavior of the oncogenic NUP98-KDM5A fusion protein. With this aim, we generated a series of Mammalian Toolkit-based plasmids (Fonseca et al., 2019) expressing fluorescently tagged NUP98-KDM5A and its variants and introduced them into cells via transient transfection, yielding cell populations with a wide range of expression levels in a single experiment (Chong et al., 2018, 2022). Live-cell imaging of U2OS cells revealed that expression of mEGFP-tagged NUP98-KDM5A leads to the formation of dense nuclear puncta, in contrast to mEGFP-tagged full-length KDM5A at similar expression levels (Figure 1 – Figure Supplement 1). However, the limited resolution of spinning-disk confocal microscopy hindered accurate segmentation of the puncta for quantitative analysis. Therefore, in our subsequent cell imaging experiments, we utilized the NSPARC scanning confocal microscope or the SoRa spinning disk confocal microscope, which have enhanced spatial resolutions ($\sim 1.4x$) to resolve neighboring NUP98-KDM5A puncta. For NUP98-KDM5A with mEGFP tagged on the C-terminus (NUP98-KDM5A-mEGFP), nuclear puncta almost all have sub-resolution-limit sizes (< 150 nm) across the expression level range (Figure 1A). Quantitative analysis of NSPARC images, followed by calibration of EGFP concentration to the microscope-acquired fluorescence signal (Figure 1 – Figure Supplement 2C and D), shows a threshold concentration of ~ 50 nM for puncta detection (Figure 1C). As the overall nuclear concentration of NUP98-KDM5A ($C_{\text{NUP98-KDM5A}}$) increases, the density of puncta rises, plateauing at $\sim 0.12 \mu\text{m}^{-2}$. When the concentration is above $\sim 1 \mu\text{M}$, approximately 40% of the total fusion protein resides in these small puncta (Figure 1C). We estimated patient cellular concentration of NUP98-KDM5A to be ~ 180 nM, which falls within the transition range (Figure 1 – Figure Supplement 3A,B; see Estimation of NUP98-KDM5A expression in patient cells in Materials and Methods). The small, dense nuclear puncta formed by NUP98-KDM5A are shaped by binding to H3K4me3 sites, which is supported by the observation that the chromatin-binding-deficient mutant (NUP98-KDM5A_{mut}-mEGFP) forms large, spherical droplets that often coincide with nucleoli (Figure 1 – Figure Supplement 2A and B).

To further test whether puncta formation involves phase separation in addition to chromatin binding, we replaced mEGFP with mCherry. While similar in size to mEGFP, mCherry has higher solubility and has previously been shown to promote the dissolution of biomolecular condensates (Brumbaugh-Reed et al., 2024; Mestrom et al., 2019). Indeed, compared to NUP98-KDM5A-mEGFP, NUP98-KDM5A-mCherry shows nuclear puncta formation at a higher concentration (~ 200 nM), and a lower fraction of fusion protein molecules was localized to these puncta. NUP98-KDM5A-mCherry nuclear puncta have a similarly sub-resolution-limited size and plateau at a similar density (Figure 1B, C). The observation that NUP98-KDM5A puncta are formed by phase separation is further supported by the fact that the KDM5A-derived fragment of the oncofusion alone, which mediates chromatin-binding, does not form nuclear puncta when overexpressed in U2OS cells (Figure 1 – Figure Supplement 1A). Thus, we conclude that nuclear NUP98-KDM5A puncta are condensates formed through a combination of chromatin binding and phase separation.

The plateauing of NUP98-KDM5A condensate density at high concentration suggests saturation of H3K4me3 binding sites. To further characterize the occupancy of H3K4me3 sites by NUP9-KDM5A, we performed immunofluorescence of H3K4me3 and analyzed its image correlation with NUP98-KDM5A-mEGFP (Figure 1D). Pixel intensity histograms clearly show that, at low expression levels, NUP98-KDM5A-mEGFP condensates occupy a subset of H3K4me3 sites, whereas in cells with high expression, H3K4me3 sites fully colocalize with NUP98-KDM5A-mEGFP condensates. In contrast, the chromatin-binding-deficient NUP98-KDM5A_{mut}-mEGFP shows complete anti-colocalization with H3K4me3 (Figure 1E). These trends are confirmed by the image Pearson correlation values as a function of NUP98 fusion expression inferred from mEGFP fluorescence

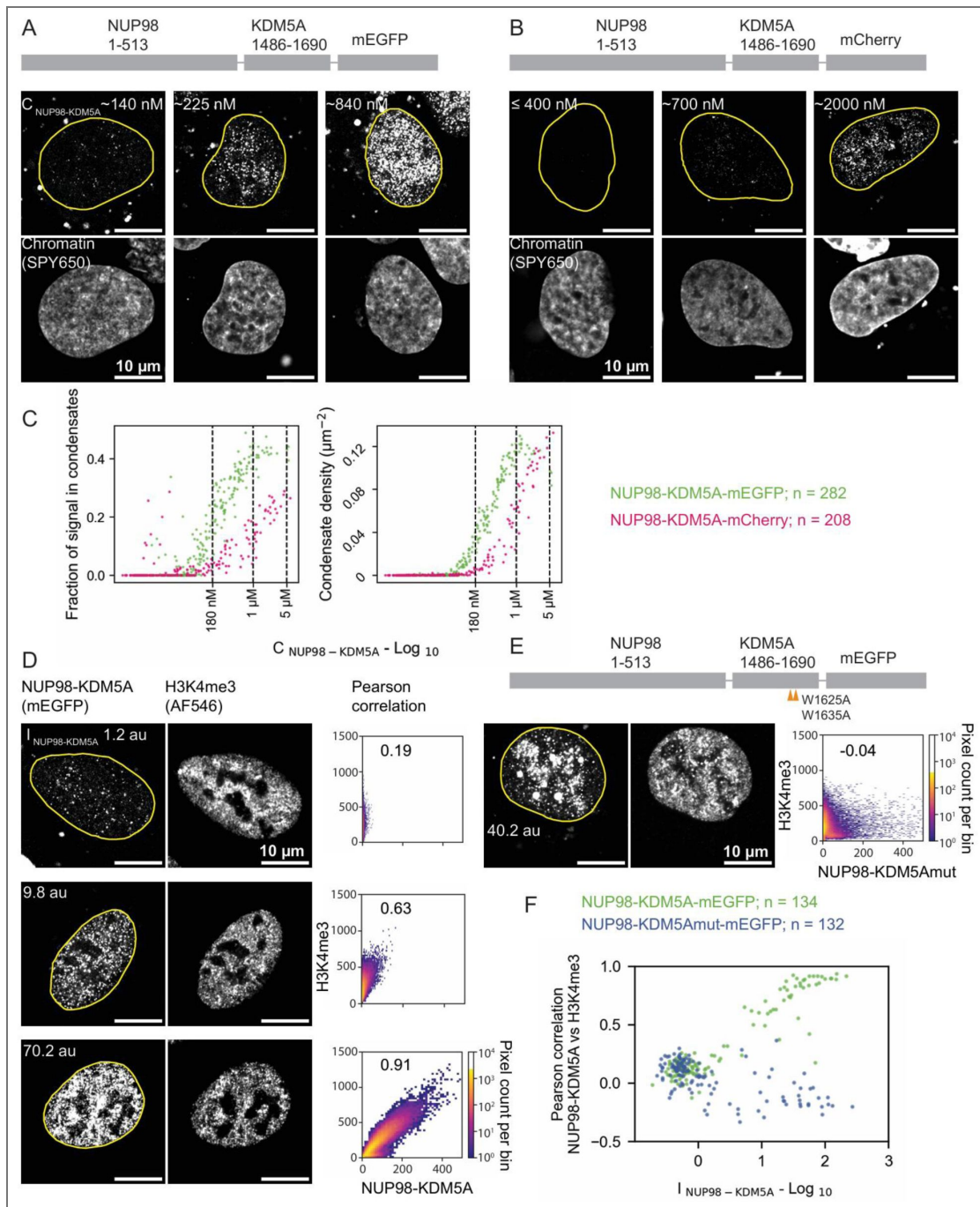


Figure 1. Concentration-dependent formation of H3K4me3-associated condensates by NUP98-KDM5A.

A. Schematic of NUP98-KDM5A construct, with the N-terminal region of NUP98 (amino acids 1-513) fused to the C-terminal region of KDM5A (1486-1690), which includes the H3K4me3-binding PHD3 domain, and NSPARC images of NUP98-KDM5A-mEGFP in U2OS nuclei at different concentrations showing sub-resolution-limit puncta. **B.** Schematic of the NUP98-KDM5A-mCherry construct and NSPARC images in U2OS nuclei at different concentrations. **C.** Fraction of signal in condensates (fluorescent protein signal within condensates relative to total nuclear signal) and condensate density (number of condensates per μm^2) as a function of nuclear protein concentration in individual cells (n = 282 cells for mEGFP tagging and n = 208 for mCherry tagging). **D.** Fixed U2OS nuclei expressing NUP98-KDM5A-mEGFP and immunostained for H3K4me3 with Alexa Fluor 546 (AF546), showing images with different mean nuclear mEGFP intensities ($I_{\text{NUP98-KDM5A}}$), 2D histograms of pixel intensities, and the Pearson correlation, R, between the images of the two channels. **E.** Schematic of the chromatin-binding-deficient NUP98-KDM5A (W1625A, W1635A) construct, and the same set of analyses as in D. **F.** Image Pearson correlation, R, as a function of expression level, $I_{\text{NUP98-KDM5A}}$, for U2OS cells expressing NUP98-KDM5A-mEGFP (n = 134 cells) or NUP98-KDM5Amut-mEGFP (n = 132).

intensity ($I_{\text{NUP98-KDM5A}}$). The correlation between NUP98-KDM5A-mEGFP and H3K4me3 rises from zero to almost unity when NUP98-KDM5A-mEGFP expression exceeds a threshold, whereas that between NUP98-KDM5A-mut-mEGFP and H3K4me3 remains low across the expression level range and has a negative trend (Figure 1F).

Taken together, our results indicate that NUP98-KDM5A forms nuclear, nanometer-sized condensates seeded at H3K4me3-marked sites through phase separation, and that both the fusion oncoprotein and its chromatin association determine the saturation concentration and biophysical properties of these condensates. As NUP98-KDM5A concentration increases, progressively more H3K4me3 sites become occupied until no additional sites remain available to support condensate formation.

H3K4me3 binding enhances the formation of NUP98-KDM5A condensates in vitro

To quantitatively evaluate the interplay between NUP98-KDM5A and H3K4me3 chromatin in condensate formation, we used a reconstitution system that contains only purified components. To establish a concentration threshold for the formation of condensates *in vitro*, purified NUP98-KDM5A was labeled with the fluorescent dye AF488, and concentration-dependent condensate formation was monitored by confocal microscopy (Figure 2 – Figure Supplement 1A). Similar to other NUP98 fusions (Chandra et al., 2022), we found that at concentrations between 5.8 and 11.7 nM, full-length NUP98-KDM5A forms condensates with irregular, submicron-sized morphological features (Figure 2 – Figure Supplement 1A).

The N-terminal FG-rich IDR of NUP98 has been previously shown to be essential for condensate formation (Ahn et al., 2021, 2025; Chandra et al., 2022; Jevtic et al., 2022; Terlecki-Zaniewicz et al., 2021). To evaluate the contributions of the sequence elements of the KDM5A fusion partner on the apparent saturation concentration (C_{sat}), we generated three additional constructs: (1) NUP98-KDM5A(1601-1662), which retains only the PHD3 domain of the KDM5A fusion partner; (2) NUP98-KDM5A(1486-1600), where the KDM5A fusion partner is truncated at its C-terminus to remove the PHD3 domain, and (3) the NUP98 segment alone (Figure 2 – Figure Supplement 1A). While the PHD3-domain containing construct NUP98-KDM5A (1601-1662) exhibited a C_{sat} range similar to the full-length NUP98-KDM5A fusion (C_{sat} : 5.8-11.7 nM), both the truncated fusion construct lacking the PHD3 domain and the NUP98 alone formed condensates at higher protein concentrations (C_{sat} : 23.4-46.9 nM) (Figure 2 – Figure Supplement 1A). These findings suggest that, in addition to the previously characterized interactions between the FG domains of NUP98, the PHD3 domain of KDM5A significantly impacts the condensate behavior of NUP98-KDM5A.

We next evaluated the ability of the fusion protein to recognize the H3K4me3 modification by assessing binding to a fluorescently labelled H3K4me3 10mer peptide using a fluorescence polarization (FP)-based assay (Figure 2A). We also introduced an H3K4me3-binding-deficient PHD3 mutant into the fusion protein to modulate its ability to recognize the modification (Figure 2A). We found that the WT NUP98-KDM5A construct bound the H3K4me3 peptide with a dissociation constant (K_D) of 1.28 μM (95% CI, 0.43-3.74 μM), which is consistent with K_D values reported for the isolated PHD3 domain of KDM5A (Figure 2A) (Zhang et al., 2022). In contrast, the chromatin-binding-deficient mutant NUP98-KDM5A W1625A exhibited no measurable binding with increasing protein concentrations of up to 20 μM (Figure 2A). To assess the specificity of histone tail-binding, we performed a competition-based FP assay in which increasing concentrations of unlabeled H3K4me3 peptide were used to displace a fluorescently labeled H3K4me3 peptide (Figure 2B). The unlabeled H3K4me3 peptide effectively displaced the fluorescently labeled ligand from H3K4me3-bound NUP98-KDM5A and MBP-PHD3 with an IC_{50} of 0.45 μM (95% Confidence Interval (C.I.), 0.32-0.62 μM) or 1.20 μM (95% C.I., 0.95-1.55 μM), respectively (Figure 2B). Together, these results indicate that the PHD3 domain in the NUP98-KDM5A fusion protein retains the ability to recognize the H3K4me3 mark in the context of the histone H3 peptide.

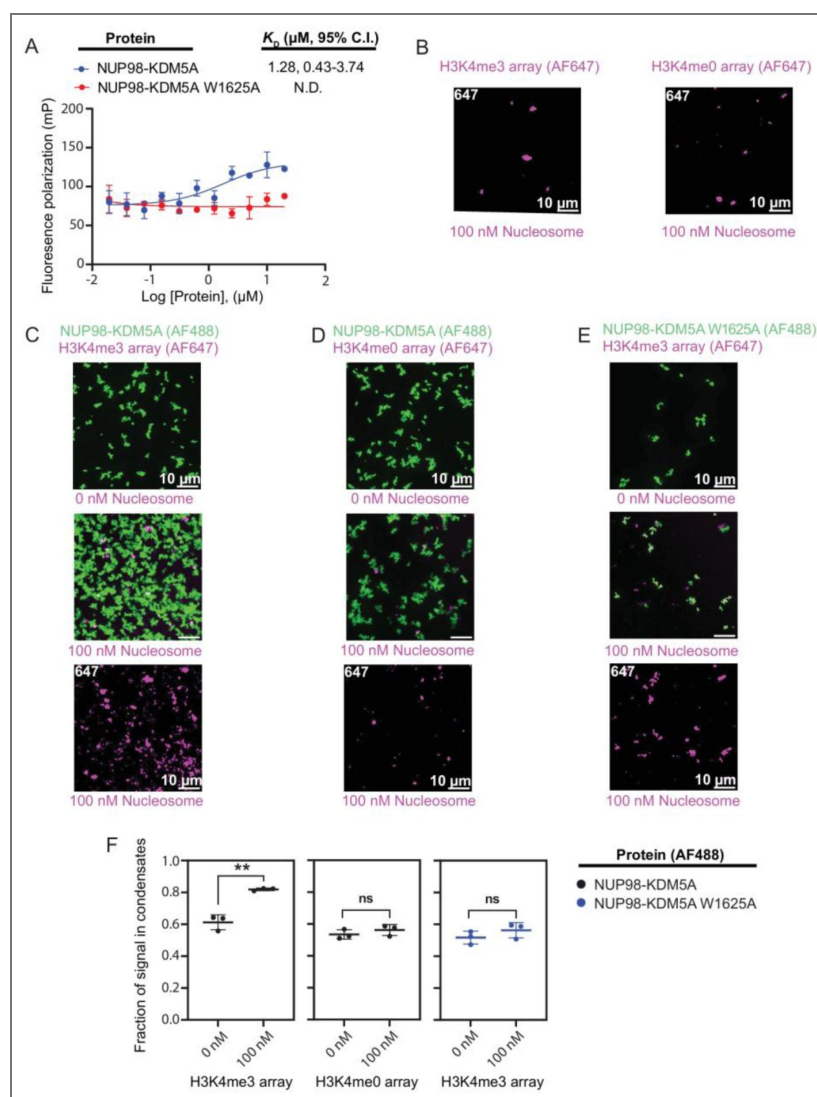


Figure 2. Binding of the PHD3 domain of NUP98-KDM5A to H3K4me3 promotes the formation of NUP98-KDM5A condensates.

A. Fluorescence polarization (FP) assays to determine the K_D of NUP98-KDM5A WT or W1625A mutant for a 10-mer FAM-labelled H3K4me3 peptide (10 nM). Error bars represent standard deviations from technical triplicates. K_D values show means and 95% C.I.. N.D.: Not determined because no binding was observed. **B.** Competitive displacement of 2 μ M NUP98-KDM5A (WT) or MBP-PHD3 (KDM5A residues 1601-1662) bound to 10-mer FAM-H3K4me3 (10 nM) by increasing concentrations of unlabeled 10-mer H3K4me3 peptide. Error bars represent standard deviations from technical triplicates. IC_{50} values show means and 95% C.I.. **C-F.** Representative confocal images of **C** 100 nM AF647 labelled H3K4me3 or H3K4me0 nucleosomal arrays, **D** AF488 labelled NUP98-KDM5A (800 nM) with AF647 H3K4me3 nucleosomal arrays (0 or 100 nM), **E** AF488 NUP98-KDM5A (800 nM) with AF647 H3K4me0 nucleosomal arrays (0 or 100 nM), or **F** AF488 NUP98-KDM5A W1625A (800 nM) with AF647 H3K4me3 nucleosomal arrays (0 or 100 nM). All images are maximum intensity projections (scale bar = 10 μ m). **G.** The fraction of protein fluorescence signal in condensates relative to total signal in the image for the conditions in **D-F**. $n = 3$ technical replicates; ns: $P > 0.05$; *: $P \leq 0.05$; **: $P \leq 0.01$, parametric unpaired t-test.

We then sought to examine the effect of NUP98-KDM5A-nucleosome interactions on the behavior of NUP98-KDM5A condensates. In these experiments, we carried out *in vitro* phase separation assays using either chicken erythrocyte polynucleosomes as a substrate with a heterogeneous pattern of histone modifications, or reconstituted nucleosomal arrays with a uniform pattern of histone modifications (H3K4me3 or -me0). In the presence of chicken erythrocyte polynucleosomes, there is a roughly 15% increase in the fraction of NUP98-KDM5A incorporated into condensates relative to reactions lacking chicken erythrocyte polynucleosomes (Figure 2 - Figure Supplement 1B and D [↗](#)). In contrast, the addition of chicken erythrocyte polynucleosomes to NUP98-KDM5A W1625A had no significant effect on the fraction of molecules partitioned into condensates (Figure 2 - Figure Supplement 1C and D [↗](#)). These observations imply chromatin binding enhances NUP98-KDM5A condensate formation.

It is well characterized that FG-repeat IDRs from NUP98 and other nucleoporins can form gel-like condensates, which are important for controlling nuclear transport by the nuclear pore (Schmidt & Görlich, 2015 [↗](#)). The irregular morphology of reconstituted NUP98-KDM5A condensates also indicates their gel-like states. Indeed, fluorescence recovery after photobleaching (FRAP) measurements of NUP98-KDM5A condensates show little recovery both in the absence of polynucleosomes (mobile fraction, $M_F = 0.043$) and with the addition of chicken erythrocyte polynucleosomes ($M_F = 0.072$) (Figure 2 - Figure Supplement 2A [↗](#)). Consistent with reconstituted condensates, live cell FRAP measurements of NUP98-KDM5A-mEGFP also show a low mobile fraction ($M_F = 0.085$) (Figure 2 - Figure Supplement 2A [↗](#)). Interestingly, switching the mEGFP tag to N-terminus dramatically increases the fluidity of NUP98-KDM5A condensates in living cells ($M_F = 0.48$ and half-time of recovery $T_{1/2} = 1.9$ sec) (Figure 2 - Figure Supplement 2A [↗](#)), despite a similarly small, punctate appearance of the condensates, comparable fraction of signal in condensates and condensate density regardless of the position of the mEGFP tag (Figure 2 - Figure Supplement 2B-D [↗](#)). A possible explanation for this effect is the insolubility of the mEGFP tag in fluorescence protein-repeat condensates (Frey et al., 2018 [↗](#)), which disrupts gel networks when these two components are immediately adjacent in the N-terminal tagged construct. This observation may explain the variability of fluidity reported for NUP98 oncofusion condensates (Chandra et al., 2022 [↗](#); Ibáñez De Opakua et al., 2022 [↗](#); Schmidt & Görlich, 2015 [↗](#)). It also provides another vivid example of how tagging could affect the biophysical properties of biomolecular condensates (Fatti et al., 2025 [↗](#); Zhou et al., 2025 [↗](#)). Based on these findings, we used C-terminal mEGFP tagging for all our live-cell experiments.

To further interrogate the role of chromatin ligands in the formation and the morphology of NUP98-KDM5A condensates, we evaluated the influence of substrates with a homogeneous pattern of histone modifications. For this, we prepared nucleosomal arrays with histone octamers containing either unmodified (H3K4me0) or trimethylated (H3K4me3) H3 reconstituted onto AF647-labeled DNA consisting of a previously characterized 3.5 kb DNA sequence from the 5' end of mouse gene *Cyp3a11*, which contains approximately 16 positioned nucleosomes. (Luger et al., 1999 [↗](#); Moore et al., 2025 [↗](#)). Consistent with previous studies, we found that nucleosomal arrays, in the absence of NUP98-KDM5A or crowding agents, can form condensates under physiological concentrations of monovalent salts (Moore et al. 2025 [↗](#), Gibson et al. 2019) (Figure 2C [↗](#)) (Figure 2 - Figure Supplement 3A [↗](#) and C [↗](#)). The reconstituted H3K4me3 or H3K4me0 nucleosome arrays were then mixed with AF488-labeled NUP98-KDM5A at a physiological salt concentration, and condensates were monitored by confocal microscopy (Figure 2D and E [↗](#), Figure 2 - Figure Supplement 3A [↗](#) and B [↗](#)). Our results show that *in vitro* mixtures of NUP98-KDM5A with H3K4me3 or H3K4me0 nucleosomes showed a similar degree of colocalization, with the highest degree of colocalization at 12.5 nM nucleosomes (Figure 2D, E [↗](#)) (Figure 2 - Figure Supplement 3A, B [↗](#)). We also observed that mixtures of NUP98-KDM5A W1625A with H3K4me3 nucleosomes (Figure 2F [↗](#)) (Figure 2 - Figure Supplement 3C [↗](#)) exhibited a similar degree of colocalization as mixtures of NUP98-KDM5A with H3K4me3 (Figure 2D [↗](#)) (Figure 2 - Figure Supplement 3A [↗](#)) or H3K4me0 nucleosomes (Figure 2E [↗](#)) (Figure 2 - Figure Supplement 3B [↗](#)). Interestingly, we found that the presence of H3K4me3 nucleosomes led to approximately a 20% increase in the fraction of NUP98-KDM5A signal in condensates, which appear irregularly shaped, highly interconnected, and gel-like (Figure 2D, G [↗](#)). In contrast, no significant changes were observed in the fraction of

signal incorporated into condensates when the WT NUP98-KDM5A was incubated with the H3K4me0 nucleosome array (Figure 2E, G [G](#)), or when the NUP98-KDM5A W1625A mutant protein was incubated with the H3K4me3 nucleosome array (Figure 2F, G [G](#)). Together, our findings reveal that the interaction between the PHD3 domain in the NUP98-KDM5A fusion protein and the H3K4me3 mark is a critical modulator of the spatial organization and behavior of NUP98-KDM5A condensates.

NUP98-KDM5A preferentially forms condensates at H3K4me3-enriched loci

A survey of publicly available H3K4me3 ChIP-seq data revealed that HOX clusters, which are key regulators of developmental processes and contain prominent leukemogenic genes such as HOXA9, are among the most densely H3K4-trimethylated regions across many cell types. We thus tested whether high local H3K4me3 density leads to preferential formation of NUP98-KDM5A condensates. We performed fluorescence *in situ* hybridization (FISH) combined with immunofluorescence in HEK293 cells to quantify the colocalization between condensates formed by C-terminally FLAG-tagged NUP98-KDM5A (NUP98-KDM5A-FLAG) and various genomic loci, including the HOXA/B/D clusters with high H3K4me3 density, inactive TNF but with H3K4me3 from nearby genes, as well as the inactive gene MUC4 and the active MYC, both having very low total H3K4me3 within a ± 50 kb window (Figure 3A [A](#)). For each locus, we used two differently colored sets of FISH probes targeting contiguous sequences to identify specific signals as paired spots in NSPARC enhanced-resolution 3D images (centroid distance < 300 nm). We then quantified NUP98-KDM5A-FLAG condensate association by measuring the local NUP98-KDM5A signal, defined as the maximum local intensity (within 200 nm 3D radius from the loci centroid) minus the neighboring background (the 20th-percentile intensity within 800 nm 3D radius) (Figure 3B [B](#)). Figure 3C [C](#) shows this local NUP98-KDM5A signal for each identified locus in the FISH image as a function of the NUP98-KDM5A expression level, which is represented by the mean nuclear fluorescence signal ($I_{\text{NUP98-KDM5A}}$).

All six scatter plots in Figure 3C [C](#) show a general trend that a higher expression level leads to more loci exhibiting a high local NUP98-KDM5A signal, i.e., association with a NUP98-KDM5A condensate, which is consistent with the trend shown in Figure 1D [D](#). At a low expression level, on the other hand, a substantial fraction of HOXA/B/D loci still have high local NUP98-KDM5A signal, whereas the vast majority of TNF, MUC4, and MYC loci show low local NUP98-KDM5A signal. This difference is more clearly shown in the box plots (Figure 3D [D](#)) by dividing cells into low-expression ($I_{\text{NUP98-KDM5A}} = 0 \sim 20$ au) and high-expression ($I_{\text{NUP98-KDM5A}} = 20 \sim 200$ au) groups. As a control, we performed the same analysis on the DAPI DNA stain channel in the same images (Figure 3 - Figure Supplement 1 [S1](#)), which show low values for all local DAPI signals and flat trends.

Taken together, these results indicate that, although a high expression level of NUP98-KDM5A can drive condensate formation at all gene loci, a low expression level leads to preferential formation of condensate at loci with high H3K4me3 density, such as the leukemogenic HOX gene clusters.

The local density of H3K4me3 correlates with expression up-regulation in patient cells

Having established that NUP98-KDM5A forms condensates at H3K4me3-rich loci in a concentration-dependent manner in HEK293 cells, we next examined how NUP98-KDM5A affects gene expression at potential target loci under patient-relevant NUP98-KDM5A expression levels and conditions. We analyzed published single-cell RNA-seq and H3K4me3 CUT&RUN datasets from primary leukemia patients harboring the NUP98-KDM5A oncofusion (Umeda et al., 2025 [U](#)). We analyzed each patient separately, i.e., one with acute myeloid leukemia (Patient 1; AMKL) and one patient with acute lymphoblastic leukemia (Patient 2; ALL), as the transcriptome profiles of these leukemias are distinct (Umeda et al., 2025 [U](#)).

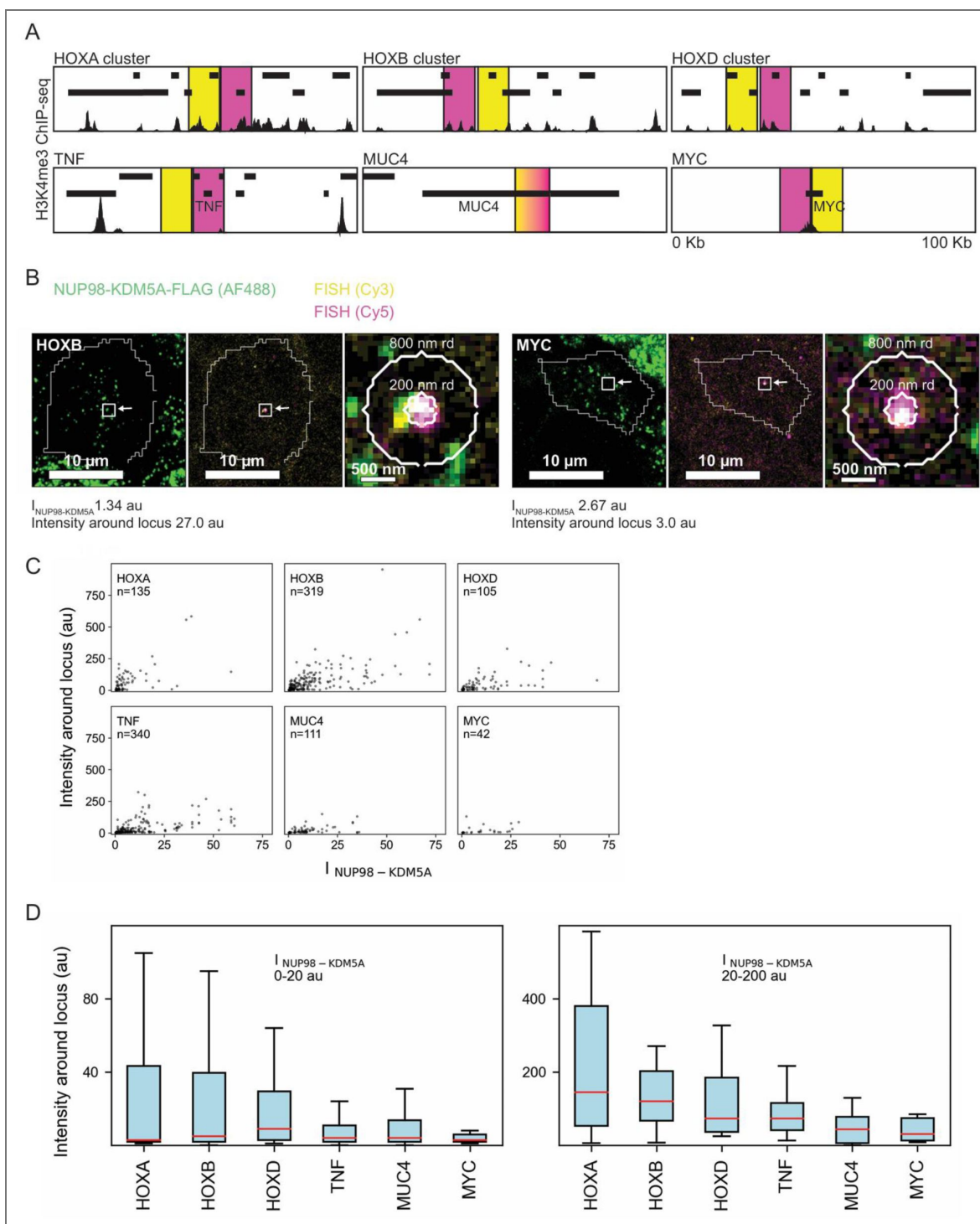


Figure 3. NUP98-KDM5A condensate formation at loci with varying H3K4me3 density.

A. Schematic of the 100 kb genomic region surrounding each FISH-labeled locus. Bars indicate genes. Black peaks at the bottom show H3K4me3 ChIP-seq signal in HEK293 cells. FISH probe targeted regions are marked in magenta (Cy5-labeled probes) and yellow (Cy3-labeled probes). **B.** Images of combined FISH (magenta and yellow) and NUP98-KDM5A immunofluorescence (green) in HEK293 cells for the HOXB and MYC loci, showing a single z-slice with white lines marking the nucleus based on DAPI signal segmentation. Circles in the zoomed-in views of the boxed regions mark the radii for maximum and background signal calculation. **C.** Local NUP98-KDM5A signal (intensity around the locus) across expression levels as measured by mean NUP98-KDM5A nuclear intensity, $I_{\text{NUP98-KDM5A}}$ (HOXA, $n = 135$ loci; HOXB, $n = 319$; HOXD, $n = 105$; TNF, $n = 340$; MUC4, $n = 111$; MYC, $n = 42$). **D.** Boxplots of local NUP98-KDM5A signal for cells with low ($I_{\text{NUP98-KDM5A}} = 0 \sim 20$ au) and high ($I_{\text{NUP98-KDM5A}} = 20 \sim 200$ au) expression levels. The median is shown in red. The box extends from the first to the third quartile, and whiskers extend to the most extreme data points located within 1.5x the interquartile range.

First, we assessed fold changes in gene expression in patients compared to age- and cell-matched healthy individuals, individually for each patient due to the inherent heterogeneity of leukemia transcriptomes. Loci were defined as 100 kb windows (± 50 kb) centered on transcription start sites (TSSs) (Figure 4 - Figure Supplement 1A [↗](#)). Interestingly, previous studies have described broad 60-100 kb as critical descriptors of cell identity (Benayoun et al., 2014 [↗](#)). Volcano plots indicate that most genes in both patients are upregulated relative to healthy individuals. Although most loci are upregulated, consistent with the proposed gene-activating function of NUP98-KDM5A condensates, a subset is downregulated, likely reflecting complex interactions within gene regulatory networks that lead to repression at specific loci (Figure 4A [↗](#)).

Next, we examined how H3K4me3 levels at loci correlated with fold changes in gene expression in the two patients in our study. In both patients, the majority of dysregulated genes were heavily H3K4me3-marked (Figure 4A [↗](#)). We observed a clear positive correlation, with the most highly H3K4me3-marked loci also being the most upregulated (Figure 4B [↗](#)). To better visualize the relationship between the presence of H3K4me3 and gene expression changes at target loci, we ranked TSS-centered loci according to their H3K4me3 content (for all genes filtered by FDR-adjusted p -value ≤ 0.05 ; rank 1 representing the highest intensity) and quantified their fold-change in expression relative to healthy individuals. This analysis revealed a clear positive correlation, with the most highly H3K4me3-marked loci (left) exhibiting the strongest upregulation. Notably, the loci with the highest H3K4me3 levels corresponded to the HOXA and HOXB clusters. Furthermore, fold changes at the HOXA and HOXB clusters exceeded those of other upregulated genes (Figure 4C [↗](#)). Notably, plots from the two patients reveal distinct transcriptional programs in NUP98-KDM5A-driven cancers, likely reflecting differences in developmental stage and unique passenger mutations in the precursor cells.

Taken together, our results show that dysregulated genes in the patients are also heavily H3K4me3-marked, predominantly up-regulated and associated with leukemogenesis. Among these, the HOXA and HOXB clusters stand out, exhibiting the highest H3K4me3 levels in patient cells, consistent with their well-established role in NUP98-KDM5A-driven leukemia (Noort S et al., 2021). This correlative trend, together with the observation that gene expression saturates over a length scale of approximately ± 50 kb (Figure 4 - Figure Supplement 1 [↗](#)), further supports selective enhancement of condensate formation in the presence of the NUP98-KDM5A oncofusion until available H3K4me3 sites become saturated.

Discussion

Our results revealed key interactions associated with the formation of NUP98-KDM5A condensates in euchromatin loci relevant for leukemogenic transformation. Specifically, we showed that the formation of NUP98-KDM5A condensates involves a combination of chromatin binding and phase separation, and that the interplay between these distinct types of interactions enhances NUP98-KDM5A condensation. Additionally, our findings address long-standing questions on how condensates of NUP98 fusions with chromatin readers are targeted to specific chromatin loci. Despite being a pathological protein, condensates formed by NUP98-KDM5A show several features that could be generalizable to physiological chromatin-associated condensates, e.g. those formed by transcription factors. First, NUP98-KDM5A condensates show clear gel-like properties with low exchange dynamics, contrasting earlier impressions in the biomolecular condensate field that liquid-like fluidity or fast exchange dynamics are required for their functions. In fact, it is now recognized that certain chromatin-protein co-condensates can be strongly solid-like, such as those formed in olfactory neurons to enable the silencing of all but one olfactory receptor genes (add reference: Pulupa Nature). Interestingly, instead of silencing, NUP98-KDM5A condensates formation actually leads to the activation of targeted genes. The gel-like properties of NUP98-KDM5A condensates may enable this role not only by enriching activating factors, but also by excluding repressive factors, analogous to the selectivity of gel-like condensates in the nuclear pore formed by the same FG-repeat IDRs from NUP98 and other nucleoporins (Hulsmann et al., 2012 [↗](#)). This exclusion could be easily overlooked because it is more difficult to measure by proteomics methods than enrichment.

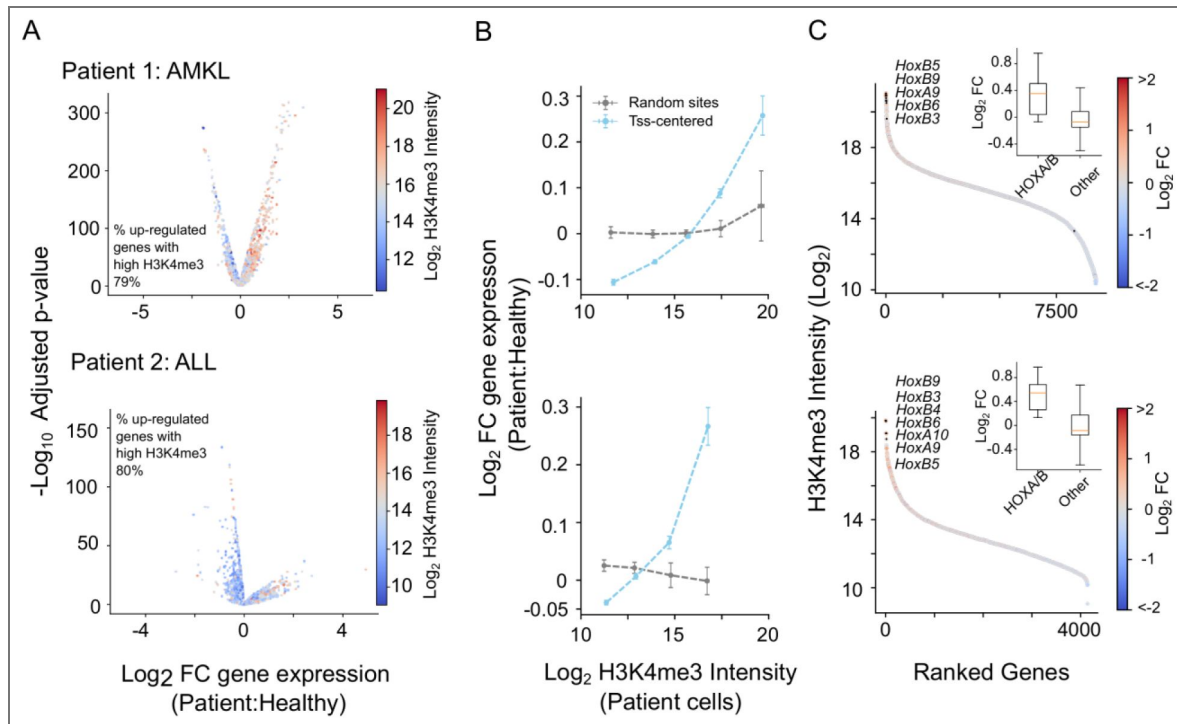


Figure 4. Degree of expression correlates with H3K4me3 intensity for the dysregulated genes in NUP98:KDM5A leukemia patients.

A. Volcano plots indicate a change in gene expression in patients compared to healthy donors, with color encoding the degree of H3K4me3 intensity. Rows correspond to patients. **B.** Correlation plot of change in gene-expression with H3K4me3 intensity in patients with genes binned by methylation signal and expression fold change. *Blue*: intensities computed across TSS-centered windows, and *Gray*: across random sites for comparable window size. Error bars: SEM. **C.** Rank-ordered H3K4me3 intensity plot with color representing the fold-change in the gene-expression. Top ranked HOXA and HOXB protein coding genes with $\log_2$ fold change ≥ 0.5 are indicated. *Inset*: boxplot represents the fold-change in gene-expression for HOXA/B coding genes compared to all the other genes in the ranked set (Patient 1: $n = 16$ for HOXA/B and 9205 for the rest; Patient 2: $n = 12$ for HOXA/B and 4125 for the rest).

Second, NUP98-KDM5A condensates are dispersed into small, sub-diffraction-limit sizes that are associated with the cognate H3K4me3 domains. Interestingly, in the cell nucleus, this small size does not rely on their gel-like properties, because the N-terminally tagged mEGFP-NUP98-KDM5A shows fast FRAP recovery but similar morphologies as the gel-like C-terminally tagged NUP98-KDM5A-mEGFP. Instead, chromatin binding is key for this appearance because nuclear condensates formed by the chromatin-binding-deficient mutant are substantially larger (lower condensate density but similar fraction of proteins in condensates). It could be imagined that binding to specific chromatin sites could also enable other transcription factors to form sub-diffraction-limit condensates (or clusters) that are challenging to detect by conventional fluorescence microscopy techniques.

Finally, the formation of small, chromatin-associated NUP98-KDM5A condensates at specific loci depends on both expression level and binding-site density. For instance, it remains poorly understood how chromatin readers recognize critical target loci. Although histone marks are known to mediate the association of epigenetic readers with chromatin, these marks are often widespread across the genome, and the sources of binding specificity remain unclear (Lukauskas et al., 2024 [↗](#); Yun et al., 2011 [↗](#)). Binding of epigenetic readers at specific loci is frequently the first step in feedback loops that impact gene expression and ultimately cell fate decisions (Atlasi & Stunnenberg, 2017 [↗](#); Blackledge & Klose, 2021 [↗](#); Meissner, 2010 [↗](#)). Here, we show that chromatin association at specific loci emerges from the combined effects of NUP98-KDM5A expression level and local H3K4me3 density. At low expression levels, NUP98-KDM5A preferentially associates with highly H3K4me3-marked loci, whereas this preference diminishes as expression levels increase. These observations indicate a high degree of specificity of NUP98-KDM5A for loci enriched in the H3K4me3 mark. Furthermore, this suggests a potential general mechanism underlying transcription factor and epigenetic reader specificity and highlights the importance of tightly controlling the expression levels of chromatin-associated factors, as the abundance of these proteins serves as a critical regulatory knob for chromatin binding and the activation of downstream gene expression programs.

Materials and Methods

DNA constructs for cell expression

We used the Mammalian Toolkit (MTK) system (Fonseca et al., 2019 [↗](#)) to generate constructs expressing variants of the NUP98-KDM5A oncofusion in tissue culture. MTK enables Golden Gate-based assembly of modular parts into complete transcriptional units for expression via transient transfection. First, we generated a library of modular parts (NUP98 fragment, KDM5A-derived fragments, fluorescent proteins) either by PCR amplification or IDT DNA synthesis, ensuring that none of the sequences contained Esp3I or BsaI recognition sites. This “domestication” step is required because Esp3I and BsaI are the type IIS restriction enzymes used for cloning modular parts and for assembling transcriptional units, respectively. Esp3I digestion enables insertion of PCR-amplified sequences into domesticated parts. In contrast, BsaI digestion of the domesticated parts generates complementary sticky ends—designated 3a, 3b, and 4a—that direct ordered assembly of the transcriptional unit into the MTK acceptor expression vector (MTK0). The acceptor vector we used carries a pGK promoter, polyadenylation tail, and ampicillin resistance cassette. During the Golden Gate assembly, the 3a overhang anneals to its corresponding site in the acceptor vector, the 3b part ligates between the 3a and 4a overhangs, and the reaction closes with the 4a overhang, producing a correctly assembled circular expression vector suitable for expression of NUP98-KDM5A upon transient transfection. A list of acceptor vectors used in this study is provided in [Supplementary Table 1](#) [↗](#). Construct sequences are provided in folders with the suffix `_DNASequences` in Dryad repository. Plasmids were purified using the BioBasic EZ-10 Spin Column Plasmid DNA Mini-Preps Kit (#C861B24). Cloning was performed in NEB Stable competent E.coli (NEB; C3040H).

The NUP98-KDM5A fusion sequence was derived from the cDNA sequences retrieved from NCBI Nucleotide (GenBank) for NUP98 isoform 1 (accession: NM_016320.5) and KDM5A (accession: NM_001042603.3). The fusion junction was selected as indicated in [Figure 3](#) [↗](#) of (Van Zutven et al.,

2006 [\[1\]](#)), incorporating amino acids 1-513 from NUP98 and 1486-1690 from KDM5A into the oncofusion. Synonymous nucleotide substitutions were introduced to the constructs to ensure compatibility with MTK Golden Gate cloning. Plasmid sequence files for plasmids used in mammalian cell expression experiments are stored in Figure1_DNASequences and Figure3_DNASequences folders in the Dryad repository.

Cell lines and tissue culture

For experiments in the results section “NUP98-KDM5A forms condensates at H3K4me3-marked chromatin in a concentration-dependent manner”, we used U2OS cells (human osteosarcoma) purchased from ATCC (Catalog Number 50-238-4124). For results sections “NUP98-KDM5A preferentially forms condensates at H3K4me3-enriched loci” we used HEK 293 FT cells (human embryonic kidney) purchased from Invitrogen (Catalog Number R70007). Cell lines were maintained under standard culture conditions in standard T-25 or T-75 tissue culture-treated surface flasks in Dulbecco’s Modified Eagle Medium (DMEM; Gibco #10566024) supplemented with 10% Refiltered Fetal Bovine Serum (FBS; UCSF Media Production #CCFAP004) and 1% Penicillin-Streptomycin (11 mg/mL; UCSF Media Production #CCFGK004). Cultures were maintained in a mammalian cell culture humidified incubator at 37 °C adjusted to 5% CO₂. Cells were re-plated every 3-5 days after one Phosphate Buffered Saline (PBS) (Gibco #20012027) wash and 5 min enzymatic cell detachment using Trypsine-EDTA 0.25% in Saline A (UCSF Media Production #CCFGP003), followed by gentle resuspension in culture media.

In vitro reconstituted condensates: DNA constructs, protein expression, and purification

pET28a-based expression vectors were generated to produce the full-length NUP98-KDM5A oncofusion and other C-terminally truncated variants of NUP98-KDM5A ([supplementary table 2](#) [\[1\]](#)). All expression constructs were generated by fast cloning using synthetic geneblock fragments (IDT) that were codon optimized for expression in BL21 (DE3) *E. coli* cells (C. Li et al., 2011 [\[2\]](#)). Plasmid sequence files for plasmids used in bacterial protein purification experiments are stored in Figure2_DNASequences folder in the Dryad repository.

In vitro reconstituted condensates: Protein expression and purification

The protein expression and purification strategies for NUP98-KDM5A and other NUP98-containing constructs were carried out as previously described for the related NUP98-HOXA9 oncofusion and the NUP98 segment alone (Chandra et al., 2022 [\[3\]](#); Ibáñez De Opakua et al., 2022 [\[4\]](#); Schmidt & Görlich, 2015 [\[5\]](#)). All proteins used in this study were expressed in BL21(DE3) competent *E. coli* cells (NEB) that were transformed with with pET28a-based plasmids and grown at 37 °C to an optical density at 600 nm (OD₆₀₀) of approximately 0.8 in Luria-Bertani (LB) medium supplemented with Kanamycin. At an OD₆₀₀ of ~0.8, protein expression was induced by the addition of 1 mM IPTG (GoldBio) and the bacterial cultures were incubated at 37 °C for 4 hours. For protein purification, 8 liters of bacterial cultures were harvested by centrifugation (6,000 × g for 15 minutes at 4 °C). Bacterial cells were resuspended in Buffer A (20 mM Tris, 500 mM NaCl, 1 mM (tris(2-carboxyethyl)phosphine) (TCEP), 0.1 % Triton X-100, pH 7.4) and lysed by sonication on ice. The cell lysates were clarified by centrifugation at 30,000 × g at 4 °C and the supernatant was discarded as all the constructs of interest were found in the insoluble fraction. The inclusion bodies obtained were resuspended in extraction buffer (Buffer A supplemented with 6 M guanidine hydrochloride), followed by sonication and centrifugation. After the second centrifugation step, the supernatant of the inclusion bodies dissolved in extraction buffer was loaded onto a nickel-nitrilotriacetic acid (Ni-NTA) column (Thermo Fisher Scientific). The bound proteins were washed with 5 column volumes of extraction buffer containing 50 mM Imidazole and eluted with extraction buffer containing 500 mM Imidazole. Fractions containing the proteins of interest were concentrated and purified by size exclusion chromatography using a Superdex 200 16/600 column that was equilibrated with extraction buffer. High purity fractions were

Supplementary Table 1.

Construct	Plasmid	Used in Figures
NUP98(1-513)-KDM5A(1486-1690)-mEGFP	MTK0-pGK	Figure 1, Figure 2
mEGFP-NUP98(1-513)-KDM5A(1486-1690)	MTK0-pGK	Figure 2
NUP98(1-513)-KDM5Amut(1486-1690)-mEGFP	MTK0-pGK	Figure 1
NUP98(1-513)-KDM5A(1486-1690)-mCherry	MTK0-pGK	Figure 1
NUP98(1-513)-KDM5A(1486-1690)-FLAG	MTK0-pGK	Figure 3
mEGFP-KDM5A(full length)	MTK0-pGK	Figure 1
mEGFP-KDM5Amut(full length)	MTK0-pGK	Figure 1
mEGFP-KDM5A(1486-1690)	MTK0-pGK	Figure 1

Supplementary Table 2.

Construct	Plasmid	Used in figures
6xHis-NUP98 (1-513)-KDM5A (1486-1690)	Pet28a	Figure 2
6xHis-NUP98 (1-513)-KDM5A (1486-1690) W1625A	Pet28a	Figure 2
6xHis-MBP-PHD3 (1601-1662)	Pet28a	Figure 2
6xHis-NUP98 (1-513)-KDM5A (1601-1662)	Pet28a	Figure 2
6xHis-NUP98 (1-513)-KDM5A (1486-1600)	Pet28a	Figure 2
6xHis-NUP98 (1-513)	Pet28a	Figure 2

identified by SDS-PAGE, concentrated by ultracentrifugation using a 30 kDa molecular weight cutoff membrane, flash frozen and stored at -80°C . Protein concentrations were determined using the following extinction coefficients at 280 nm: NUP98 (1-513)-KDM5A (1486-1690): $27430\text{ M}^{-1}\text{cm}^{-1}$, NUP98 (1-513)-KDM5A (1601-1662): $18950\text{ M}^{-1}\text{cm}^{-1}$, NUP98 (1-513)-KDM5A (1486-1600): $12950\text{ M}^{-1}\text{cm}^{-1}$ and NUP98 (1-513): $5960\text{ M}^{-1}\text{cm}^{-1}$.

***In vitro* reconstituted condensates: AF488 labelling of proteins**

All proteins were labelled using Alexa Fluor 488 TFP ester reactive dyes according to the manufacturer's protocol. Briefly, proteins were diluted to 1 mg/mL using PBS (136.9 mM NaCl, 2.68 mM KCl, 10 mM Na_2HPO_4 1.7 mM KH_2PO_4 , pH 7.4) and then mixed 10:1 with 1 M sodium bicarbonate. The labelling reactions were carried out at a dye:protein molar ratio of 10:1. The reaction mixtures were then incubated for 15 minutes at room temperature. To remove any unbound dye from the protein solution, the reaction mixtures were dialyzed for 16 hours at 4°C against the extraction buffer at 5,000 times the volume of the sample using a 2 kDa molecular weight cutoff membrane. For labelled proteins, the concentration and degree of labelling are based on extinction coefficients at 280 nm and were determined by measuring the absorbance at 280 nm and 494 nm.

***In vitro* reconstituted condensates: Fluorescence polarization assay**

The binding of unlabeled NUP98-KDM5A or NUP98-KDM5A W1625A to a H3K4me3 10mer peptide was measured either by direct or competition-based fluorescence polarization assays (FP). For the direct FP binding assays, increasing protein concentrations of up to 20 μM were incubated for 30 minutes at room temperature with 10 nM of a C-terminal fluorescently labelled H3K4me3 peptide. To determine the dissociation constant (K_d) for the H3K4me3 peptide, the raw data from direct FP measurements were fitted to a one site-total Nonlinear regression curve fit equation (GraphPad Prism 9.5.1, GraphPad Software, La Jolla, CA). For competition-based FP assays, 2 μM NUP98-KDM5A or MBP-PHD3 (KDM5A residues 1601-1662) was incubated for 30 minutes at room temperature with 10 nM of C-terminal fluorescently labeled H3K4me3 peptide in the presence of increasing concentrations of unlabeled competitor peptides. IC_{50} values were determined by fitting the raw data from competition-based FP measurements to an inhibitor versus response (three parameters) nonlinear regression curve fit equation (GraphPad Prism 9.5.1, GraphPad Software, La Jolla, CA). Direct or competition-based FP assays were carried out in 1X PBS (136.9 mM NaCl, 2.68 mM KCl, 10 mM Na_2HPO_4 , 1.7 mM KH_2PO_4 , pH 7.4) supplemented with 2.5 mM DTT and performed in triplicates.

Nucleosome assembly and *in vitro* phase separation assays:

Generation of AF647 labelled H3K4me3 or H3K4me0 nucleosomes

DNA was purified, labelled and assembled onto histone octamers as previously described (Luger et al., 1999 [\[4\]](#); Moore et al., 2025 [\[5\]](#)). In brief, a 3.5 kb insert consisting of 3.2 kb from the 5' end of the mouse *Cyp3a11* gene, 300 bp of plasmid DNA, and a 5' overhang, was purified by size-exclusion chromatography and labeled with Alexa Fluor 647-aha-dCTP (Thermo Fisher Scientific) using the Klenow fragment (Moore et al. 2025 [\[5\]](#)). Unmodified (H3K4me0) or histone H3 lysine 4 trimethylated (H3K4me3) histone octamers (two each of human histones H2A, H2B, H3, and H4), made from recombinant histones expressed in *E. coli*, were purchased from EpiCypher. AF647 labelled DNA was assembled onto H3K4me0 or H3K4me3 histone octamers using a salt gradient dialysis at a histone octamer:DNA molar ratio of 17.7:1 (Luger et al. 1999 [\[4\]](#)). Following assembly, samples were dialyzed against TCS buffer (20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT) and stored at 4°C . DNA concentration was determined after assembly via Nanodrop and the molecular weight of the double stranded sequence, along with the histone octamer:DNA molar ratio, was used to calculate the nucleosome concentrations. Nucleosome assembly was assessed by protection of DNA from enzymatic digestion by Micrococcal nuclease.

Nucleosome assembly and *in vitro* phase separation assays:

Preparation of microscopy plates

384-well glass bottom SenoPlates (Greiner Bio-One) were mPEGylated and passivated with Bovine serum albumin (BSA) as previously described (Hihara et al. 2012). Briefly, microscopy plates were washed and incubated with 2% Hellmanex for 1 hour at room temperature. After several washes with H₂O, wells were incubated with 0.5 M NaOH for 30 minutes and then rinsed with H₂O. For mPEG passivation, wells were incubated overnight at 4 °C with a solution of 5K mPEG-silane (Laysan Bio) at 20 mg/mL dissolved in 95% Ethanol. The wells were washed twice with 95% Ethanol, followed by three washes with H₂O, prior to the addition of a 100 mg/mL solution of BSA dissolved in H₂O. To remove the BSA solution, wells were washed six times with H₂O. After three additional washes with phase separation assay buffer (1X PBS (136.9 mM NaCl, 2.68 mM KCl, 10 mM Na₂HPO₄, 1.7 mM KH₂PO₄, pH 7.4) supplemented with 2.5 mM DTT), the plates were ready for use.

Nucleosome assembly and *in vitro* phase separation assays: *In vitro* phase separation assay

In vitro assays were prepared as previously described with minor modifications and protein concentrations expressed as mixtures of AF488 labeled:unlabeled proteins at a ratio 1:10 mixtures (Chandra et al. 2022 [DOI](#), Ruan et al. 2025). Briefly, unlabeled stocks were first diluted to 300 μM in extraction buffer and then gradually diluted to 1.5 M guanidine hydrochloride with H₂O before mixing with labeled samples to make a 37.5X mixture at 1:10 (labeled to unlabeled). The 1:10 protein mixture was further diluted to 4.75X (0.1M guanidine hydrochloride) in H₂O and serial dilutions in the same solvent were carried out. The 1:10 protein mixtures were then mixed with phase separation assay buffer (1X PBS (136.9 mM NaCl, 2.68 mM KCl, 10 mM Na₂HPO₄, 1.7 mM KCl, pH 7.4) supplemented with 2.5 mM DTT) to achieve the final concentrations and negligible levels of guanidine hydrochloride (0.02 M).

For all *in vitro* phase separation reactions, samples were incubated at room temperature overnight before imaging in mPEG/BSA passivated 384-well SensoPlates (Greiner Bio-One). Phase separation reactions with increasing concentrations of proteins (0-750 nM, Supplementary Figure 2A) were prepared as described above, where 2-fold dilutions were carried out from the 1:10 protein mixture at 4.75X prior to mixing with phase separation assay buffer. For reactions in supplementary figure 2 B and C, proteins (NUP98-KDM5A or NUP98-KDM5A W1625A) at a concentration of 750 nM were mixed with 0 or 2 μg chicken erythrocyte polynucleosomes in phase separation assay buffer. Similarly, *in vitro* FRAP experiments (Supplementary Figure 2E) were carried out with 1 μM NUP98-KDM5A mixed with 0 or 2 μg chicken erythrocyte polynucleosomes. Reactions with increasing concentrations of AF647-labeled H3K4me3 or H3K4me0 nucleosomal arrays (0-100 nM) (Figure 2 D-F [DOI](#) and Supplementary Figure 2 F-H) were prepared by mixing varying concentrations of protein (0-800 nM of NUP98-KDM5A or NUP98-KDM5A W1625A) in phase separation assay buffer.

Nucleosome assembly and *in vitro* phase separation assays:

Microscopy

All images for *in vitro* phase separation assays in this study were acquired at the Center for Advanced Light Microscopy at UCSF. Confocal microscopy images were acquired using a CREST X-Light V2 large field of view spinning disk confocal unit, 100 × 1.40 NA oil objective, and Prime 95B 25mm sCMOS camera. Photobleaching was carried out with an Opti-Microscan FRAP unit using a 405nm laser line at 10% of the maximum laser intensity. Confocal images of *in vitro* reconstitutions are located in folder Figure2_ConfocalImaging_inVitro in the Dryad repository.

Cell FRAP

Confocal live imaging coupled with fluorescence recovery after photobleaching (FRAP) of U2OS cells expressing mEGFP-NUP98-KDM5A and NUP98-KDM5A-mEGFP was performed in an imaging chamber maintained at 37 °C and 5% CO₂, mounted on an Andor Borealis CSU-W1 Spinning Disk Confocal microscope. We used a Plan Apo VC 100x/1.4 Oil objective. Imaging was carried out in a single channel using a 488 nm laser at 40% power (GFP channel, 200 ms exposure) with 1 × 1 binning. Images were acquired using a Zyla sCMOS camera (16-bit bit depth). For each field of view, time-lapse movies of 2048 × 2048 pixel images (pixel size 0.065 μm) were collected at a temporal resolution of approximately 220 ms per frame for over 30 s.

Photobleaching was performed using a 405 nm laser (60 mW) with a 200 ms exposure time. Bleaching was applied approximately 3 s after the start of acquisition (~10 frames) using four sequential laser pulses targeting a rectangular region of interest (~150 × 50 pixels). Raw images used for Figure 2 - Figure Supplement 2 [2](#) - A can be found in Figure2_ConfocalImaging_inVitro\Figure2-FigureSupplement2TIFF\A in the Dryad repository.

Image analysis

Image analysis of condensates formed *in vitro* was carried out using FIJI (Version 1.54p) (Schindelin et al. 2012). Mean pixel intensities for protein condensates were calculated in FIJI from .tiff files collected under identical microscopy settings. Condensates were picked using an automatic threshold (3%) and the fraction of signal in condensates was determined by dividing the mean fluorescent protein intensity within condensates relative to the intensity in the entire image.

FRAP recovery curves for the photobleached region of interest (ROI) in both *in vitro* and in-cell experiments were generated by subtracting background fluorescence from the measured recovery signal and normalizing the data to the pre-bleach intensity. The FRAP recovery curves were fitted by non-linear regression using a one-phase exponential association model (GraphPad Prism 9.5.1, GraphPad Software, La Jolla, CA). In this model, the mobile fraction (M_F) corresponds to the plateau value, while the half-life ($T_{1/2}$) represents the time required for the fluorescence intensity to reach 50% of the maximum of the plateau level.

Transfection procedure for cell imaging experiments

Sample preparation was performed as follows: 15 000–25 000 cells were plated into each of the wells of a Cellvis 8-chambered coverglass system (#1.5 polymer coverslip; Product#: C8-1.5P) for confocal and SoRa imaging, or Lab-Tek II 8-chambered coverglass system (#1.5 borosilicate coverglass; 155409) for Nikon-NSPARC live-imaging. Chambers were pre-treated with Poly-L-lysine solution (Sigma-Aldrich; P4707). Cells were plated in 200 μL culture medium and incubated for 18–24 h to allow attachment. The following day, 150–200 ng of plasmid DNA were transfected using the jetOPTIMUS DNA Transfection Kit (Item No. 101000025) according to the manufacturer's short protocol (<https://www.polyplus-sartorius.com/products/jetoptimus> [2](#)), adjusted for a 200 μL media volume. Cells were incubated for 4–6 hours in the transfection medium, after which the transfection medium was replaced with fresh culture medium and the cells were allowed to recover overnight. The following day, cells were either prepared for live-cell imaging experiments or fixed and processed for immunostaining and FISH protocols.

Live cell imaging conditions

For live imaging, cells transfected with the appropriate NUP98–KDM5A constructs were incubated for 1 hour in FluoroBrite DMEM medium (Gibco #A1896701) supplemented with GlutaMAX (Gibco #35050-061) and SPY650-DNA stain (Cytoskeleton Cat. #CY-SC501). Samples were then imaged at 37 °C and X% CO₂ in cell-culture imaging chambers mounted on the corresponding imaging microscope.

Immunostaining

For immunostaining, cells transfected with the appropriate NUP98–KDM5A–mEGFP or NUP98–KDM5A^{mut}–mEGFP constructs were fixed with 4% PFA (Electron Microscopy 32% Paraformaldehyde, aqueous solution; #50980494) at room temperature for 10 min, PBS washed, and incubated for 1.5 hours at 37 C with a primary rabbit monoclonal antibody against the H3K4me3 histone mark (Cell Signaling Technology; C42D8) diluted 1:200. After PBS washes, samples were incubated for 1 hour at room temperature with a goat anti-rabbit secondary polyclonal antibody conjugated with Alexa Fluor 546 (Thermo Fisher; A11035) diluted 1:1000. Then, cells were stained with DAPI and post-fixed for 10 min at room temperature with 2% paraformaldehyde and 0.1% glutaraldehyde (Electron Microscopy Sciences; Cat #16019). All antibody incubations were performed in an immunostaining that contains PBS, 0.2% Triton X-100 (Sigma; T9284) and 3% Bovine Serum Albumin (BSA; Jackson Immuno Research; Code: 001-000-162), as described in the protocol available on

Figure1_NUP98KDM5A_H3K4me3_correlation\ImmunostainingProtocol in Dryad repository.

Combined FISH and Immunostaining

For FISH combined with immunostaining, cells transfected with a NUP98-KDM5A-FLAG construct were fixed with 4% PFA for 10 minutes at room temperature, washed with PBS, and processed through a FISH protocol based on (Mateo et al., 2019 [DOI](#)). First, samples were gently permeabilized for 10 minutes with 0.5% Triton-X, followed by a 1-hour RNase A (Monarch; Catalog# T3018L) treatment at 37 C and a 5-minute deproteinization step in 0.1 M HCl. After PBS washing, samples were incubated at 42 C for 30 minutes in a pre-hybridization buffer composed of 50% formamide (Millipore; #344206), saline-sodium citrate buffer (SSC; Invitrogen; Ref AM9763), Tween 20 (Sigma; P9416). Next, samples were transferred into a hybridization buffer composed of 50% formamide, SSC, Tween 20, and dextrane sulfate solution (Millipore; S4030) containing two sets of FISH probes targeting contiguous loci (5 ng/μL each; details in FISH probe design section). Genomic DNA was denatured on a heat block at 93 C for 5 minutes, after which samples were incubated at 42 C overnight. The following day, formamide was removed through a series of SSC washes, and the samples were immunostained by incubating them for 1.5 hours at 37 C with DYKDDDDK (Flag) Tag recombinant rabbit monoclonal antibody conjugated with Alexa Fluor Plus 488 (Thermo Fisher; 8H8L17) diluted 1:200 in immunostaining buffer. We performed a post-hybridization post-fixation step using 2% paraformaldehyde and 0.1% glutaraldehyde for 10 minutes at room temperature. Finally, we hybridized the secondary pairs of FISH probes (100 nM each), tagged with Cy3 and Cy5 and corresponding to each contiguous set of primary probes, by incubating the samples in ethylene carbonate (Sigma Aldrich; E26258) buffer. Samples were then washed with SSC buffer and 30% formamide, and stained with DAPI. For detailed protocol, see Figure3_FISHProtocol in Dryad repository.

FISH probe design

FISH probes were designed by first retrieving pairs of 10-kb contiguous genomic regions corresponding to portions of the loci of interest: HOXA cluster (chr7:27197026-27207050 and chr7:27186295-27196321), HOXB cluster (chr17:46653759-46663990 and chr17:46665068-46675304), HOXD cluster (chr2:176957019-176967205 and chr2:176968092-176978322), TNF (chr6:31529154-31539396 and chr6:31539895-31550008), MUC4 (this probe is the oligo CCCCTCTCCTGTACCGACA CTCCTCAGCATCCACAGG, a repetitive sequence distributed among 11 kb within MUC4) and MYC (chr8:128739417-128749599 and chr8:128749901-128760113). Genomic regions were selected according to their H3K4me3 density in HEK293 cells, as assessed using the public ChIP-seq dataset “HEK293 H3K4me3 Histone Modifications by ChIP-seq from ENCODE/University of Washington” (Sabo et al., 2004 [DOI](#)). Genomic region sequences were retrieved from UCSC Genome Browser on Human (GRCh37/hg19) (Perez et al., 2025 [DOI](#)) and each region downloaded as a FASTA file. We then used the bioinformatic tool OligoMinerApp (<http://oligominerapp.org/> [DOI](#)) (Passaro et al., 2020 [DOI](#)) to generate approximately 100 probes per 10-kb region, using the following parameters: MIN length 37 bp; MAX length 42 bp; MIN GC 35 %; MAX GC 70 %; MIN Tm 42 C; MAX Tm 47 C; Prohibited

sequences AAAAA,TTTTT,CCCCC,GGGGG; MIN space 0 bp; [Na⁺] 390 mM; [Formamide] 50%. All other parameters were set to default. Finally, we incorporated short overhang sequences into the primary probes to enable hybridization with fluorophore-labeled secondary probes. The overhangs were designed to be complementary to the secondary probe sequences: the “GCTTTAAG GCCGGTCCTAGC-Cy5” and “Cy5-AGCTCACCTATTAGCGGCTAAGG” secondary probes (conjugated to Cy5) hybridize to the overhangs flanking one set of primary probes, whereas the “AGCTGATCGTGG CGTTGATG-Cy3” and “Cy3-TGGACGGTTCCAATCGGATC” secondary probes (conjugated to Cy3) hybridize to the overhangs flanking the second, contiguous set of primary probes. Primary probes were synthesized as 50 pmole oligo pools, and secondary probes were synthesized as 100 nmole modified DNA oligos labeled with Cy3 or Cy5. Both sets of probes were synthesized by IDT. Synthesized probe sequences are provided in the [gene]-IDT .xlsx files within each [gene] folder in folder Figure3_DNASequences\FISHProbes in Dryad repository.

Live imaging with a regular confocal microscope

Confocal live imaging of U2OS cells expressing the NUP98-KDM5A-related constructs shown in Figure 1SA [\[link\]](#) was performed in an imaging chamber maintained at 37 °C and 5% CO₂, mounted on an Andor Borealis CSU-W1 Spinning Disk microscope controlled by the open-source software Micromanager (D. Edelstein et al., 2014 [\[link\]](#)). We used a Plan Apo VC 100x/1.4 Oil objective. Imaging was carried out in two sequential channels using a 488 nm laser at 20% power (GFP channel, 200 ms exposure) and a 640 nm laser at 50% power (Cy5 channel, 500 or 800 ms exposure), with 1 x 1 binning. Images were acquired using an Andor Zyla sCMOS 4.2 camera (16-bit bit depth, rolling shutter, 200 MHz readout speed). For each field of view, 2048 × 2048 pixel images (0.065 μm pixel size) were collected as Z-stacks spanning 14 μm centered around the focal center, with a step size of 0.26 μm, resulting in a total of 54 slices. Raw images in Figure 1 - Figure Supplement 1 [\[link\]](#) are in folder Figure1_Confocal_live_imaging in Dryad repository.

Live imaging with the SoRa spinning disk confocal microscope

SoRa-enhanced resolution live imaging of U2OS cells expressing NUP98-KDM5A constructs tagged with mEGFP at either the N- or C-terminus was performed in an imaging chamber maintained at 37 °C and 5% CO₂, mounted on a Yokagawa CSU-W1 SoRa Spinning Disk Confocal Microscope controlled by the open-source software Micromanager. We used an HC PL APO 100x/1.47 OIL CORR TIRF objective. Imaging was carried out in two sequential channels using a 488nm laser at 1.5 V (30% power) (GFP channel, 200 ms exposure) and a 640 nm laser at 1.5 V (30% power) (Cy5 channel, 200 ms exposure), with 1 x 1 binning. Images were acquired using a Prime BSI (PCIe) sCMOS camera (16-bit bit depth, open shutter) coupled with SoRa optics. For each field of view, 2300 × 2100 pixel images (0.0262 μm pixel size) were collected as Z-stacks spanning 8 μm centered around the focal center, with a step size of 0.26 μm, resulting in a total of 31 slices. Raw images used for Figure 2 - Figure Supplement 2B [\[link\]](#) and C [\[link\]](#) are stored in folder Figure2_SoRa_live_imaging in Dryad repository.

Live imaging with the NSPARC scanning confocal microscope

NSPARC-enhanced resolution live imaging of U2OS cells expressing the NUP98-KDM5A-related constructs was performed in an imaging chamber maintained at 37 °C and 5% CO₂, mounted on a Nikon Ti2-E scanning laser microscope equipped with the Nikon Spatial Array Confocal (NSPARC) detector. We used an oil objective Plan APO λD 60 × oil OFN25 DIC N2-1.42 NA. Imaging was performed sequentially in two emission channels: 524 nm (EGFP channel) or 593 nm (mCherry channel), and 677 nm (Cy5 channel), each at a laser power setting of 3.0. Laser scanning was performed in bidirectional mode with an integration setting of 2 and a pixel dwell time of 0.3 μs, and fluorescence emission was collected using a quad-band emission filter. The virtual pinhole size was set to 1.0, yielding an optical resolution of 0.19 μm. Imaging was performed with a zoom factor of 3.9, resulting in a pixel size of 0.073 μm (Nyquist sampling: 0.083 μm), and a scan area of 1024 × 1024 pixels (76 μm × 76 μm). NSPARC super-resolution mode was enabled, and Z-stacks spanning 6 μm were acquired with 13 slices at 0.5 μm step size. The frame acquisition time was 2.3

s per slice, resulting in a total stack acquisition time of approximately 30 s. Raw images for [Figure 1](#) (A, B, and C), and [Figure 1 - Figure Supplement 2A](#) and [B](#) are stored in folder [Figure1_NSPARC_live_imaging](#) in Dryad repository.

NSPARC imaging of immunostaining samples

NSPARC-enhanced resolution imaging of U2OS cells expressing NUP98-KDM5A-mEGFP or NUP98-KDM5A-mut-mEGFP and immunostained for H3K4me3 was performed using an NSPARC scanning laser microscope. We used an oil objective Plan APO λ D 60 $\times$ oil OFN25 DIC N2-1.42 NA. Imaging was performed sequentially in three emission channels: 452 nm (DAPI channel, laser power 1.0), 524 nm (EGFP channel, laser power 4.0), and 593 nm (AF546 channel, laser power 2.0). Laser scanning was performed in bidirectional mode with an integration setting of 4 and a pixel dwell time of 0.3 μ s, and fluorescence emission was collected using single emission filters. The virtual pinhole size was set to 1.0, yielding an optical resolution of 0.19 μ m. Imaging was performed with a zoom factor of 3.9, resulting in a pixel size of 0.073 μ m (Nyquist sampling: 0.083 μ m), and a scan area of 1024 $\times$ 1024 pixels (76 μ m $\times$ 76 μ m). NSPARC super-resolution mode was enabled, and Z-stacks spanning 4 μ m were acquired with 9 slices at 0.5 μ m step size. Raw images for [Figure 1 D](#), [E](#) and [F](#) can be found in folder [Figure1_NUP98KDM5A_H3K4me3_correlation](#) in Dryad repository.

NSPARC imaging of combined FISH and immunofluorescence samples

NSPARC-enhanced resolution imaging of HEK293FT cells immunostained for NUP98-KDM5A-FLAG and labeled by FISH at specific genomic loci was performed using an NSPARC scanning laser microscope. We used an oil objective Plan APO λ D 60 $\times$ oil OFN25 DIC N2-1.42 NA. Imaging was performed sequentially in four emission channels: 452 nm (DAPI, laser power 3.0), 524 nm (AF488, laser power 3.0), 593 nm (Cy3, laser power 4.0), 699 nm (Cy5, laser power 4.0). Laser scanning was performed in unidirectional mode with an integration setting of 2 and a pixel dwell time of 0.4 μ s, and fluorescence emission was collected using single emission filters. The virtual pinhole size was set to 1.0, yielding an optical resolution of 0.19 μ m. Imaging was performed with a zoom factor of 2, resulting in a pixel size of 0.073 μ m (Nyquist sampling: 0.083 μ m), and a scan area of 2048 $\times$ 2048 pixels (150 μ m $\times$ 150 μ m). NSPARC super-resolution mode was enabled, and Z-stacks spanning 10 μ m were acquired with 53 slices at 0.195 μ m step size (Nyquist sampling on Z). Raw images for [Figure 3B](#), [C](#) and [D](#) and [Figure 3 - Figure Supplement 1A](#) and [B](#) can be found under [gene] folders in [Figure3_NSPARC_FISH_Imaging](#) in Dryad repository.

Calibration between mEGFP and mCherry signals

For mEGFP-to-mCherry calibration, we generated a DNA construct expressing the C-terminal domain of KDM5A (amino acids 1486–1690), which contains the PHD3 chromatin-binding domain and does not form condensates, tagged with mEGFP at the N-terminus and mCherry at the C-terminus. Transient transfection was performed in U2OS cells, taking advantage of the broad range of expression levels produced by transient expression. Transfected cells were imaged using the same protocol and settings employed for NSPARC live-cell imaging. Raw images for [Figure 1](#) - [Figure Supplement 2C](#) are located in [Figure1_mEGFP_mCherry_calibration](#).

Cells were segmented using Cellpose 2.0 (Pachitariu & Stringer, 2022), and mean fluorescence intensities in the mEGFP and mCherry channels were measured within the segmented nuclear masks across a range of expression levels using the script [Figure1_Code\NSPARC-mEGFPmCherry-Calibration.ipynb](#). Output files with the suffix `_cells` and `_parameters` were saved in folder [Figure1_mEGFP_mCherry_calibration\analysis_20241011-NSPARC_U2OSlive-mEGFP-KDM5A-mCherry](#). Then, we used the script [Figure1_Code\Figure1A.ipynb](#) to calculate a regression line relating mEGFP and mCherry intensity values acquired under NSPARC live imaging settings, and generated [Figure 1 - Figure Supplement 2C](#). Paths correspond to the Dryad repository.

EGFP intensity to concentration calibration

Sample preparation for intensity-to-concentration calibration was performed as follows: Sonicated TetraSpeck beads of 0.1 μm diameter (Invitrogen; Catalog #T7279) were diluted 1:1000 in PBS. Purified EGFP (Recombinant Purified Protein; ChromoTek; Cat. EGFP) was then diluted into the PBS-bead solution to final concentrations of 0 nM, 25 nM, 50 nM, 100 nM, 200 nM, 400 nM, 500 nM for SoRa analysis; and 200 nM, 100 nM, 50 nM for NSPARC analysis. EGFP has nearly identical molecular brightness as mEGFP. Each EGFP dilution was plated into a Lab-Tek II 8-chambered coverglass system (#1.5 borosilicate coverglass; 155409) for imaging under the same settings used for SoRa live imaging or NSPARC live imaging. Raw SoRa EGFP calibration data used for Figure 2 - Figure Supplement 2D [is](#) stored in Figure2_SoRa_EGFP_concentration_calibration. Raw NSPARC EGFP calibration data used for Figure 1 - Figure Supplement 2D [is](#) in Figure1_NSAPARC_EGFP_concentration_calibration.

For each concentration of SoRa EGFP concentration to intensity analysis, we recorded one biological replicate and four technical replicates. Each replicate was recorded as 2300×2100 pixel images over a Z-stack of 41 slices with a step size of 0.5 μm , spanning a total of 20 μm , around the bottom of the chamber, as shown by the presence of TetraSpeck beads. Image analysis was performed in a Python-based Jupyter Notebook script Figure2_Code\SoRa-EGFPCalibration.ipynb. For each image in the Z-stack, a 1000×1000 pixel region centered within the field of view was extracted. The mean fluorescence intensity was calculated for each z-slice, and the 20th slice—corresponding approximately to the bottom of the chamber, as indicated by the presence of Tetraspeck beads—was selected. Fluorescence intensity was then quantified as the mean intensity of the 1000×1000 pixel region in the z-slice located 8 slices (approximately 4 μm) above this reference plane, ensuring measurement within the EGFP solution. Background subtraction was performed by subtracting the mean intensity of a 1000×1000 pixel region from a z-slice located 8 planes (approximately 4 μm) below the bottom of the chamber. We read output files, and the script Figure1_Code\FigureS1.ipynb to generate Figure 2 - Figure Supplement 2D [is](#).

As for NSPARC EGFP concentration to intensity analysis, we recorded two biological replicates and three technical replicates for all concentrations. Each replicate was recorded as 1024×1024 pixel images over a Z-stack of 63 slices with a step size of 0.195 μm , spanning a total of 12 μm , around the bottom of the chamber, as shown by the presence of TetraSpeck beads. Image analysis was performed in a Python based Jupyter notebook script Figure1_Code\NSPARC-EGFPCalibration.ipynb. For each image in the Z-stack, a 600×600 pixel region centered within the field of view was extracted. The mean fluorescence intensity was calculated for each z-slice, and the slice with the highest mean intensity—corresponding approximately to the bottom of the chamber as indicated by the presence of tetraspeck beads—was identified. Fluorescence intensity was then quantified as the mean intensity of the 600×600 pixel region in the z-slice located 20 slices (approximately 4 μm) above this reference plane, ensuring measurement within the EGFP solution sample. Because regions outside the sample yield zero intensity values under NSPARC imaging conditions, background subtraction was not required. An output file was generated and stored in Figure1_NSAPARC_EGFP_concentration_calibration\20241202-NSPARC-EGFP_NSAPARCcalibration_output.txt.

We used the output data to generate Figure 1 - Figure Supplement 2D [is](#) through the script Figure1_Code\Figure1A.ipynb. Paths correspond to the Dryad repository.

Image analysis: U2OS NSPARC live imaging

Raw image files were read and analyzed using custom scripts in a Python-based Jupyter Notebook script Figure1_Code\NSPARC-CondensatesQuantification.

For each dataset, the central z-slice was selected manually, and a single slice per dataset was used for downstream analysis. Nuclei were segmented using Cellpose 2.0 on the DAPI channel (Cell pose model: cyto), and nuclear masks were generated. Condensates were segmented from the corresponding fluorescence channel using a Difference of Gaussians (DoG) filter (low sigma: 1.5 pixels; high sigma: 2 pixels), followed by thresholding in DoG space (gaussian threshold for

mEGFP: 1e-05, gaussian threshold for mCherry: 4e-06). Segmented condensates smaller than 4 pixels were excluded from further analysis. Remaining condensates were assigned to individual nuclear masks, and quantitative statistics were computed for both nuclei and condensates. Nuclear measurements included x and y coordinates, total nuclear fluorescence intensity (raw_exp), nuclear area (cell_area), background intensity (background), nucleoplasmic intensity (nuc_intensity), and nucleoplasmic area (nuc_area), where the nucleoplasm was defined as the portion of the nucleus excluding condensates after a 3-pixel erosion. Condensate measurements included x and y coordinates, integrated intensity (raw_mass), size (size), and eccentricity (ecc). The resulting measurements were exported for downstream analysis into separate output .txt files with the suffixes _cells, _condensates, and _parameters, which contain details of the analysis. All files were saved within the Analysis folder of each dataset. Images of the segmented nuclei were saved in the Images subfolder. Figure 1C [Figure 1C](#) and Figure 1 - Figure Supplement 2B [Figure 1 - Figure Supplement 2B](#) were generated with the output files and the script Figure1_Code\Figure1A.ipynb. Paths correspond to the Dryad repository.

Image analysis: U2OS SoRa live imaging

Segmentation and quantification of SoRa-imaged condensates were performed as described for the NSPARC datasets, with modifications, using script Figure2_Code\SoRa-CondensatesQuantification. Nuclei were segmented using Cellpose 2.0 on the Cy5 channel (Cellpose model: cyto). Condensates were segmented using a Difference of Gaussians filter (low sigma: 1.5 pixels; high sigma: 4.5 pixels), followed by thresholding in DoG space (Gaussian threshold for mEGFP: 1e-04). Segmented condensates smaller than 4 pixels were excluded. The nucleoplasm was defined as the portion of the nucleus excluding condensates after a 5-pixel erosion. Nuclear and condensate measurements were identical to those used for U2OS NSPARC live imaging analysis. The resulting measurements were exported for downstream analysis into separate output .txt files with the suffixes _cells, _condensates, and _parameters, which contain details of the analysis. All files were saved within the Analysis folder of each dataset. Images of the segmented nuclei were saved in the Images subfolder. We read output files and generated Figure 2 - Figure Supplement 2C [Figure 2 - Figure Supplement 2C](#) with the script Figure1_Code\FigureS1.ipynb. Paths correspond to the Dryad repository.

Image analysis: Colocalization analysis between NUP98-KDM5A and H3K4me3

Raw image files were read and analyzed using the Python-based Jupyter Notebook script Figure1_Code\NSPARC-NK-H3K4me3-Correlation.ipynb.

For NSPARC microscopy of U2OS cells expressing NUP98-KDM5A-mEGFP or NUP98-KDM5Amut-mEGFP and immunostained for H3K4me3, images were analyzed using Python-based Jupyter Notebook scripts. After manual selection of the central Z-slice, channel alignment between NUP98-KDM5A and H3K4me3 was corrected using phase cross-correlation. This method leverages the cross-power spectrum and inverse Fourier transform to estimate inter-channel offsets with subpixel accuracy. Following channel alignment, nuclear masks were generated using Cellpose 2.0 (model: cyto) applied to the H3K4me3 channel after Gaussian smoothing. Fluorescence intensities from the H3K4me3 and NUP98-KDM5A channels were extracted from each individually segmented nucleus, flattened into one-dimensional vectors, and zero-valued pixels were discarded. Pearson's correlation coefficient was then computed to quantify pixel-level linear association between the two signals. For visualization, two-dimensional histograms of pixel intensity values were generated (x-axis: mEGFP-NUP98-KDM5A intensity; y-axis: H3K4me3-AF546 intensity) and rendered using logarithmic color normalization to emphasize differences in pixel-density distributions. The highest density typically appeared near the origin, corresponding to the background signal. We generated output files named megfp_546_with_iou.txt and megfpmut_546_with_iou.txt, which contain tabulated measurements including the Pearson correlation coefficient between channels (Pearson_r), mean nuclear fluorescence intensity (Expression), and the intersection-over-union (IoU) coefficient between the two channels. Two-

dimensional histograms were saved in folders named 2D_histograms, located in the same directory as the corresponding output files. These files are located in Figure1_NUP98KDM5A_H3K4me3_correlation\NSPARC_Nup98_KDM5A_Correlation_Analysis. Output files were analyzed with Python-Jupyter notebook script Figure1_Code\Figure1B.ipynb to generate Figure 1F [\[link\]](#). Paths correspond to the Dryad repository.

Image analysis: FISH analysis

NSPARC images of immunostaining and DNA FISH samples were analyzed. Images were acquired at 2048×2048 pixels, corresponding to a field of view of $150 \mu\text{m} \times 150 \mu\text{m}$, as Z-stacks spanning $10 \mu\text{m}$ with 53 slices. Each dataset contained four channels: DAPI (nuclear marker), AF488 (NUP98–KDM5A condensates), and Cy3 and Cy5 channels corresponding to contiguous FISH probes labeling specific genomic loci. For selected datasets, a minimum 3D filter was applied with Fiji/ImageJ to enhance nuclear contrast prior to segmentation (kernel size: $x = 3.0$, $y = 3.0$, $z = 1.0$). These datasets are denoted by “m3d” in the segmentation mask file title. Three-dimensional nuclear segmentation was then performed using Cellpose version 4.0.3 in 3D mode with the CPSAM mode (Pachitariu et al., 2025 [\[link\]](#)). Masks were saved as .tiff files with the suffix_cp_masks.

To detect FISH-labeled genomic loci, we used the Python-based package Big-FISH (Imbert et al., 2021) implemented in Jupyter Notebook scripts. Big-FISH was applied to the Cy3 and Cy5 channels to localize contiguous genomic loci in three dimensions, extracting their x, y, and z coordinates from the full Z-stacks for downstream spatial analyses. In short, the Cy3 and Cy5 channels were first thresholded at 2 au, followed by three-dimensional Gaussian filtering ($\sigma = 1$ pixel in x, y, and z) for denoising. Local maxima detection was then applied to identify Gaussian-like intensity spots corresponding to FISH signals. A maximum was defined as a pixel whose value remained unchanged before and after application of the local maxima filter, with a minimum distance of 4 pixels enforced between detected maxima. Detected maxima were further filtered by intensity to retain only the brightest peaks ($\text{Cy3} \geq 3$ au; $\text{Cy5} \geq 5$ au). To identify legitimate probe pairs hybridized to contiguous genomic loci, only Cy3 and Cy5 peaks located within 300 nm of each other were retained, based on voxel dimensions of $0.073 \mu\text{m} \times 0.073 \mu\text{m} \times 0.195 \mu\text{m}$ (x, y, z). For each valid probe pair, the midpoint between the Cy3 and Cy5 coordinates in three-dimensional space was defined as the final x, y and z positions of the genomic locus, denoted as cy5cy3_x, cy5cy3_y, cy5cy3_z. Output files, named with the suffix_fish and_fish_supplement were generated for downstream analyses. This part of the analysis was performed using Python-based Jupyter Notebook script Figure3_Code\FISHPairs.ipynb.

To study the association between NUP98–KDM5A and FISH-labeled genomic loci, we read nuclear segmentation masks generated by Cellpose (files with the suffix_cp_masks) and corresponding locus coordinate files (files with the suffix_fish) containing the three-dimensional positions of genomic loci. Nuclear masks were applied to the AF488 channel, which shows NUP98–KDM5A condensates, and the mean NUP98–KDM5A fluorescence intensity within each nucleus was measured (Mean NUP98–KDM5A intensity in cells). FISH spots were then assigned to their corresponding nuclei based on 3D spatial overlap within the cellpose masks, enabling quantification of NUP98–KDM5A signal in relation to the labeled genomic loci.

For quantification of NUP98–KDM5A intensity around individual genomic loci, ellipsoidal volumes were modeled around each detected FISH spot. Because of anisotropic resolution of optical imaging, a spherical neighborhood in physical space corresponds to a flattened ellipsoid along the Z axis in voxel space. Voxel dimensions were $0.073 \mu\text{m}$ in X, $0.073 \mu\text{m}$ in Y, and $0.195 \mu\text{m}$ in Z, and ellipsoid dimensions were adjusted accordingly to accurately sample the local NUP98–KDM5A signal surrounding each locus. For each locus, we computed the mean, maximum, and 20th percentile value of NUP98–KDM5A fluorescence intensities within ellipsoids of radii 100, 200, 300, 400, 800, and 1000 nm. These measurements were saved in output .txt files with the suffix_radial and_radial_supplement. Images of individual FISH spots and surrounding regions were saved in folders named after the corresponding dataset, located alongside the _radial files. This analysis was performed using the script Figure3_Code\RadialAnalysis.ipynb. Figure 3B [\[link\]](#) was generated using a modified version of the script Figure3_Code\RadialAnalysis-Figure3.ipynb.

Finally, the intensity around locus, shown in Figures 3C and D [\[4\]](#), was calculated as the maximum pixel intensity within a 200 nm ellipsoid—representing the local condensate signal—minus the 20th percentile intensity within an 800 nm ellipsoid, which served as an estimate of the local background fluorescence. As a control, the same analysis was repeated using the unrelated DAPI channel instead of the AF488 channel (Figure 3SA [\[4\]](#) and Figure S3B). As expected, the resulting measurements did not follow the same trends observed for AF488. Plots in Figure 3 [\[4\]](#) and Figure 3 - Figure Supplement 1 [\[4\]](#) were generated using the script `Figure3_Code\Figure3.ipynb`. Paths correspond to the Dryad repository.

Single cell RNA seq analysis: Datasets

Single-cell RNA-seq data were analyzed from published datasets comprising pediatric leukemias with NUP98 onco-fusions (Umeda et al., 2025 [\[4\]](#), GEO GSE287716) and bone marrow cells from healthy donors (H. Li et al., 2025 [\[4\]](#), GEO: GSE189161). Two patients with NUP98-KDM5A fusion (SJALL040121 and SJAML061263) were considered, and healthy donor published dataset (H. Li et al., 2025 [\[4\]](#)), matched for age (pediatric patients aged 2, 4, 10, and 12 years). Annotated cell-state were selected based on the dominant cell-types in each patient. The matched comparator for SJAML061263 (referred to as patient 1 here) was composed of erythroid/megakaryocytic progenitors together with early stem/progenitor compartments (i.e., ‘Mk/E-MPP’, ‘Mk-Prog’, ‘E-Prog-1’, ‘E-Prog-2’, ‘E-Prog-3’, ‘My-MPP’, ‘MPP-1’, ‘MPP-2’, ‘MPP-3’, donor age: 2 and 4). While, patient SJALL040121 (referred to as “patient 2”, here) was composed of lymphoid progenitors and early stem/progenitor compartments (i.e., ‘Ly-Prog-1’, ‘Ly-Prog-2’, ‘Ly-Prog-3’, ‘Ly-Prog-4’, ‘Ly-Prog-5’, ‘MPP-1’, ‘MPP-2’, ‘MPP-3’ with donor age: 10 and 12 years). These matched subsets were used for all the downstream analysis. Relevant data was stored in the file `Figure4_CombinedData_Tables_Fig4_Fig4S1_Fig1S3` for further analysis.

Single cell RNA seq analysis: Data processing

Standard single-cell quality control procedures were applied, cells with high mitochondrial transcripts were excluded by (>20%), cells with low gene detection (i.e., < 200 detected genes per cell) and genes detected in at least 3 cells were included. Filtered counts were normalized for sequencing depth to a library size (10,000 counts per cell) and log transformed. No additional correction was applied.

Single cell RNA seq analysis: Estimation of NUP98-KDM5A expression in patient cells

To quantify the relative abundance of fusion partner transcripts we computed KDM5A to NUP98 ratios in each group. This was computed as the ratio of mean expression of KDM5A to the mean expression of NUP98 on library-size normalized counts. Because transcripts are quantified by poly-A capture, transcript identity is determined by the 3' end. As a result, the NUP98-KDM5A fusion transcript is classified as KDM5A by our analysis pipeline. This shifts the KDM5A-to-NUP98 ratio from approximately 1 in healthy donors, who harbor two NUP98 and two KDM5A alleles, to a ratio of nearly 2 in Patient 1 (AMKL) and Patient 2 (ALL), both of which carry one NUP98 allele, one KDM5A allele, and one NUP98-KDM5A fusion allele that we quantified as a second KDM5A allele (Figure 1 - Figure Supplement 3A [\[4\]](#)). We also computed per-cell ratios for cells with detectable NUP98 abundance (>0) and assessed statistical significance using the Wilcoxon test for paired groups, evaluating differences in relative abundance between healthy and patient donors as well as within healthy and patient individuals (Figure 1 - Figure Supplement 3B [\[4\]](#), `Figure4_CombinedData_Tables_Fig4_Fig4S1_Fig1S3`; Data Table 3). Because the KDM5A-to-NUP98 ratio in patients is close to 2, this suggests that NUP98-KDM5A expression is approximately half that of wild-type NUP98. This inference is supported by the fact that NUP98-KDM5A expression is driven by the NUP98 promoter. Given OpenCell measurements indicating that endogenous NUP98 expression in HEK293T cells is ~360 nM (Cho et al., 2022 [\[4\]](#)), we infer that NUP98-KDM5A is expressed at ~180 nM in this system. We observed condensate formation at comparable concentrations in both U2OS and HEK293FT cells.

Single cell RNA seq analysis: Differential gene expression analysis

Differential gene expression was assessed by comparing each patient sample to its matched healthy reference subset using Wald test (diffxpy in scanpy). Genes with FDR-adjusted p-value ≤ 0.05 and absolute fold-change (i.e. Log2 fold-change) ≥ 0.50 defined as differentially expressed genes (DEG files stored at Figure4_CombinedData_Tables_Fig4_Fig4S1_Fig1S3; Data Table 1-2). To account for variability from donors in the healthy cohort compared to a patient for statistical rigor and robustness we additionally did the following: a) DEG analysis for each patient with healthy donors (n=2), and b) with respect to each donor in two ways: i. considering all cells, ii. bootstrapping a sub-sample from both healthy and patients. Thus, for each patient, gene-expression profiles were compared for three scenarios for concurrence in the direction of change in fold expression, and observed high concurrence in our results (data not shown).

Single cell RNA seq analysis: H3K4me3 intensity quantification from CUT&RUN assays

Total H3K4me3 signal per gene in patients was quantified using the published CUT&RUN data from Umeda et al. 2025 [\[1\]](#) (GEO: GSE287298). Gene annotations and transcription start site (TSS) coordinates were obtained from a GTF file (GRCh38.p14). For each gene, a defined window centered on the annotated TSS was used. H3K4me3 peak coordinates from BED files (broad peaks, 2 replicates) were intersected with this window, and gene-level H3K4me3 intensity was calculated from BigWig tracks as the sum of signal intensities across all overlapping peaks within the defined window.

Genes were defined as H3K4me3 marked genes, if detected within ± 1 kbp of the TSS of a gene since H3K4me3 is promoter-proximal mark. Total H3K4me3 intensities for H3K4me3 marked genes was computed by integrating the intensities across all the peaks in the Transcription start site (TSS-centered windows). For the downstream analysis, H3K4me3 marked genes were considered.

Single cell RNA seq analysis: Window-size selection

Total window size was varied from ($\pm$ TSS centered symmetric) 1kbp -75 kbp i.e. (2kbp to 150kb in total) for H3K4me3 intensities calculations, since some genes span H3K4me3 breadth up to 60-100kbp (Benayoun et al., 2014 [\[2\]](#)). Cut-off or saturating window size was determined as the window size that first crosses the saturating H3K4me3 intensity mark defined as 90% of maximal value. High H3K4me3 intensity threshold was arbitrarily defined as 60th percentile from the pooled H3K4me3 signal distribution across all the window sizes.

Single cell RNA seq analysis: Correlation between Gene-expression and H3K4me3 methylation

Genes defined as H3K4me3 marked genes were used to evaluate the relationship between the gene-specific H3K4me3 intensity in patients (at ± 50 kb window centered around TSS), and the relative change in its expression. Data was binned, mean and standard error of mean computed within each bin. For a control baseline, for each gene, we randomly sampled a loci on the same chromosome with the same window size, excluding the TSS centered window and similarly computed the binned statistics (Data as in Figure 4B [\[3\]](#); Figure4_CombinedData_Tables_Fig4_Fig4S1_Fig1S3; Data Table 4). Next, we ranked all the H3K4me3 marked genes by their intensity (at ± 50 kb with highest intensity ranked first, filtered by adjusted p-value from differential gene expression by FDR-adjusted p-value ≤ 0.05) with color representing the log-fold-change in the gene expression. Only protein coding HOXA/HOXB genes are considered to compute statistics, while all the other genes in the ranked set is grouped under “other genes”. Data for these genes with H3K4me3 intensity in 2000 and 100 kbp window-size and corresponding change in gene expression (Data as in Figure 4C [\[4\]](#); Figure4_CombinedData_Tables_Fig4_Fig4S1_Fig1S3; Data Table 5 and 6 for each patient).

Figure supplements

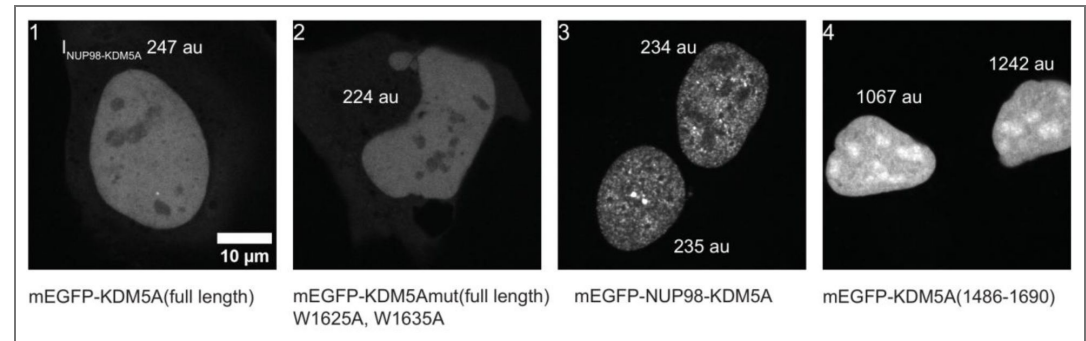


Figure 1 - Figure Supplement 1. Live imaging of U2OS cell nuclei expressing NUP98-KDM5A-related constructs. 1. Nucleus expressing full-length KDM5A construct fluorescently tagged on the N-terminus. 2. Chromatin-binding-deficient version of the construct in (1). 3. Condensates formed by NUP98-KDM5A tagged with mEGFP on the N-terminus. 4. Nuclei expressing only the chromatin-binding C-terminal fragment of the NUP98-KDM5A fusion, derived from KDM5A (1486–1690). 1-3 show cells with comparable expression levels, whereas 4 shows two cells with high expression but no chromatin-associated puncta formation. All images are a single z-slice acquired on a standard confocal microscope.

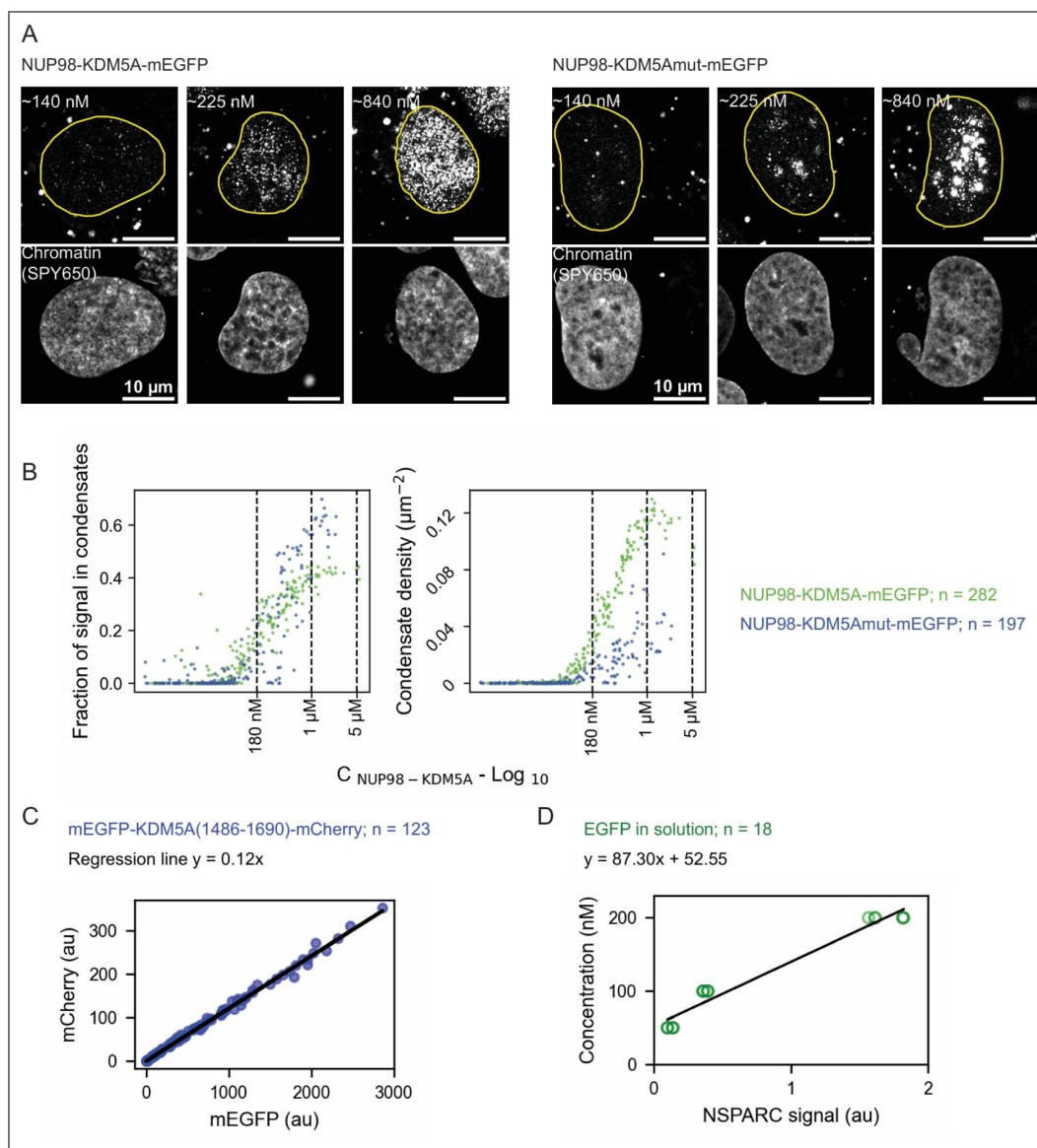


Figure 1 - Figure Supplement 2. NSPARC enhanced-resolution live imaging of NUP98-KDM5A condensates in U2OS cells.

A. Concentration-dependent formation of (Left) NUP98-KDM5A-mEGFP and (Right) NUP98-KDM5A-mut-mEGFP condensates, showing single z-slices from 3D stacks. **B.** Quantification of NUP98-KDM5A-mEGFP and NUP98-KDM5A-mut-mEGFP condensate formation across expression levels for the fraction of signal in condensates and the number density of condensates. **C.** Calibration between mEGFP and mCherry fluorescence intensities under the NSPARC microscopy settings used for live-cell imaging (see Materials and Methods). **D.** Calibration between EGFP fluorescence intensity and EGFP molar concentration in solution under the NSPARC microscopy settings used for live-cell imaging (see Materials and Methods).

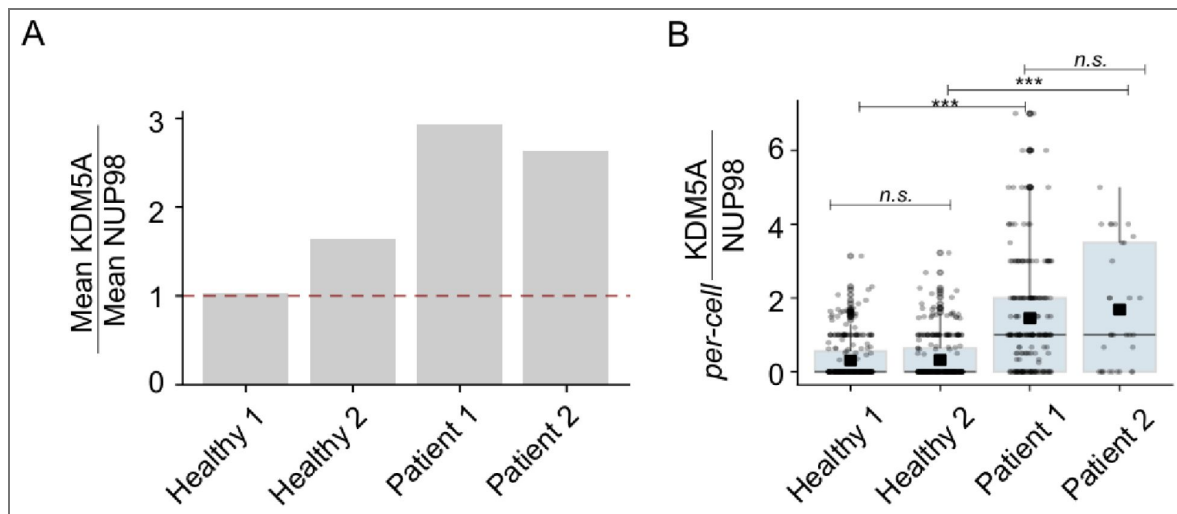


Figure 1 - Figure Supplement 3. Estimation of the relative abundance of KDM5A to NUP98 in patient and healthy donors.

A. Group level relative abundance of KDM5A: NUP98 transcripts quantified as the ratio of the mean expression across cells in each group. **B.** Distribution of per-cell KDM5A: NUP98 ratios for cells with detectable NUP98 expression (NUP98 > 0). Squares represent 'mean'. Wilcoxon rank sum test, ns: not significant, *p<0.05, ***p< 0.001.

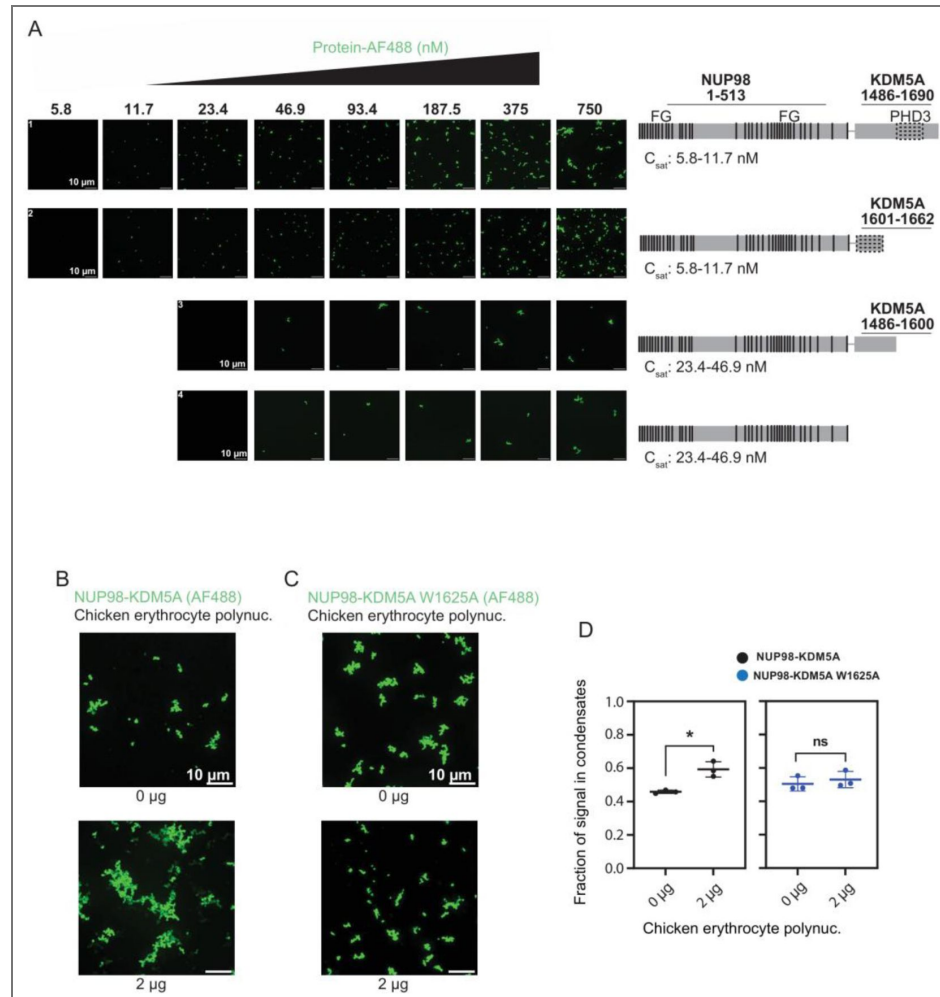


Figure 2 - Figure Supplement 1. *In vitro* reconstitution of NUP98-KDM5A condensates.

A. Spinning disk confocal images of condensates formed *in vitro* with increasing concentration of the following AF488 labelled constructs: NUP98 (1-513)-KDM5A (1486-1690), the full length oncofusion; NUP98 (1-513)-KDM5A (1601-1662), which contains only the PHD3 domain of the KDM5A fusion partner; NUP98 (1-513)-KDM5A (1486-1690), where the C-terminal truncation of the KDM5A fusion partner removes the PHD3 domain; NUP98 (1-513), the NUP98 segment alone. **B.** *In vitro* phase separation assay of AF488 labelled NUP98-KDM5A (750 nM) with chicken erythrocyte polynucleosomes (0 or 2 μ g). **C.** *In vitro* phase separation assay of AF488 labelled NUP98-KDM5A W1625A (750 nM) with chicken erythrocyte polynucleosomes (0 or 2 μ g). All images in **A-C** show maximum intensity projections of Z-stacks (scale bar = 10 μ m; n = 3). **D.** Fraction of signal in condensates for reactions with 750 nM AF488 labelled NUP98-KDM5A with 0 or 2 μ g chicken erythrocyte polynucleosomes (left panel) or 750 nM AF488 labelled NUP98-KDM5A W1625A with 0 or 2 μ g chicken erythrocyte polynucleosomes (right panel). The fraction of signal in condensates corresponds to the fluorescence intensity measured inside condensates relative to the total image intensity, reflecting the proportion of molecules partitioned into condensates. A parametric unpaired t-test was carried out to compare the average for the fraction of signal in condensates (n=3; ns, $P > 0.05$; *, $P \leq 0.05$; **, $P \leq 0.01$).

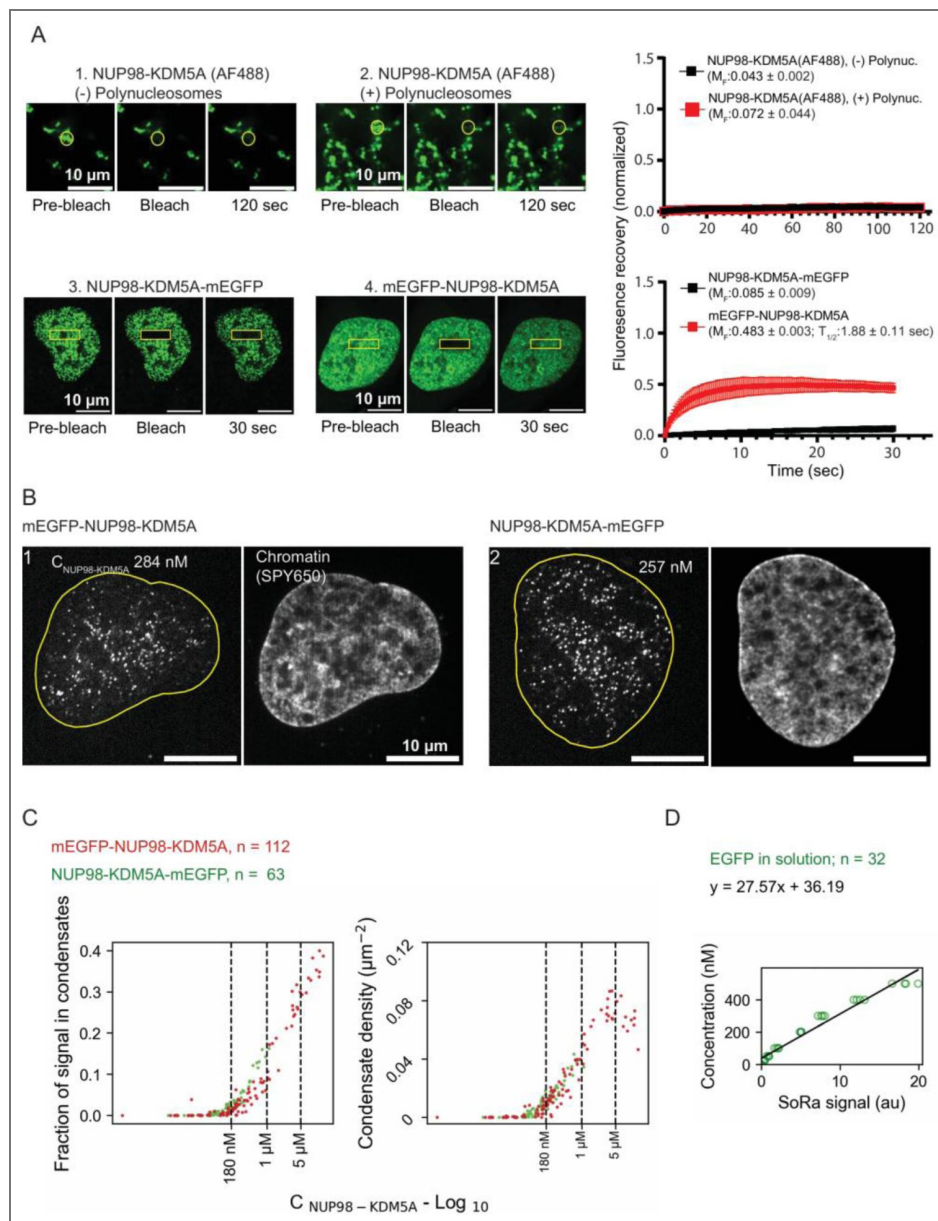


Figure 2 - Figure Supplement 2. Dynamics of NUP98-KDM5A condensates.

A. Representative images of the fluorescence recovery (inside yellow circle) of AF488 labelled NUP98-KDM5A (1 μ M) mixed with 0 or 2 μ g chicken erythrocyte polynucleosomes following photobleaching. Fluorescence recovery (inside yellow box) of U2OS cells expressing NUP98-KDM5A constructs tagged with mEGFP on either the C- or N-terminus after photobleaching. Normalized fluorescence recovery curves of photobleached regions (inside yellow marks) are plotted. Error bars are standard deviations of the means. **B.** SoRa-enhanced resolution live imaging of U2OS cells expressing NUP98-KDM5A constructs with an N- or C-terminal mEGFP tag. **C.** Quantification of mEGFP-NUP98-KDM5A and NUP98-KDM5A-mEGFP condensate formation across expression levels ($C_{\text{NUP98-KDM5A}} - \text{Log}_{10}$) for the fraction of signal in condensates and condensate density (μm^{-2}). **D.** Calibration between EGFP fluorescence intensity and EGFP molar concentration in solution under the SoRa microscopy settings used for live-cell imaging (see Materials and Methods).

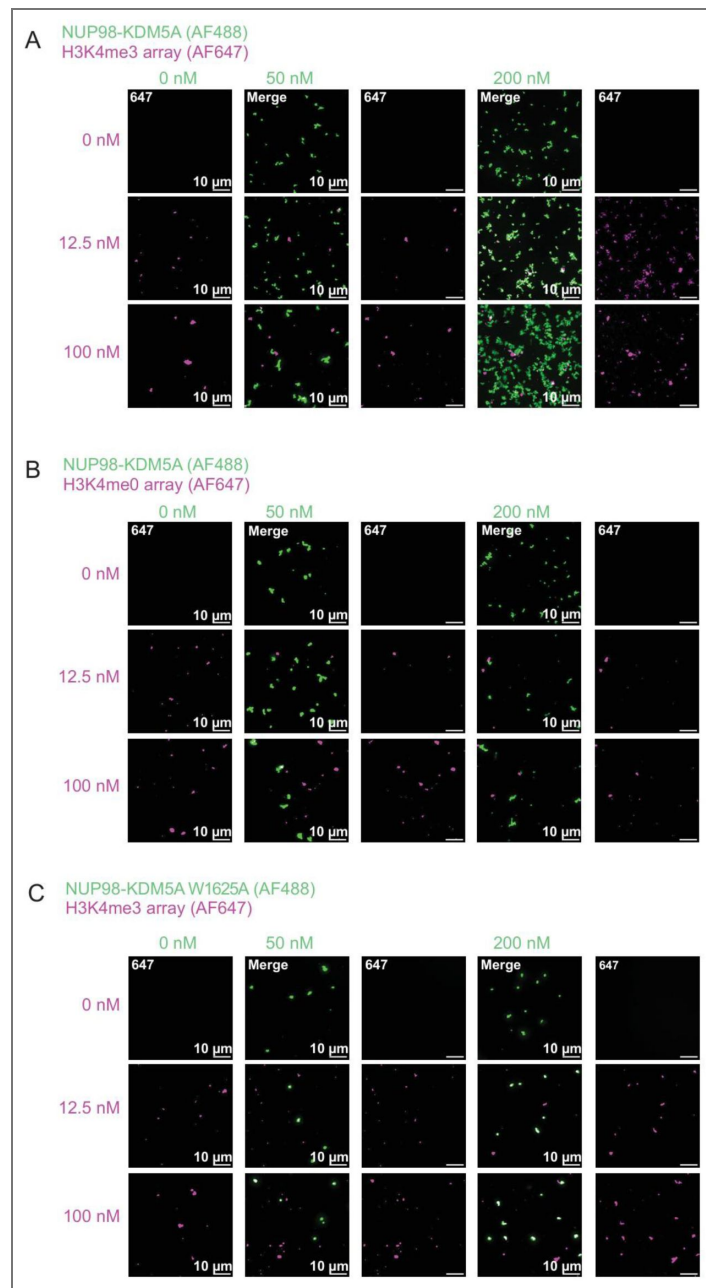


Figure 2 - Figure Supplement 3. Representative images of *in vitro* NUP98-KDM5A and chromatin co-phase separation assays.

A. AF488 labelled NUP98-KDM5A (0, 50 or 200 nM) with AF647 H3K4me3 nucleosomal arrays (0, 12.5 or 100 nM nucleosomes).
B. AF488 labelled NUP98-KDM5A (0, 50 or 200 nM) with AF647 H3K4me0 nucleosomal arrays (0, 12.5 or 100 nM nucleosomes).
C. AF488 labelled NUP98-KDM5A W1625A (0, 50 or 200 nM) with AF647 H3K4me3 nucleosomal arrays (0, 12.5 or 100 nM nucleosomes). Micrographs are presented as maximum intensity projections (scale bar = 10 μ m; n = 3).

Figure 3 - Figure Supplement 1. Negative control for the local signal analysis.

A. Intensity around the FISH-labeled loci (HOXA, HOXB, HOXD, TNF, MUC4, and MYC) on the DAPI channel across a range of NUP98-KDM5A-FLAG expression levels in HEK293FT cells. **B.** Boxplots showing the DAPI channel intensity surrounding FISH-labeled loci (HOXA, HOXB, HOXD, TNF, MUC4, and MYC) across binned NUP98-KDM5A-FLAG expression levels in HEK293FT cells.

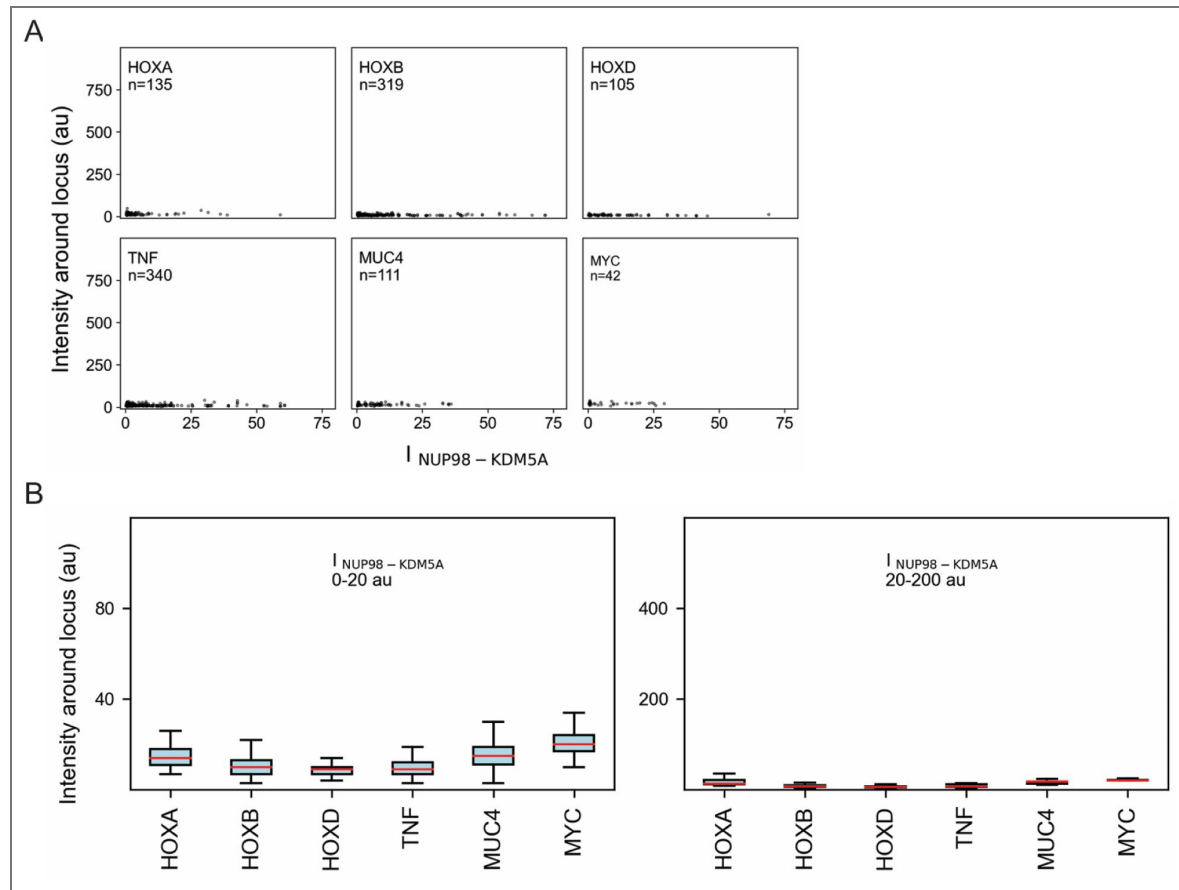
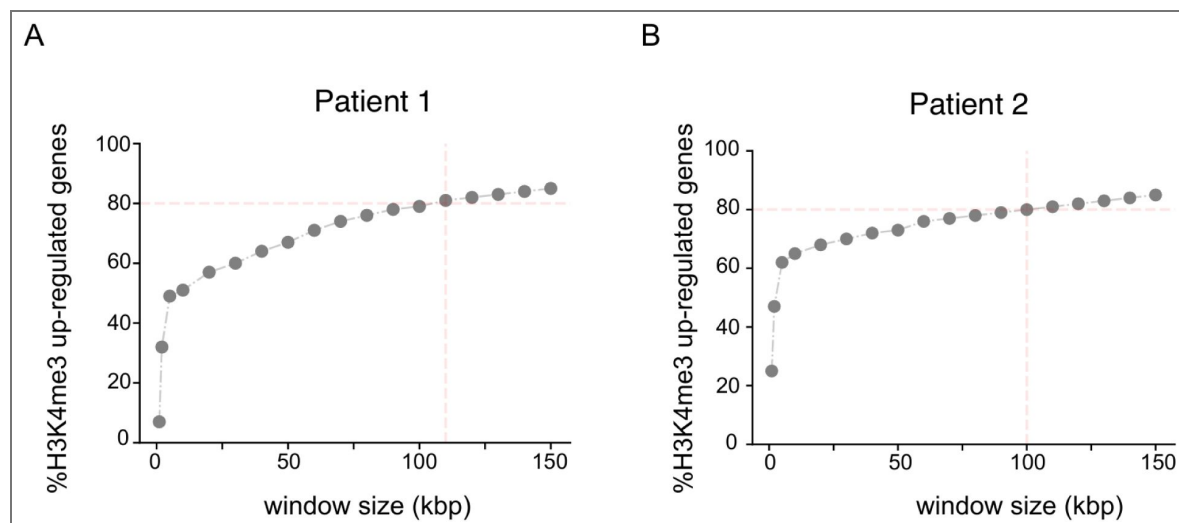


Figure 4 - Figure Supplement 1. Selection of window size for H3K4me3 quantification.

Window size selection for the quantification of gene-specific H3K4me3 intensity.



Data availability

Supplemental data files are available in the Dryad repository "Preferential formation of NUP98-KDM5A condensates at specific H3K4me3-rich loci drives leukemogenic gene expression - Data and Code" (DOI: 10.5061/dryad.m63xsj4hb). Specific file paths are indicated throughout the Materials and Methods section of the manuscript.

Acknowledgements

Data for this study were acquired at the Center of Advanced Light Microscopy at UCSF on a W1-CSU Confocal obtained using NIH S10 Shared Instrumentation grant (1S10OD017993-01A1), on CREST/C2 Confocal obtained using grants from the UCSF Program for Breakthrough Biomedical Research funded in part by the Sandler Foundation and the UCSF Research Resource Fund Award, and at the UCSF Innovation Core at the Weill Institute for Neurosciences on a Yokagawa CSU-W1 SoRA Spinning Disk Confocal Microscope.

Data for this study were acquired in the Nachury laboratory at UCSF using a Nikon Ti2-E scanning laser microscope equipped with a Nikon Spatial Array Confocal (NSPARC) detector. We thank Maxence Nachury and members of the Nachury lab for providing access to their Nikon NSPARC system.

We thank members of the Huang, Fujimori, and Li laboratories at UCSF, particularly Daniel Foyt, Nitya Kopparapu, Yiming Kuang, Catherine Taylor, and Aivy David, for helpful discussions and support. We also thank Alistair Boettiger and members of the Boettiger laboratory at Stanford University for guidance in establishing our FISH protocol, in particular Kavvya Gupta and Derek Le. We further acknowledge the faculty of the Cold Spring Harbor Laboratory course "*Quantitative Imaging: From Acquisition to Analysis*" for insightful advice on imaging analysis.

B.H. and G.N. are supported by the NIH 4D Nucleome Consortium (U01DK127421). J.E.S. is supported by NIH Institutional Research and Academic Career Development Award (IRACDA) K12GM081266 through the UCSF IRACDA Scholars Program. D.G.F. is supported by the Bowes Biomedical Research Award. N.K. and H. L. are supported by NIH (R01AG083524, R33AG078793). B.H. and H.L. are Biohub San Francisco Investigators.

Additional information

Author Contributions

AB - Conceptualization, Data curation, Software, Formal analysis, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing, Project administration

JES - Conceptualization, Data curation, Software, Formal analysis, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing, Project administration

NK - Data curation, Software, Formal analysis, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing

AM - Data curation, Software, Formal analysis, Visualization, Investigation, Methodology

TW - Formal analysis, Investigation, Methodology

CM - Formal analysis, Investigation, Methodology, Writing – review and editing

HL – Resources, Supervision, Funding acquisition, Writing – review and editing

GJN – Resources, Supervision, Funding acquisition, Writing – review and editing

DGF – Conceptualization, Resources, Supervision, Funding acquisition, Methodology, Writing – original draft, Writing – review and editing, Project administration

BH - Conceptualization, Resources, Supervision, Funding acquisition, Methodology, Writing – original draft, Writing – review and editing, Project administration

Funding

Funder	Grant reference number	Author
HHS National Institutes of Health (NIH)	U01DK127421	Geeta J Narlikar Bo Huang
HHS National Institutes of Health (NIH)	K12GM081266	Jonathan E Sandoval
HHS National Institutes of Health (NIH)	R01AG083524	Hao Li Neha Khetan
HHS National Institutes of Health (NIH)	R33AG078793	Neha Khetan Hao Li
Biohub San Francisco	Investigator Program	Hao Li Bo Huang
University of California, San Francisco (UCSF)	Bowes Biomedical Research Award	Danica G Fujimori

Author ORCID iDs

Geeta J Narlikar:  <https://orcid.org/0000-0002-1920-0147>

Bo Huang:  <https://orcid.org/0000-0003-1704-4141>

References

Ackermann BE, Debelouchina GT (2019) Heterochromatin Protein HP1 α Gelation Dynamics Revealed by Solid-State NMR Spectroscopy. *Angewandte Chemie International Edition* **58**:6300-6305
<https://doi.org/10.1002/anie.201901141> | PubMed

Ahn JH, Davis ES, Daugird TA, Zhao S, Quiroga IY, Uryu H, Li J, Storey AJ, Tsai Y.-H, Keeley DP, *et al.* (2021) Phase separation drives aberrant chromatin looping and cancer development. *Nature* **595**:591-595
<https://doi.org/10.1038/s41586-021-03662-5> | PubMed

Ahn JH, Guo Y, Lyons H, Mackintosh SG, Lau BK, Edmondson RD, Byrum SD, Storey AJ, Tackett AJ, Cai L, *et al.* (2025) The phenylalanine-and-glycine repeats of NUP98 oncofusions form condensates that selectively partition transcriptional coactivators. *Molecular Cell* **85**:708-725.e9.
<https://doi.org/10.1016/j.molcel.2024.12.026> | PubMed

Atlasi Y, Stunnenberg HG (2017) The interplay of epigenetic marks during stem cell differentiation and development. *Nature Reviews Genetics* **18**:643-658 <https://doi.org/10.1038/nrg.2017.57> | PubMed

Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, Karra K, Wong ED, Devarajan K, Daugherty AC, Kundaje AB, Mancini E, *et al.* (2014) H3K4me3 Breadth Is Linked to Cell Identity and Transcriptional Consistency. *Cell* **158**:673-688 <https://doi.org/10.1016/j.cell.2014.06.027> | PubMed

Bertrums EJM, Smith JL, Harmon L, Ries RE, Wang Y.-CJ, Alonzo TA, Menssen AJ, Chisholm KM, Leonti AR, Tarlock K, *et al.* (2023) Comprehensive molecular and clinical characterization of NUP98 fusions in pediatric acute myeloid leukemia. *Haematologica* **108**:2044-2058
<https://doi.org/10.3324/haematol.2022.281653> | PubMed

Blackledge NP, Klose RJ (2021) The molecular principles of gene regulation by Polycomb repressive complexes. *Nature Reviews Molecular Cell Biology* **22**:815-833 <https://doi.org/10.1038/s41580-021-00398-y> | PubMed

Brumbaugh-Reed EH, Gao Y, Aoki K, Toettcher JE (2024) Rapid and reversible dissolution of biomolecular condensates using light-controlled recruitment of a solubility tag. *Nature Communications* **15**:6717 <https://doi.org/10.1038/s41467-024-50858-0> | PubMed

Calvo KR, Sykes DB, Pasillas MP, Kamps MP (2002) Nup98-HoxA9 immortalizes myeloid progenitors, enforces expression of Hoxa9, Hoxa7 and Meis1, and alters cytokine-specific responses in a manner similar to that induced by retroviral co-expression of Hoxa9 and Meis1. *Oncogene* **21**:4247-4256 <https://doi.org/10.1038/sj.onc.1205516> | PubMed

Chandra B, Michmerhuizen NL, Shirnekhi HK, Tripathi S, Pioso BJ, Baggett DW, Mitrea DM, Iacobucci I, White MR, Chen J, *et al.* (2022) Phase Separation Mediates NUP98 Fusion Oncoprotein Leukemic Transformation. *Cancer Discovery* **12**:1152-1169 <https://doi.org/10.1158/2159-8290.CD-21-0674> | PubMed

Cho NH, Cheveralls KC, Brunner A.-D, Kim K, Michaelis AC, Raghavan P, Kobayashi H, Savy L, Li JY, Canaj H, *et al.* (2022) OpenCell: Endogenous tagging for the cartography of human cellular organization. *Science* **375**:eabi6983 <https://doi.org/10.1126/science.abi6983> | PubMed

Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, Heckert A, Banala S, Lavis L, Darzacq X, *et al.* (2018) Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**:eaar2555 <https://doi.org/10.1126/science.aar2555> | PubMed

Chong S, Graham TGW, Dugast-Darzacq C, Dailey GM, Darzacq X, Tjian R (2022) Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. *Molecular Cell* **82**:2084-2097.e5. <https://doi.org/10.1016/j.molcel.2022.04.007> | PubMed

D. Edelstein A, A. Tsuchida M, Amodaj N, Pinkard H, D. Vale R, Stuurman N (2014) Advanced methods of microscope control using µManager software. *Journal of Biological Methods* **1**:1 <https://doi.org/10.14440/jbm.2014.36> | PubMed

Deshpande P, Prentice E, Vidal Ceballos A, Casaccia P, Elbaum-Garfinkle S (2024) Epigenetic marks uniquely tune the material properties of HP1α condensates. *Biophysical Journal* **123**:1508-1518 <https://doi.org/10.1016/j.bpj.2024.04.020> | PubMed

Fatti E, Khawaja S, Weis K (2025) The dark side of fluorescent protein tagging—The impact of protein tags on biomolecular condensation. *Molecular Biology of the Cell* **36**:br10 <https://doi.org/10.1091/mbc.E24-11-0521> | PubMed

Fonseca JP, Bonny AR, Kumar GR, Ng AH, Town J, Wu QC, Aslankoohi E, Chen SY, Dods G, Harrigan P, *et al.* (2019) A Toolkit for Rapid Modular Construction of Biological Circuits in Mammalian Cells. *ACS Synthetic Biology* **8**:2593-2606 <https://doi.org/10.1021/acssynbio.9b00322> | PubMed

Frey S, Rees R, Schünemann J, Ng SC, Fünfgeld K, Huyton T, Görlich D (2018) Surface Properties Determining Passage Rates of Proteins through Nuclear Pores. *Cell* **174**:202-217.e9. <https://doi.org/10.1016/j.cell.2018.05.045> | PubMed

Gough SM, Lee F, Yang F, Walker RL, Zhu YJ, Pineda M, Onozawa M, Chung YJ, Bilke S, Wagner EK, *et al.* (2014) NUP98-PHF23 Is a Chromatin-Modifying Oncoprotein That Causes a Wide Array of Leukemias Sensitive to Inhibition of PHD Histone Reader Function. *Cancer Discovery* **4**:564-577 <https://doi.org/10.1158/2159-8290.CD-13-0419> | PubMed

Gough SM, Slape CI, Aplan PD (2011) NUP98 gene fusions and hematopoietic malignancies: Common themes and new biologic insights. *Blood* **118**:6247-6257 <https://doi.org/10.1182/blood-2011-07-328880> | PubMed

Hamamoto K, Zhu G, Lai Q, Lesperance J, Luo H, Li Y, Nigam N, Sharma A, Yang F.-C, Claxton D, *et al.* (2025) HoxB1nc lncRNA reprograms CTCF-independent TADs to drive leukemic transcription and HSC dysregulation in NUP98-rearranged leukemia. *Journal of Clinical Investigation* **135**:e184743 <https://doi.org/10.1172/JCI184743> | PubMed

Heikamp EB, Henrich JA, Perner F, Wong EM, Hatton C, Wen Y, Barwe SP, Gopalakrishnapillai A, Xu H, Uckelmann HJ, *et al.* (2022) The menin-MLL1 interaction is a molecular dependency in NUP98 - rearranged AML. *Blood* **139**:894-906 <https://doi.org/10.1182/blood.2021012806> | PubMed

Heikamp EB, Martucci C, Henrich JA, Neel DS, Mahendra-Rajah S, Rice H, Wenge DV, Perner F, Wen Y, Hatton C, *et al.* (2024) NUP98 fusion proteins and KMT2A-MENIN antagonize PRC1.1 to drive gene expression in AML. *Cell Reports* **43**:114901 <https://doi.org/10.1016/j.celrep.2024.114901> | PubMed

- Hulsmann BB, Labokha AA, Gorlich D (2012) The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell* **150**:738-751 <https://doi.org/10.1016/j.cell.2012.07.019> | PubMed
- Ibáñez De Opakua A, Geraets JA, Frieg B, Dienemann C, Savastano A, Rankovic M, Cima-Omori M.-S, Schröder GF, Zweckstetter M (2022) Molecular interactions of FG nucleoporin repeats at high resolution. *Nature Chemistry* **14**:1278-1285 <https://doi.org/10.1038/s41557-022-01035-7> | PubMed
- Jevtic Z, Matafora V, Casagrande F, Santoro F, Minucci S, Garre' M, Rasouli M, Heidenreich O, Musco G, Schwaller J, et al. (2022) SMARCA5 interacts with NUP98-NSD1 oncofusion protein and sustains hematopoietic cells transformation. *Journal of Experimental & Clinical Cancer Research* **41**:34 <https://doi.org/10.1186/s13046-022-02248-x> | PubMed
- Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S, Narlikar GJ (2017) Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **547**:236-240 <https://doi.org/10.1038/nature22822> | PubMed
- Li C, Wen A, Shen B, Lu J, Huang Y, Chang Y (2011) FastCloning: A highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnology* **11**:92 <https://doi.org/10.1186/1472-6750-11-92> | PubMed
- Li H, Côté P, Kuoch M, Ezike J, Frenis K, Afanassiev A, Greenstreet L, Tanaka-Yano M, Tarantino G, Zhang S, et al. (2025) The dynamics of hematopoiesis over the human lifespan. *Nature Methods* **22**:422-434 <https://doi.org/10.1038/s41592-024-02495-0> | PubMed
- Luger K, Rechsteiner TJ, Richmond TJ (1999) Expression and Purification of Recombinant Histones and Nucleosome Reconstitution. In: Becker PB (Ed). *Chromatin Protocols* **119** Humana Press. pp. 1-16 <https://doi.org/10.1385/1-59259-681-9:1> | PubMed
- Lukauskas S, Tvardovskiy A, Nguyen NV, Stadler M, Faull P, Ravnsborg T, Özdemir Aygenli B, Dornauer S, Flynn H, Lindeboom RGH, et al. (2024) Decoding chromatin states by proteomic profiling of nucleosome readers. *Nature* **627**:671-679 <https://doi.org/10.1038/s41586-024-07141-5> | PubMed
- Mateo LJ, Murphy SE, Hafner A, Cinquini IS, Walker CA, Boettiger AN (2019) Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**:49-54 <https://doi.org/10.1038/s41586-019-1035-4> | PubMed
- Meissner A (2010) Epigenetic modifications in pluripotent and differentiated cells. *Nature Biotechnology* **28**:1079-1088 <https://doi.org/10.1038/nbt.1684> | PubMed
- Mestrom L, Marsden SR, Dieters M, Achterberg P, Stolk L, Bento I, Hanefeld U, Hagedoorn P.-L (2019) Artificial Fusion of mCherry Enhances Trehalose Transferase Solubility and Stability. *Applied and Environmental Microbiology* **85**:e03084-18 <https://doi.org/10.1128/AEM.03084-18> | PubMed
- Michmerhuizen NL, Heikamp EB, Iacobucci I, Umeda M, Arthur B, Mishra V, Park CS, Di Giacomo D, Hiltenbrand R, Gao Q, et al. (2025) KAT6A and KAT7 Histone Acetyltransferase Complexes Are Molecular Dependencies and Therapeutic Targets in NUP98-Rearranged Acute Myeloid Leukemia. *Cancer Discovery* **15**:2096-2116 <https://doi.org/10.1158/2159-8290.CD-24-1772> | PubMed
- Michmerhuizen NL, Klco JM, Mullighan CG (2020) Mechanistic insights and potential therapeutic approaches for NUP98-rearranged hematologic malignancies. *Blood* **136**:2275-2289 <https://doi.org/10.1182/blood.2020007093> | PubMed
- Moore C, Wong E, Kaur U, Chio US, Zhou Z, Ostrowski M, Wu K, Irklyenko I, Wang S, Ramani V, et al. (2025) ATP-dependent remodeling of chromatin condensates reveals distinct mesoscale outcomes. *Science* **390**:eadr0018 <https://doi.org/10.1126/science.adr0018> | PubMed
- Noort S, Wander P, Alonzo TA, Smith J, Ries RE, Gerbing RB, Dolman MEM, Locatelli F, Reinhardt D, Baruchel A, et al. (2021) The clinical and biological characteristics of NUP98-KDM5A in pediatric acute myeloid leukemia. *Haematologica* **106**:630-634 <https://doi.org/10.3324/haematol.2019.236745> | PubMed
- Oka M, Otani M, Miyamoto Y, Oshima R, Adachi J, Tomonaga T, Asally M, Nagaoka Y, Tanaka K, Toyoda A, et al. (2023) Phase-separated nuclear bodies of nucleoporin fusions promote condensation of MLL1/CRM1 and rearrangement of 3D genome structure. *Cell Reports* **42**:112884

<https://doi.org/10.1016/j.celrep.2023.112884> | PubMed

Pachitariu M, Rariden M, Stringer C (2025) Cellpose-SAM: Superhuman generalization for cellular segmentation. *bioRxiv* <https://doi.org/10.1101/2025.04.28.651001>

Pachitariu M, Stringer C (2022) Cellpose 2.0: How to train your own model. *Nature Methods* **19**:1634-1641 <https://doi.org/10.1038/s41592-022-01663-4> | PubMed

Passaro M, Martinovic M, Bevilacqua V, Hershberg EA, Rossetti G, Beliveau BJ, Bonnal RJP, Pagani M (2020) OligoMinerApp: A web-server application for the design of genome-scale oligonucleotide in situ hybridization probes through the flexible OligoMiner environment. *Nucleic Acids Research* **48**:W332-W339 <https://doi.org/10.1093/nar/gkaa251> | PubMed

Perez G, Barber GP, Benet-Pages A, Casper J, Clawson H, Diekhans M, Fischer C, Gonzalez JN, Hinrichs AS, Lee CM, *et al.* (2025) The UCSC Genome Browser database: 2025 update. *Nucleic Acids Research* **53**:D1243-D1249 <https://doi.org/10.1093/nar/gkae974> | PubMed

Plys AJ, Davis CP, Kim J, Rizki G, Keenen MM, Marr SK, Kingston RE (2019) Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes & Development* **33**:799-813 <https://doi.org/10.1101/gad.326488.119> | PubMed

Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, *et al.* (2018) Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**:eaar3958 <https://doi.org/10.1126/science.aar3958> | PubMed

Sabo PJ, Hawrylycz M, Wallace JC, Humbert R, Yu M, Shafer A, Kawamoto J, Hall R, Mack J, Dorschner MO, *et al.* (2004) Discovery of functional noncoding elements by digital analysis of chromatin structure. *Proceedings of the National Academy of Sciences* **101**:16837-16842 <https://doi.org/10.1073/pnas.0407387101> | PubMed

Sanulli S, Trnka MJ, Dharmarajan V, Tibble RW, Pascal BD, Burlingame AL, Griffin PR, Gross JD, Narlikar GJ (2019) HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature* **575**:390-394 <https://doi.org/10.1038/s41586-019-1669-2> | PubMed

Schmidt HB, Görlich D (2015) Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity. *eLife* **4**:e04251 <https://doi.org/10.7554/eLife.04251> | PubMed

Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH (2017) Phase separation drives heterochromatin domain formation. *Nature* **547**:241-245 <https://doi.org/10.1038/nature22989> | PubMed

Strom AR, Kim Y, Zhao H, Chang Y.-C, Orlovsky ND, Košmrlj A, Storm C, Brangwynne CP (2024) Condensate interfacial forces reposition DNA loci and probe chromatin viscoelasticity. *Cell* **187**:5282-5297.e20. <https://doi.org/10.1016/j.cell.2024.07.034> | PubMed

Tatavosian R, Kent S, Brown K, Yao T, Duc HN, Huynh TN, Zhen CY, Ma B, Wang H, Ren X (2019) Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. *Journal of Biological Chemistry* **294**:1451-1463 <https://doi.org/10.1074/jbc.RA118.006620> | PubMed

Terlecki-Zaniewicz S, Humer T, Eder T, Schmoellerl J, Heyes E, Manhart G, Kuchynka N, Parapatics K, Liberante FG, Müller AC, *et al.* (2021) Biomolecular condensation of NUP98 fusion proteins drives leukemogenic gene expression. *Nature Structural & Molecular Biology* **28**:190-201 <https://doi.org/10.1038/s41594-020-00550-w> | PubMed

Tian J, Zhu Y, Li J, Yang G, Weng X, Huang T, Zhao L, Sun H, Yan Z, Zhang S (2024) The landscape of NUP98 rearrangements clinical characteristics and treatment response from 1491 acute leukemia patients. *Blood Cancer Journal* **14**:81 <https://doi.org/10.1038/s41408-024-01066-y> | PubMed

Tripathi S, Shirnekhi HK, Gorman SD, Chandra B, Baggett DW, Park C.-G, Somjee R, Lang B, Hosseini SMH, Pioso BJ, *et al.* (2023) Defining the condensate landscape of fusion oncoproteins. *Nature Communications* **14**:6008 <https://doi.org/10.1038/s41467-023-41655-2> | PubMed

- Troester S**, Eder T, Wukowits N, Piontek M, Fernández-Pernas P, Schmoellerl J, Haladik B, Manhart G, Allram M, Maurer-Granofszky M, *et al.* (2025) Transcriptional and epigenetic rewiring by the NUP98::KDM5A fusion oncoprotein directly activates CDK12. *Nature Communications* **16**:4656 <https://doi.org/10.1038/s41467-025-59930-9> | PubMed
- Umeda M**, Hiltenbrand R, Michmerhuizen NL, Barajas JM, Thomas Iii ME, Arthur B, Walsh MP, Song G, Ma J, Westover T, *et al.* (2025) Fusion oncoproteins and cooperating mutations define disease phenotypes in NUP98 - rearranged leukemia. *Blood* **146**:2102-2118 <https://doi.org/10.1182/blood.2025028993> | PubMed
- Van Zutven LJCM**, Önen E, Velthuisen SCJM, Van Drunen E, Von Bergh ARM, Van Den Heuvel-Eibrink MM, Veronese A, Mecucci C, Negrini M, De Greef GE, *et al.* (2006) Identification of NUP98 abnormalities in acute leukemia: JARID1A (12p13) as a new partner gene. *Genes. Chromosomes and Cancer* **45**:437-446 <https://doi.org/10.1002/gcc.20308> | PubMed
- Wang GG**, Cai L, Pasillas MP, Kamps MP (2007) NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nature Cell Biology* **9**:804-812 <https://doi.org/10.1038/ncb1608> | PubMed
- Wang GG**, Song J, Wang Z, Dormann HL, Casadio F, Li H, Luo J.-L, Patel DJ, Allis CD (2009) Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature* **459**:847-851 <https://doi.org/10.1038/nature08036> | PubMed
- Wang L**, Gao Y, Zheng X, Liu C, Dong S, Li R, Zhang G, Wei Y, Qu H, Li Y, *et al.* (2019) Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism. *Molecular Cell* **76**:646-659.e6. <https://doi.org/10.1016/j.molcel.2019.08.019> | PubMed
- Yun M**, Wu J, Workman JL, Li B (2011) Readers of histone modifications. *Cell Research* **21**:564-578 <https://doi.org/10.1038/cr.2011.42> | PubMed
- Zhang MY**, Yang H, Ortiz G, Trnka MJ, Petronikolou N, Burlingame AL, DeGrado WF, Fujimori DG (2022) Covalent labeling of a chromatin reader domain using proximity-reactive cyclic peptides. *Chemical Science* **13**:6599-6609 <https://doi.org/10.1039/D2SC00555G> | PubMed
- Zhang Y**, Brown K, Yu Y, Ibrahim Z, Zandian M, Xuan H, Ingersoll S, Lee T, Ebmeier CC, Liu J, *et al.* (2021) Nuclear condensates of p300 formed through the structured catalytic core can act as a storage pool of p300 with reduced HAT activity. *Nature Communications* **12**:4618 <https://doi.org/10.1038/s41467-021-24950-8> | PubMed
- Zhen CY**, Duc HN, Kokotovic M, Phiel CJ, Ren X (2014) Cbx2 stably associates with mitotic chromosomes via a PRC2- or PRC1-independent mechanism and is needed for recruiting PRC1 complex to mitotic chromosomes. *Molecular Biology of the Cell* **25**:3726-3739 <https://doi.org/10.1091/mbc.e14-06-1109> | PubMed
- Zhou Z**, Hong K, Huang B, Narlikar GJ (2025) Understanding how genetically encoded tags and crowding agents affect phase separation by heterochromatin protein HP1a. *Cell Reports Methods* **5**:101029 <https://doi.org/10.1016/j.crmeth.2025.101029> | PubMed
- Umeda M**, Kiko J (2025) Fusion oncoproteins and cooperating mutations define disease phenotypes in NUP98-rearranged leukemia. NCBI Gene Expression Omnibus. ID GSE287716 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE287716>
- Li H**, Butty VL, Kuoch M, Cote P, Chou ST, Rowe RG (2024) Hematopoiesis at single cell resolution spanning human development and maturation. NCBI Gene Expression Omnibus. ID GSE189161 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE189161>

Peer reviews

Reviewer #1 (Public review):

Leukemia-driving NUP98 oncofusion proteins form chromatin-associated biomolecular condensates in the nucleus, and these structures are important for oncogenic transformation.

Most NUP98 fusions do not contain domains that mediate the recognition of specific DNA elements. Instead, they entail domains that are important for chromatin regulation. For instance, the NUP98::KDM5A fusion features a fusion of the NUP98 N-terminus with the third PHD domain of the histone demethylase KDM5A. As PHD domains are critical for the recognition of methylated histones without any sequence specificity, it is not clear what controls the condensation and chromatin binding of NUP98::KDM5A, leading to the induction of oncogenic transcriptional programs.

In this work, the authors use a combination of cellular and in vitro studies to show that biomolecular condensation of NUP98::KDM5A is dependent on H3K4me3 binding. Their model proposes that concentration-dependent chromatin-associated condensation of NUP98::KDM5A depends on local densities of H3K4me3 on chromatin and the levels of the fusion oncoprotein. In line with this, the analysis of gene expression data from NUP98::KDM5A-positive AML cells shows a positive correlation between differentially expressed genes and H3K4me3 levels.

This is an interesting manuscript that aims to dissect the molecular mechanisms underlying biomolecular condensation of the NUP98::KDM5A oncoprotein. The work is solid, and the results are well explained and presented in a logical order. However, the study suffers from several weaknesses that if addressed would improve the study.

Major points:

(1) All cellular experiments are performed in settings of transient transfection of NUP98::KDM5A in non-hematopoietic cell types. These conditions are not physiologically relevant, as these cells do not depend on the fusion oncogene. Therefore, any claims about concentration-dependent effects on condensation need to be validated in AML cells that are driven by NUP98::KDM5A. While this may not be possible in primary patient-derived cells, several groups have published AML models of NUP98::KDM5A-driven AML that could be used.

(2) The results presented in Figure 4 are not entirely supportive of the mechanism. It is known that active gene expression correlates with high H3K4me3 levels; therefore, the correlations shown by the authors are expected. Yet, the authors do not discuss the fact that many H3K4me3-positive genomic regions do not show NUP98::KDM5A binding. This should be elaborated on in the discussion section.

(3) While the focus of the manuscript is on NUP98::KDM5A, this oncofusion is part of a family of >30 fusions that join the NUP98 N-terminus to a variety of factors with roles in epigenetic control and transcription. While the repertoire of NUP98 fusion partners is diverse with regard to functional domains, they all induce a conserved set of target genes that is characteristic of this leukemia subtype. How can this be achieved in the context of NUP98 fusion proteins that do not contain a PHD domain, such as NUP98::NSD1 or NUP98::HOXA9? Please discuss this.

<https://doi.org/10.7554/eLife.111922.1.sa2>

Reviewer #2 (Public review):

In this manuscript, the authors investigate how the oncogenic fusion protein NUP98-KDM5A alters gene expression in leukemia, using a combination of cellular experiments with model and patient cell lines, as well as in vitro studies. Upon transfection of U2OS cells with mEGFP-tagged NUP98-KDM5A, the authors show that the fusion proteins form sub-micrometer puncta, whereas KDM5A alone does not. These foci are also observed at expected native expression levels (using OpenCell data). The tag has an effect here, as switching to an

mCherry tag raises the apparent saturation concentration for phase separation. Finally, the authors show via super-resolution imaging that the foci correlate with H3K4me3 distribution.

In vitro, the fusion protein forms amorphous, gel-like condensates at double-digit nanomolar concentrations. Truncation analysis identifies PHD3 of KDM5A as required for maximal phase separation, consistent with the ability of the protein to bind H3K4me3 peptides. Addition of polynucleosomes increases the amount of fusion protein partitioning into the condensate in an H3K4me3-binding-dependent manner. Condensates are gel-like with slow internal dynamics in vitro; in cells, however, the dynamics depend on the position of the EGFP tag (no corresponding experiments with mCherry are shown). Reconstitution with H3K4me3- and H3K4me0-modified arrays shows colocalization with both wild-type NUP98-KDM5A and the binding mutant. Here, H3K4me3 arrays recruit ~20% more protein and yield gel-like structures in a manner dependent on the PTM and on the PHD finger.

In cells, the fusion protein colocalizes with H3K4me3-marked loci, including the HOX clusters, as confirmed by FISH. Finally, re-analysis of published expression datasets from patient cells shows that genes are predominantly upregulated and that the upregulated genes are H3K4me3-marked.

This is a well-executed mechanistic study. The data convincingly establish that NUP98-KDM5A forms sub-micrometer foci at realistic expression levels, that these foci correlate with H3K4me3-marked sites, that the PHD3-H3K4me3 interaction mediates chromatin binding while the NUP98 moiety drives phase separation in vitro, that foci in cells overlap genes heavily decorated with H3K4me3, and that H3K4me3-marked genes are those found to be upregulated in patient datasets. These are important mechanistic findings and of interest to the community.

Still, the functional/causal link is a bit more tentative, as the data is mostly correlative, since it is not directly established that there is feedback between H3K4 methylation, NUP98-KDM5A recruitment, phase separation and target gene overexpression. An experiment that could further bolster this claim would be a direct test of whether NUP98-KDM5A expression drives overexpression of bound genes, e.g. expression of the fusion protein vs PHD- and NUP98-mutant variants, followed by qPCR of target genes, such as the HOX cluster, and possibly H3K4me3 ChIP at the same loci. As all the constructs and cell lines exist, this could be feasible and would substantially strengthen the manuscript.

<https://doi.org/10.7554/eLife.111922.1.sa1>

Author response:

We greatly appreciate the positive and constructive comments from the reviewers, which recognized the intellectual contributions of our manuscript and also captured its limitations. Below is our response to the three major points from Reviewer #1 and the comments from Reviewer #2.

Response to Reviewer #1

(1) Use of transient transfection in non-hematopoietic cells. We appreciate the reviewer's point regarding physiological relevance. Our goal in these cellular microscopy experiments was to dissect the biophysical principles of NUP98::KDM5A (including its mutants) condensate formation under controlled expression levels (concentration). While AML model systems driven by NUP98::KDM5A are available, they do not offer this possibility because of pre-existing NUP98::KDM5A expression. The suspension culture of hematopoietic cells also brings practical challenges for correlative FISH+IF and high-resolution microscopy analysis. We agree that validating our observed behaviors in AML models would be valuable, e.g. by creating HSPC lines with inducible expression of tagged NUP98::KDM5A and its mutants, but

such experiments fall outside the scope of the current study. We will add text acknowledging this limitation and clarifying that our mechanistic conclusions are grounded in biophysical principles that should generalize across cell types.

(2) Correlation between H3K4me3 and gene activation. We agree that active transcription correlates with H3K4me3, and that this baseline relationship must be considered. Our analysis explicitly uses fold-change between patient cells and healthy controls as the readout. This comparison inherently accounts for the activating effect of H3K4me3 itself. Regarding the reviewer's comment that "many H3K4me3-positive genomic regions do not show NUP98::KDM5A binding", we would like to clarify that this point is exactly what our manuscript aims to explain. Our cell line studies demonstrate that, at a patient-relevant expression level, NUP98::KDM5A condensates form preferentially at H3K4me3 locus with high local mark density. Considering that NUP98::KDM5A concentration in the nucleus is lower than the K_D between KDM5A PHD3 and H3K4me3, this means that H3K4me3 loci with lower mark density will not see NUP98::KDM5A binding without the high local concentration of the fusion protein (as a result of condensate formation). This is consistent with our observation that genes with the highest local density show disproportionately stronger upregulation. We will further clarify this point in the revised manuscript.

(3) Generalization to other NUP98 fusions lacking PHD domains. We appreciate this important conceptual question. We will expand the discussion to note that many NUP98 fusions, despite diverse partner domains, produce similar transcriptional programs. Our current hypothesis is that the highly active status of the HOX cluster genes in HSPC attracts NUP98 oncofusions targeting H3K4me3 (e.g. KDM5A and PHF23), while NUP98::HOXA9 and other transcription factor fusions directly target the HOX cluster via DNA sequence recognition. This convergence and the broad targets of the dysregulated HOX transcription factors explain the transcriptional program similarity, sustained by the previously described enrichment of transcriptional co-activators by NUP98 oncofusion condensates. We will explicitly discuss this hypothesis in our revised manuscript.

Response to Reviewer #2

We thank the reviewer for the thoughtful evaluation and agree that the causal link between H3K4me3 recognition, condensate formation, and gene activation remains partly correlative. Experiments such as qPCR or ChIP following expression of WT versus mutant constructs would indeed strengthen the causal chain. However, performing these assays across multiple constructs and loci in a physiologically relevant system is not feasible within the current revision cycle. We have added text acknowledging this limitation and clarifying that our study focuses on establishing the biophysical mechanism of targeting, while functional consequences are inferred from patient datasets rather than new perturbation experiments.

<https://doi.org/10.7554/eLife.111922.1.sa0>