Phase separation drives aberrant chromatin looping and cancer development

Jeong Hyun Ahn[1,2], **Eric S. Davis**[3], **Timothy A. Daugird**[4], **Shuai Zhao**[1,2], **Ivana Yoseli Quiroga**[5], **Hidetaka Uryu**[1,2], **Jie Li**[1,6], **Aaron J. Storey**[7], **Yi-Hsuan Tsai**[1], **Daniel P. Keeley**[8], **Samuel G. Mackintosh**[7], **Ricky D. Edmondson**[7], **Stephanie D. Byrum**[7], **Ling Cai**[1,2,9], **Alan J. Tackett**[7], **Deyou Zheng**[10], **Wesley R. Legant**[4,11], **Douglas H Phanstiel**[1,3,5,12,§], **Gang Greg Wang**[1,2,4,6,§]

[1]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC 27599, USA.

[2]Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC 27599, USA.

[3]Curriculum in Bioinformatics and Computational Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.

[4]Department of Pharmacology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC 27599, USA.

[5]Thurston Arthritis Research Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.

[6]Curriculum in Genetics and Molecular Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

[7]Department of Biochemistry and Molecular Biology, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA.

[8]UNC Neuroscience Center and Carolina Institute for Developmental Disabilities, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.

[9]Department of Genetics, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC 27599, USA.

Correspondence and requests for materials should be addressed to D.H.P. (douglas_phanstiel@med.unc.edu) or G.G.W. (greg_wang@med.unc.edu).
§These authors jointly supervised this work.

[10]Departments of Genetics, Neurology, and Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[11]Joint Department of Biomedical Engineering, University of North Carolina, Chapel Hill, and North Carolina State University, Raleigh, NC 27599, USA.

[12]Department of Cell Biology and Physiology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC 27599, USA.

Abstract

The development of cancer is intimately associated with genetic abnormalities that target proteins with intrinsically disordered regions (IDRs). In human hematological malignancies, recurrent chromosomal translocation of nucleoporin (NUP98 or NUP214) generates an aberrant chimera that invariably retains the nucleoporin's IDR, tandemly dispersed phenylalanine-and-glycine (FG) repeats[1,2]. However, it remains elusive how unstructured IDRs contribute to oncogenesis. We show that IDR harbored within NUP98-HOXA9, a homeodomain-containing transcription factor (TF) chimera recurrently detected in leukemias[1,2], is essential for establishing liquid-liquid phase separation (LLPS) puncta of chimera and for inducing leukemic transformation. Notably, LLPS of NUP98-HOXA9 not only promotes chromatin occupancy of chimera TFs but is also required for formation of a broad, 'super-enhancer'-like binding pattern, typically seen at a battery of leukemogenic genes, potentiating their transcriptional activation. An artificial HOX chimera, created by replacing NUP98's FG repeats with an unrelated LLPS-forming IDR of FUS[3,4], had similar enhancement effects on chimera's genome-wide binding and target gene activation. Chromosome conformation capture techniques such as Hi-C mapping further demonstrated that phase-separated NUP98-HOXA9 induces CTCF-independent chromatin looping enriched at proto-oncogenes. Together, this report describes a proof-of-principle example wherein cancer acquires mutation to establish oncogenic TF condensates via phase separation, which simultaneously enhances their genomic targeting and induces organization of aberrant three-dimensional chromatin structure during tumorous transformation. As LLPS-competent molecules are frequently implicated in diseases[1,2,4–7], this mechanism can potentially be generalized to many malignant and pathological settings.

IDRs within various proteins—including transcription factors (TFs), chromatin modulators and RNA-binding proteins—form liquid droplets via phase separation, which affects myriad biological processes ranging from organelle formation and stress tolerance to transcription[4,5,8–10]. Notably, many cancers are characterized by recurrent fusions between genes encoding IDR-containing and chromatin-binding proteins. For instance, a subset of leukemias displaying poor prognosis carry a characteristic chromosomal translocation that produces chimera combining an IDR-containing segment of nucleoporin with chromatin/DNA-binding factor[1,2,11,12]. Similarly, in Ewing's sarcoma, aberrant fusion occurs between TFs and the IDR of RNA-binding proteins[7]. Both chromatin-binding and IDR-containing domains were previously shown to be essential for tumorigenicity, supporting chromatin deregulation as a general mechanism[1,11,12]. However, it remains elusive how IDRs contribute to gene mis-regulation and oncogenesis.

IDRs induce chimeric TF phase separation

We aimed to define the role for IDR and potentially phase separation in tumorigenicity by characterizing NUP98-HOXA9, a fusion that shares similarity to other NUP98-TF chimeras identified from various leukemia subtypes[1,2]. NUP98-HOXA9 contains two structural identities from NUP98—dispersed FG repeats and a GLE2-binding sequence (GLEBS; Extended Data Fig.1a). GLEBS deletion did not interfere with NUP98-HOXA9-mediated transformation of primary hematopoietic stem/progenitor cells (HSPCs; Extended Data Fig.1b–c). Normally, NUP98 is mainly localized to the nuclear periphery. Live-cell imaging showed that the full-length and GLEBS-deleted NUP98-HOXA9 displayed a pattern of nucleoplasmic puncta (Extended Data Fig.1d). Immunoblotting showed the levels of NUP98 and NUP98-HOXA9 were comparable (Extended Data Fig.1d, right). Thus, NUP98-HOXA9-mediated HSPC transformation and condensate formation are GLEBS-independent. To dissect the role for NUP98's IDR in leukemogenesis, we mainly used GLEBS-deleted NUP98-HOXA9 (Fig.1a–b; hereafter referred to as N-IDR$_{WT}$/A9).

To determine whether N-IDR$_{WT}$/A9 puncta are established via LLPS, we employed several approaches[8–10,13]. First, we observed that N-IDR$_{WT}$/A9 puncta were sensitive to 1,6-hexanediol, a chemical used to disrupt phase-separated condensates[8–10,13] (Fig.1c). Second, the purified NUP98 IDR (N-IDR) proteins formed liquid condensates *in vitro* (Fig.1d, 38×FG). To further assess concentration dependency and importance of multivalency conferred by FG-repeats for condensate formation, we generated recombinant N-IDR proteins containing a varying number of FG-repeats (Extended Data Fig.1e–f). While N-IDR harboring 38× or 36×FG-repeats formed liquid droplets in a concentration-dependent fashion (Fig.1d), those with 27× or 11×FG-repeats failed to phase separate under same conditions (not shown). Only with assistance of a crowding agent and at higher concentrations was the 27×FG-repeat-containing N-IDR able to establish condensates *in vitro* (Fig.1d). However, when mixed with N-IDR proteins harboring 38×FG-repeats, those with 11× or 27×FG-repeats were readily incorporated into formed condensates *in vitro* (Fig.1e). Imaging of cells expressing N-IDR/A9 with the varying FG-repeat number corroborated *in vitro* findings— compared to chimeras with 38× or 36×FG-repeats, those with less FG-repeats formed less condensates in cells (27×) or could not at all (11×), resembling what was seen with the HOXA9 fusion segment alone (Extended Data Fig.1g and Fig.1f). Additionally, DNA binding is dispensable for forming LLPS-like NUP98-HOXA9 puncta. In fact, relative to N-IDR$_{WT}$/A9 puncta, those formed by its DNA-binding-defective form (carrying an N51S homeodomain mutation[14,15]) were significantly fewer in total number and much larger in size (Extended Data Fig.1a,d,h); condensates of the N51S mutants were readily detected as droplet-like nuclear structures even under the phase-contrast microscope (Fig. 1g). This indicates that chromatin binding of NUP98-HOXA9 may spatially restrict condensates from further coalescence, which occurs more readily with the N51S-mutant puncta. Condensates of NUP98-HOXA9^{N51S} were also 1,6-hexanediol-sensitive (Extended Data Fig.1h). Notably, live-cell imaging post-induction of GFP-NUP98-HOXA9^{N51S} showed events of coalescence in which multiple small condensates collided producing a larger one (Fig.1h; Supplementary Video 1), which is a characteristic of

liquid condensates[13]. Together, IDR within NUP98-HOXA9 establishes LLPS in a valency-dependent and concentration-dependent manner.

IDRs in chimeric TFs drive oncogenesis

To investigate the roles for IDR and LLPS in leukemogenesis, we mutated phenylalanine in the FG-repeats of chimera to serine (Fig.1a), a mutation previously shown to disable hydrogel formation by FG-repeats *in vitro*[16]. Such F→S mutations did not affect the protein stability but significantly abolished the nucleoplasmic droplet formation by N-IDR$_{WT}$/A9 carrying either wildtype (WT) or N51S-mutated homeodomain, supporting a critical requirement of FG-repeats for LLPS in cells (Fig.1b–c and Extended Data Fig.1h,2a–b). NUP98-HOXA9 was reported to interact, either directly or indirectly, with coactivators such as CBP/p300[17] and MLL-NSL complexes[18]. We next queried if F→S mutations perturbed such interaction networks by employing BioID and found a majority of N-IDR$_{WT}$/A9- and N-IDR$_{FS}$/A9-interacting proteins to be shared, including all reported interactors and many general transcriptional machinery proteins (Extended Data Fig. 2c and Supplementary Table 1). To further examine the relationship between IDR-mediated LLPS and leukemogenesis, we performed the retrovirus-mediated oncogene transduction and transformation assays with murine HSPCs and found that, unlike N-IDR$_{WT}$/A9 that efficiently formed nuclear condensates and displayed a potent HSPC-transforming capacity as described[19], the F→S mutant failed to establish puncta in HSPCs, failed to transform HSPCs *in vitro*, and failed to induce leukemia *in vivo* (Fig.1i–k; Extended Data Fig.2d–g). We further assessed involvement of IDR and LLPS in leukemogenesis with an artificial chimera termed F-IDR$_{WT}$/A9 by fusing HOXA9's homeodomain to an unrelated IDR of FUS that can phase separate[20,21] (Fig.1a–b). As expected, F-IDR$_{WT}$/A9 formed puncta in cells, a process suppressed by 1,6-hexanediol treatment or condensate-disrupting mutation[20] (F-IDR$_{YS}$/A9; Fig.1a–b and Extended data Fig.2h). Consistent to NUP98-HOXA9, only the IDR-intact and not Y→S mutant form of F-IDR/A9 caused leukemic transformation *in vitro* and *in vivo* (Fig.1i–j and Extended data Fig.2g,i). Altogether, LLPS-forming IDRs retained within chimeric TFs are essential for cancerous transformation.

IDRs enhance genomic binding of TF chimera

NUP98-HOXA9 binds DNA via homeodomain, causing gene deregulation during leukemogenesis. Next, we assessed the impact of IDR-mediated phase separation on chromatin targeting of NUP98-HOXA9 by chromatin immunoprecipitation-sequencing (ChIP-seq) to map genome-wide binding of LLPS-competent N-IDR$_{WT}$/A9 versus LLPS-deficient N-IDR$_{FS}$/A9 in their corresponding stable expression cells. Here, 293 cell provides a system for assessing direct gene-regulatory effects of NUP98-HOXA9, because its cellular state is relatively stable and not apparently altered post-transduction of chimera, in contrast to what was observed in HSPCs such as differentiation arrest[18,19] (Extended data Fig.2e–f). ChIP-seq using antibodies of different tags attached to N-IDR/A9 produced robust, highly correlated signals whereas ChIP-seq with non-tagged cells generated almost no binding (Extended Data Fig.3a–c). Both N-IDR$_{WT}$/A9 and N-IDR$_{FS}$/A9 displayed preferential binding to intergenic and intronic enhancers, with binding most enriched in expected motifs of HOX-related TFs (Extended data Fig.3d–g). Despite shared features

seen for their targeting, N-IDR$_{WT}$/A9 displayed a strikingly enhanced genomic occupancy relative to N-IDR$_{FS}$/A9, irrespective of peak subclasses defined by unsupervised clustering (Fig.2a). Also, the broad and dense 'super-enhancer'-like peaks are unique to N-IDR$_{WT}$/A9 (Supplementary Table 2) and enriched at development- and leukemia-associated genes (Extended Data Fig.3h), exemplified by HOX, PBX3 and MEIS1 (Fig.2b–c and Extended Data Fig.4a–e). Super-enhancer calling by N-IDR$_{WT}$/A9 or H3K27ac verified their dense binding at proto-oncogenes (Extended Data Fig.5a–c).

To further assess the role for IDR-induced LLPS in chimera's chromatin targeting, we employed several additional strategies. First, treatment of 1,6-hexanediol drastically decreased chromatin occupancy of N-IDR$_{WT}$/A9 whereas it had minimal effect on overall binding of N-IDR$_{FS}$/A9 (Fig.2d and Extended Data Fig.5d–e). 1,6-hexanediol treatment also suppressed formation of a vast majority of broad N-IDR$_{WT}$/A9 peaks (Extended Data Fig.5f; Supplementary Table 2). As a result, overall binding of N-IDR$_{WT}$/A9 post-treatment of 1,6-hexanediol more closely resembled that of LLPS-incompetent N-IDR$_{FS}$/A9, compared to N-IDR$_{WT}$/A9 without treatment (Extended Data Fig.5g). Second, we turned to F-IDR/A9 and tested whether FUS's IDR is sufficient in enhancing chimera's genomic binding. ChIP-seq revealed that these two chimeras carrying unrelated LLPS-competent IDRs showed similar binding patterns—F-IDR$_{WT}$/A9 shows significantly enhanced genomic targeting and broad binding at AML-related oncogenes, in contrast to F-IDR$_{YS}$/A9 (Fig.3a–b; Extended Data Fig.4,6a–b; Supplementary Table 3). ChIP-seq for N-IDR$_{WT}$/A9 in murine leukemias uncovered the similar 'super-enhancer'-like peaks at oncogenes, which overlapped those found in 293 cells (Extended Data Fig.6c–e). ChIP-qPCR verified the dramatically enhanced enrichment of N-IDR$_{WT}$/A9 and F-IDR$_{WT}$/A9, relative to their corresponding IDR mutant, and suppressive effect by 1,6-hexanediol on binding of N-IDR$_{WT}$/A9, but not its LLPS-defective mutant, to the tested loci (Extended Data Fig.6f–g). Thirdly, we used cells expressing NUP98-HOXA9 with varied numbers of FG-repeats, which were either LLPS-competent or LLPS-incompetent, and ChIP-qPCR detected significantly enhanced enrichment of LLPS-competent and not LLPS-incompetent fusions at loci showing broad N-IDR$_{WT}$/A9 binding (Fig.3c), indicating a critical FG-repeat number required for establishing LLPS and intensified binding of chimera. Lastly, we conducted single-molecule imaging studies to evaluate chromatin occupancy of N-IDR$_{WT}$/A9 relative to N-IDR$_{FS}$/A9. Measurements of single-molecule speed and track displacement showed N-IDR$_{WT}$/A9 to be significantly less mobile than N-IDR$_{FS}$/A9 (Extended Data Fig.7). Two-state kinetic modeling of single-molecule trajectories[22] revealed that, compared to N-IDR$_{FS}$/A9, N-IDR$_{WT}$/A9 had a greater fraction of molecules in the low-diffusion bound state and had slower diffusion coefficients (Fig.3d and Extended Data Fig.7f–g), which suggests that TF assemblies, confined within phase-separated puncta, engage target DNA sequences more tightly and generally display slower diffusion, compared to LLPS-defective TFs. Collectively, using genetic and pharmacological approaches, we demonstrated a causal role for IDR-mediated LLPS in establishing enhanced targeting of chimeric TFs, particularly those seen at super-enhancer-like peaks.

IDRs in TFs potentiate target activation

To assess relationship between NUP98-HOXA9 binding and gene activation, we conducted H3K27ac ChIP-seq and observed that increased chimera TF binding is correlated to increased H3K27ac (Fig.2a–c and Extended Data Fig.4). Immunofluorescence also revealed co-localization of N-IDR$_{WT}$/A9 'dots' with H3K27ac, in comparison to H3K9me3 (Extended Data Fig.8a–b). To further define the role for IDR in target gene regulation, we performed RNA-seq in 293 cells with stable chimera expression and identified 303 of differentially expressed genes (DEGs) significantly up-regulated by N-IDR$_{WT}$/A9, compared to mock and N-IDR$_{FS}$/A9 (Fig.3e and Supplementary Table 4), the effect confirmed by RT-qPCR (Extended Data Fig.8c). IDR-dependent gene activation was also observed in 293 cells with expression of F-IDR$_{WT}$/A9 versus F-IDR$_{YS}$/A9 (Extended Data Fig.8d; Supplementary Table 5), albeit gene activation of F-IDR$_{WT}$/A9 is less than that of N-IDR$_{WT}$/A9 (Fig.3f), in agreement with a relatively less oncogenic potency by the former *in vivo* (Fig.1i–j). Additionally, RNA-seq of murine HSPCs transduced with fusion relative to mock corroborated that N-IDR$_{WT}$/A9, but not N-IDR$_{FS}$/A9, sustains oncogenic gene-expression programs, which again include Hox, Meis and Pbx family genes and other signatures related to leukemia and HSPCs (Fig.3g and Extended Data Fig.8e–f; Supplementary Table 6); as expected, differentiation-related genesets were suppressed in the N-IDR$_{WT}$/A9 sample (Extended Data Fig.8f, bottom). Gene-regulatory effects of the artificial chimera F-IDR$_{WT}$/A9 were found to be similar to those of N-IDR$_{WT}$/A9 in HSPCs (Extended Data Fig.8g and Supplemental Table 7). Furthermore, a reduction in the FG-repeat number, which decreased LLPS competence, also significantly decreased effects of the chimera on oncogene transcription and HSPC transformation (Fig.3h–i). Thus, genomic profiling of independent models lends strong support for a critical role of IDRs in activating proto-oncogenes, many of which carry 'super-enhancer'-like elements bound by chimeric TFs and H3K27ac.

IDRs and LLPS induce chromatin looping

Increasing evidence suggests that phase separation of chromatin-associated factor can promote three-dimensional chromatin structure to modulate transcription[9,23–26]. However, so far there is little direct evidence that phase separation can form point-to-point DNA loops similar to those created by CTCF and cohesion, nor that such phase separation-driven loops have a causal role in human disease. To test the ability of NUP98-HOXA9 to form chromatin loops via LLPS, we generated Hi-C profiles of 293 cells with either N-IDR$_{WT}$/A9 or N-IDR$_{FS}$/A9, revealing 6,615 DNA loops (Fig.4a) and high correlation between replicates (Extended Data Fig.9a–b). To determine the effect of N-IDR$_{WT}$/A9 on Hi-C contact frequency, we aggregated interaction counts between the 500 most strongly N-IDR$_{WT}$/A9-occupied sites for both N-IDR$_{WT}$/A9- and N-IDR$_{FS}$/A9-expressing cells. Regions with high occupancy of N-IDR$_{WT}$/A9 exhibited elevated interactions frequencies, even between binding sites separated by great distances (>2Mb) or on different chromosomes entirely (Fig.4b). Elevated interaction frequencies were not observed between the same loci in cells expressing N-IDR$_{FS}$/A9 (Fig.4b). Differential analysis revealed 232 loops specific to N-IDR$_{WT}$/A9 and 52 specific to N-IDR$_{FS}$/A9 (DESeq2, P<0.01; Fig.4a,c–e). The majority (91%) of N-IDR$_{WT}$/A9-specific-loop anchors overlapped N-IDR$_{WT}$/A9 binding, while much

fewer (31%) overlapped a CTCF binding site (Fig.4f). Thus, N-IDR$_{WT}$/A9 loops form in a largely CTCF-independent manner, consistent with a phase separation-driven mechanism. 3C-qPCR post-treatment of 1,6-hexanediol showed that N-IDR$_{WT}$/A9-specific loop at *PBX3*, but not an unrelated CTCF loop, was significantly disrupted (Extended Data Fig. 9d–g). The vast majority (82%) of N-IDR$_{WT}$/A9-specific-loop anchors overlapped H3K27ac, in contrast to only 31% observed for non-differential loop anchors (Fig. 4f), which suggests that N-IDR$_{WT}$/A9-specific loops rewire connections between enhancers and target genes. Indeed, genes whose promoters overlapped N-IDR$_{WT}$/A9-specific-loop anchors were highly expressed, further supporting a regulatory role of these loops (Fig. 4g). The up-regulated genes at N-IDR$_{WT}$/A9-specific-loop anchors again include proto-oncogenes such as HOX and PBX3 (Fig. 4d,g and Extended Data Fig. 10a–c). These results support that IDRs of chimeric TFs induces DNA looping between 'super-enhancer-like' targeting sites and oncogenes via phase separation.

Discussion

In summary, we show a critical requirement of LLPS-competent IDR harbored within NUP98-HOXA9 oncoproteins for leukemogenesis and for activation of the oncogenic gene-expression program via its effects on (i) enhancing chimeric TF binding to genomic targets and/or (ii) promoting long-distance, enhancer-promoter looping at oncogenes (Extended Data Fig.11). We demonstrated these effects by both genetic (IDR mutagenesis or replacement with an unrelated one and changing the FG-repeats valency) and pharmacological methodologies, which provide a proof-of-principle example wherein cancer acquires mutation to establish oncogenic TF condensates for their efficient targeting onto binding sites and reorganization of 3D chromatin structure during tumorous transformation. As a range of IDR-containing LLPS-competent molecules are implicated in diseases[1,2,4–7], this mechanism can potentially be generalized to many pathological settings.

Online content methods, additional references, Nature Research reporting summaries, extended data, source data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available online at xxx.

Methods

Plasmid Construction.

The MSCV-based retroviral vector for expression of NUP98-HOXA9 fusion was previously described[27] and the mammalian expression construct containing various tagged NUP98-HOXA9 (such as GFP-NUP98-HOXA9 in an inducible expression vector[28]) was kind gifts of M. Kamps, B. Fahrenkrog and J. Schwaller. For generating various chimera constructs of N-IDR/A9 or F-IDR/A9 fusions, we synthesized the gBlocks (IDT) that contain cDNA segments of both fusion partners fused in-frame, with a 3×HA-3×FLAG tag added at the C-terminus. Each gBlock fragment was cloned into the MSCV retroviral vector with a drug selection marker (Puro or Neo). For live-cell imaging studies, we replaced the 3×HA-3×FLAG tag in fusion constructs with EGFP by subcloning. For generating a series of constructs with a varying number of NUP98 FG repeats, we used the following NUP98

portion as its fusion segment in the expression vector: aa 1–468 as 38× FG repeats, aa 1–468(132–224) as 36× FG repeats, aa 65–468 (132–224) as 27× FG repeats and aa 357–468 as 11×FG repeats. For bacterial expression of IDR, the same fragments with varying number of FG repeats were cloned into the pRSFDuet-1 vector (a kind gift of Dr J. Song). For single-molecule tracking studies, we synthesized gBlocks (IDT) that contain cDNA segments of a HaloTag with flanking enzymatic sites of MluI and XhoI to replace the 3×HA-3×FLAG tag described in the above expression vectors. All plasmids used were confirmed by sequencing before use and are listed in Supplementary Table 8.

Tissue Culture and Stable Cell Line Generation.

HEK293T (ATCC #CRL-3216) and HeLa (ATCC #CCL-2) cells were obtained from ATCC and maintained using recommended culture conditions. Authentication of cell identities, including those parental and derived lines, was ensured by the Tissue Culture Facility affiliated to UNC Lineberger Comprehensive Cancer Center with genetic signature profiling and fingerprinting analyses[29]. A routine examination for any possible mycoplasma contamination was performed every month with kits (Lonza). Cells in a passage of less than of 10 times were used in this study. Retrovirus or lentivirus was packaged and produced in 293 cells, and the stable cell lines were generated by viral infection followed by drug selection as performed before[30,31]. The 293 cell lines with stable expression of chimera with either WT or mutant IDRs were first examined by western blotting and immunofluorescence of the transgene, and the same sets of cells then used throughout this study for various assays such as live-cell imaging and genomic profiling (RNA-seq, ChIP-seq and Hi-C).

Antibodies and Western Blotting.

Immunoblotting was performed as previously described[30,31]. Affinity-purified antibodies against endogenous NUP98 (raised in rabbits against NUP98 amino acids 51 to 223 covering GLEBS) was a kind gift of JM van Deursen and used as described before[17,32]. The information of antibodies used in this study is listed in Supplementary Table 8.

Fixed Cell Immunofluorescence

293 Cells were grown on polylysine-coated coverslips (Corning, #354085) for 24 hrs at a 37°C incubator. For non-adherent mouse HPSCs, 0.1 million cells were added on top of polylysine-coated coverslips and spun down on centrifuge for 30min at 3000rpm. The cover slips were briefly washed with PBS and then fixed in 4% formaldehyde (Thermo Scientific #28908) for 10 minutes at room temperature. Fixed cell samples were washed with cold PBS three times and incubated in PBS plus 0.1% Triton X-100 for 10 min, followed by washing with PBS for three times and incubation in blocking buffer (1% BSA in PBS plus 0.1% Tween-20) for 30 min. After discarding the blocking buffer, the fixed samples were incubated with a primary antibody diluted in the blocking buffer for 2 hrs at room temperature or overnight at 4 °C in a humidified chamber, and then washed with PBS plus 0.1% Tween20 for three times (3 min each time). Lastly, the samples were incubated with the secondary antibody-conjugated to appropriate fluorophores for 2 hrs at room temperature and washed three times with PBST before adding the mounting medium (Thermo Scientific, #P36935). The slides were then dried overnight at dark before imaging on the Olympus FV1000 confocal microscope with a 100x/1.4NA Plan Apochromat oil immersion objective.

DAPI was imaged with an excitation of 405nm and emission from 430–470nm, Alexa Fluor 488 was imaged with an excitation of 488nm and emission from 505–540nm, and Alexa Fluor 594 was imaged with an excitation of 559nm and emission from 575–675nm.

Live Cell Imaging

For live cell imaging, cells were grown on 35-mm dish with 20mm glass bottom well (Cellvis, D35-20-1.5-N) for 24hrs prior to imaging. Live cell imaging was conducted on Olympus FV1000 confocal microscope using 60× and 100× oil objectives. To capture the events of coalescence where multiple small liquid condensates of chimera are fused into a single one, we used Hela cells with stable expression of doxycycline-inducible GFP-tagged NUP98-HOXA9^{N51S} for live-cell imaging upon chimera expression induction.

Chemical Treatment.

To test sensitivity of protein aggregates to 1,6-hexanediol treatment, 10% of 1,6-hexanediol (Sigma-Aldrich, #240117) were prepared in PBS. Throughout this study, the 1,6-hexanediol treatment condition was 10% for 1 min. Such 6-hexanediol treated cells, together with the vehicle-treated ones as control, were used for various experiments such as immediate imaging or fixation with 1% formaldehyde for a subsequent ChIP-seq experiment.

Recombinant Protein Purification

For bacterial expression of IDR proteins, the His6x tag-containing pRSFDuet-1 vector that contains NUP98 segment covering FG repeats was transformed into Rosetta 2™ (DE3) competent cells (Sigma, #71397). Three liters of bacterial cultures were grown at 37 °C for 12 hrs and then added with a final concentration of 0.5mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for overnight induction at 16 °C. Bacterial cells were spun down at 6000 rpm for 15 min, resuspended and lysed in 6M guanidine hydrochloride added with 20mM imidazole. After brief sonication, lysates were subject to centrifugation at 16,000rpm for 1hr at 4 °C, and supernatants were collected. Supernatants were run through Ni-column (Qiagen, #30250) and washed sequentially with the following buffers: 2M guanidine hydrochloride with 20mM imidazole, 2M guanidine hydrochloride with 1M NaCl, and 2M guanidine hydrochloride with 20mM imidazole. The His6x-tagged target proteins were eluted in 2M guanidium hydrochloride with 500mM imidazole, with 50ul of elution assessed by SDS-PAGE after ethanol precipitation. Then, protein samples were further purified on size exclusion column 10/300 SD75 (GE healthcare) using the AKTA purifier (GE healthcare, AKTA™ pure 25) in SEC buffer (2M guanidine hydrochloride). Fractions with purified target proteins were combined and concentrated using microcon-10 filter (Millipore, #MRCPRT010) to reach the sample concentration ranging from 27uM to 255uM and kept at −80 °C for storage.

In Vitro Phase Separation Assay

We first carried out the labeling of recombinant protein with the Alexa Flour 488 and 594 protein labeling kit (ThermoFisher, #A30006 and #A3008) according to manufacturer's protocols. To set up the *in vitro* phase separation assays, the labeled proteins were mixed with unlabeled ones at a ratio of 1:20, and such a mixture further diluted to a desired

concentration in the Eppendorf tubes with either TBS buffer alone (50mM Tris-HCl pH 7.5, 150mM NaCl) or TBS plus a crowding agent such as 20% of polyethylene glycol (PEG) 3350 (ThermoFisher, #NC0620958). Imaging was carried out immediately with samples transferred to a 35-mm dish with 20mm glass bottom well (Cellvis, D35-20-1.5-N) using Olympus FV3000RS Confocal microscope with 100× oil objective. For fluorescence imaging studies with a mixture of two species of N-IDR recombinant proteins containing FG-repeats in different numbers, we used those with 38× FG-repeats in the final concentration of 2.5 μM in the TBS buffer (labeled with Alexa Flour 488), which was mixed with those labeled with Alexa Flour 594, either carrying 27× FG-repeats (a final concentration of 2.5 μM) or 11× FG repeats (a final concentration of 6 μM).

Colocalization analysis

Colocalization analysis between fusion and H3K27ac or H3K9me3 was performed using the EzColocalization plugin in FIJI version 1.53[33]. Colocalization was measured using the Pearson's Correlation Coefficient (PCC). An a-priori power analysis of pilot data was performed in G*Power (Z-tests, two independent Pearson r's) and showed that a sample size of at least 388 cells would be required to determine significance at $p > 0.05$ given an effect size of 0.24. For analysis, nuclei were manually segmented by hand tracing with the polygon selection tool, then converted into binary masks used in the EzColocalization plugin to restrict colocalization analysis to the nuclei. PCC values for each cell were averaged and the calculated means were compared with an independent two-tailed Student's T-test.

Purification, Transduction, and Cultivation of Primary Murine Hematopoietic Stem/ Progenitor Cells (HSPCs)

Primary bone marrow cells are harvested from femur and tibia of Balb/C mice and then subject to a lineage-negative (Lin⁻) enrichment protocol to remove differentiated cell populations as described before[32,34]. Lin⁻ enriched HSPCs were first stimulated in the base medium (OptiMEM [Invitrogen, cat#31985]) supplemented with 15% of FBS (Invitrogen, cat#16000–044), 1% of antibiotics, 50 μM of β-mercaptoethanol and a cytokine cocktail that contains 10 ng/mL each of murine SCF (Peprotech), Flt3 ligand (Flt3L; Sigma), IL3 (Peprotech) and IL6 (Peprotech) for 4 days as described[11,32,34]. Two days post-infection with retrovirus, murine HSPCs were subject to drug selection and then plated for assaying proliferation and differentiation in the same liquid base medium with SCF alone as described before[11,32,34]. These *in vitro* cultured HSPC cells were routinely monitored under microscopy and cellular morphology examined by Wright-Giemsa staining as described[11,32,34]. For HSPCs transduced with a bicistronic GFP-containing retroviral construct, we also scored relative proliferation of GFP-positive HSPCs by FACS every 2–3 days post-infection.

Flow Cytometry (FACS) Analysis

Cells were washed once in the cold FACS buffer (PBS with 1% of FBS added) and then resuspended and incubated in the FACS buffer added with the respective antibodies (1:100 dilution) for 30 min on ice. The cell pellets were washed with FACS buffer and the stained cells were subject to analysis with the FACS machine (Attune Nxt, Thermo Fisher; available in UNC Flow Cytometry Core Facility). Data were analyzed using FlowJo software.

In Vivo Leukemogenic Assay

Determination of potential leukemogenic properties of the oncogene was carried out as described before[32,35]. In brief, 0.5 million of freshly infected and selected murine HSPCs were transplanted to syngeneic balb/C mice via tail vein injection (carried out by Animal Studies Core of UNC Cancer Center). Mice were regularly monitored with complete blood counting (CBC) with the collected peripheral blood and abdomen palpation for early signs of leukemia such as lethargy, increased white blood cell (WBC) counts and enlarged spleen[30]. Mice exhibiting leukemic phenotypes were euthanized followed by pathological and histological analyses as described[32,35]. Haematoxylin Eosin (H&E) staining of spleen sections was carried by UNC Pathology Core as described before[36].

BioID

A BirA cDNA sequence (a kind gift from B Strahl) was inserted into N-terminus of target protein in the MSCV based retroviral vector, followed by viral production and establishment of 293 stable expression cells. Proximity-dependent labeling of interacting proteins or BioID was conducted as previously described[37–39]. In brief, 293 stable cells were harvested from five of 15cm plates post-treatment with 50uM of biotin for 24hrs, and then washed twice with cold PBS. The cell pellets were resuspended in 1mL of RIPA lysis buffer (10% glycerol, 25mM Tris-HCl pH 8, 150mM NaCl, 2mM EDTA, 0.1% SDS, 1% NP-40, 0.2% Sodium Deoxycholate), and lysates were added with 1ul of Benzonase (Sigma-Aldrich, #E1014) followed by incubation on ice for one hour. After centrifugation at max speed for 30 min at 4 °C, the supernatant was collected and incubated with Neutravidin beads (Thermo Fisher #29204) overnight at 4 °C. The Neutravidin beads were then washed twice with the RIPA buffer and TAP lysis buffer (10% glycerol, 350mM NaCl, 2mM EDTA, 0.1% NP-40, 50mM HEPES pH 8) sequentially. Lastly, the beads were washed three times with the ABC buffer (50mM Ammonium bicarbonate pH 8) and subjected to mass spectrometry-based analysis.

Mass Spectrometry-based Protein Identification

Proteins were eluted from beads by adding 50μL 2× Laemmli buffer (Boston Bioproducts) and heating at 95°C for 5 minutes. A total of 50μL of each sample was resolved by SDS-PAGE using a 4–20% Tris-Glycine Wedge Well gel (Invitrogen) and visualized by Coomassie staining. Each SDS-PAGE gel lane was sectioned into 12 segments of equal volume. Each segment was subjected to in-gel trypsin digestion as follows. Gel slices were destained in 50% methanol (Fisher), 50 mM ammonium bicarbonate (Sigma-Aldrich), followed by reduction in 10 mM Tris [2-carboxyethyl] phosphine (Pierce) and alkylation in 50 mM iodoacetamide (Sigma-Aldrich). Gel slices were then dehydrated in acetonitrile (Fisher), followed by addition of 100 ng porcine sequencing grade modified trypsin (Promega) in 50 mM ammonium bicarbonate (Sigma-Aldrich) and incubation at 37°C for 12–16 hours. Peptide products were then acidified in 0.1% formic acid (Pierce). Tryptic peptides were separated by reverse phase XSelect CSH C18 2.5 um resin (Waters) on an in-line 150 × 0.075 mm column using a nanoAcquity UPLC system (Waters). Peptides were eluted using a 30 min gradient from 97:3 to 67:33 buffer A:B ratio (Buffer A = 0.1% formic acid, 0.5% acetonitrile; buffer B = 0.1% formic acid, 99.9% acetonitrile). Eluted peptides

were ionized by electrospray (2.15 kV) followed by MS/MS analysis using higher-energy collisional dissociation (HCD) on an Orbitrap Fusion Tribrid mass spectrometer (Thermo) in top-speed data-dependent mode. MS data were acquired using the FTMS analyzer in profile mode at a resolution of 240,000 over a range of 375 to 1500 m/z. Following HCD activation, MS/MS data were acquired using the ion trap analyzer in centroid mode and normal mass range with precursor mass-dependent normalized collision energy between 28.0 and 31.0. Proteins were identified by searching the UniProtKB database restricted to *Homo Sapiens* using Mascot (Matrix Science) with a parent ion tolerance of 3 ppm and a fragment ion tolerance of 0.5 Da, fixed modifications for carbamidomethyl of cysteine, and variable modifications for oxidation on methionine and acetyl on N-terminus. Scaffold (Proteome Software) was used to verify MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established with less than 1.0% false discovery by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established with less than 1.0% false discovery and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm[40]. Proteins were filtered out if they had a spectral count < 8 in all sample groups and the counts were normalized to log2 normalized spectral abundance factor (NSAF) values. Significant interacting proteins were defined with a cut-off of a log2 fold change over 2 in the experimental versus control samples.

Chromatin Immunoprecipitation (ChIP) followed by sequencing (ChIP-seq).

ChIP-seq was carried out as before[30,41]. In brief, cells were fixed in 1% formaldehyde (Thermo Scientific #28908) for 10 min, followed by quenching with 125 mM glycine for 5 min. Cells were then washed twice with cold PBS added with protease inhibitors (Sigma-Aldrich, #4693132001), and then subjected to resuspension and incubation in LB1 buffer (50mM HEPES-KOH pH 7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% TritonX-100), LB2 buffer (10mM Tris-HCl pH 8.0, 200mM NaCl, 1mM EDTA, 0.5mM EGTA), and LB3 buffer (10mM Tris-HCl pH 8.0, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1% Sodium Deoxycholate, 0.5% N-lauroylsarcosine). The cell nuclei were collected for sonication using Bioruptor sonicator (Diagenode, #B01020001; at high-energy setting for 45 cycles with 30 sec on and 30 sec off). After treatment with Triton X-100 (1% as a final concentration), the supernatant was collected after centrifugation (20,000g for 10min at 4 °C) for incubation with the dynabeads (Invitrogen, #11204D) that are pre-bound with antibodies for ~8 hrs at 4 °C. After a series of wash, the chromatin-protein complexes bound to beads were eluted, subject to reverse crosslink overnight at 65 °C, and treated with RNase (Roche, #11119915001; 1 hr at 37 °C) and then protease K (Roche, #03115828001; 2 hrs at 55 °C). The final DNA sample, as well as 1% of input chromatin, was recovered using PCR purification kit (Qiagen, #28106). The ChIP-seq library was prepared using NEBNext Ultra II kit (NEB, #E7645L) following the instructions of the manufacture's product manual. ChIP-seq libraries were sequenced on the Nextseq 550 system using Nextseq 550 High Output Kit v2.5 (Illumina, #20024906). For ChIP-Seq of HA-tagged N-IDR/A9 (with either WT or mutated IDRs), we used the matched input signals for signal normalization, which also displayed a good correlation to H3K27ac; for ChIP-Seq of GFP-tagged N-IDR/A9 (with either WT or mutated IDRs), we used signals of Drosophila spike-in chromatin for

normalization following the procedure described before[42] (Active Motif spike-in ChIP-seq reagents, cat #53083 and 61686).

ChIP-seq Data Analysis.

ChIP-seq data alignment, filtration, peak calling and assignment, and cross-sample comparison were performed as previously described[30,41] with slight modifications. In brief, ChIP-seq reads were aligned to human genome build GRCh37/hg19 or to mouse genome build GRCm38/mm10 using STAR version 2.7.1a[43]. The MACS2 software was used for peak identification with data from input as controls and default parameters[44]. Homer (ver 4.10.0) "annotatePeaks" and "findMotifsGenome" functions were used to annotate the called peaks and to find enriched motifs in these called peaks. Alignment files in the bam format were also transformed into read coverage files (bigWig format) using DeepTools[45]. Genomic binding profiles were generated using the deepTools "bamCompare" functions with options [--operation ratio ---pseudocount 1 -binSize 10 --extendReads 250] and normalized to the matched input. The resulting bigWig files were visualized in the Integrative Genome Viewer (IGV). Heatmaps for ChIP-seq signals were generated using the deepTools "computeMatrix" and "plotHeatmap" functions. ROSE were used for defining super-enhancers[46], with input signals used as control for normalization and peaks at +/− 2.5kb from transcriptional start site (TSS) excluded. Homer mergePeaks was used to determine overlap of ChIP-seq peaks with default settings.

RNA Sequencing (RNA-seq) and Data Analysis

RNA-seq was performed as described before[41,47]. In brief, total RNAs were purified using RNeasy Plus kit (Qiagen, #74136) and further processed with Turbo DNA-free kit (Thermo Fisher, #AM1907) to ensure the purity of RNA sample. For RNA-seq, the RNA samples were either sent to Novogene or processed using NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, #E7490) and NEBNext Ultra II RNA library Prep kit (NEB, #E7770) following the instructions of the manufacture's product manual. The multiplexed RNA-seq libraries were subjected for deep sequencing using the Illumina NextSeq500 platform (available in the UNC Sequencing Facility) with the Nextseq 550 High Output Kit v2.5 (Illumina, #20024906). For data analysis, RNA-seq reads were mapped to the reference genome followed by differential gene expression analysis as described before[41,47]. In brief, RNA-seq reads were mapped using MapSplice[48] and quantified using RSEM[49]. Read counts were upper-quantile normalized and log2 transformed. Raw read counts were used for differential gene expression analysis by DESeq[50]. Gene Ontology (GO) analysis was done using the C5 gene set of Molecular Signature Database (MsigDB) collections available in GSEA website[51].

ChIP or RT Followed by Quantitative PCR (ChIP-qPCR or RT-qPCR)

ChIP-qPCR or RT-qPCR was carried out as described before[36,41]. ChIP DNA was prepared with the same above procedure described for ChIP-seq whereas total RNA sample was used to generate cDNA with the iScript cDNA Synthesis kit (Biorad, #1708890) for qPCR.

Single Molecule Tracking, Lattice Light Sheet Microscopy, and Data Analysis

3D lattice light sheet microscopy movies of cells were acquired on a modified version of the lattice light sheet system as described before[52] using a square lattice excitation with numerical apertures of 0.4 (outer) and 0.3 (inner). Time intervals and imaging duration are specified in the legends for each dataset presented. Single molecule tracking was performed on the same system by focusing on a single plane within the nucleus of cells expressing Halo-tag protein fusions. Prior to imaging, cells were incubated win 1 nM of Halo Tag-Janelia Fluor 549 ligand for 20 minutes and then washed in PBS[53]. After transferring to the microscope, single planes within the nucleus of each cells were imaged under the same lattice illumination parameters above for a total of 20,000–40,000 frames with 20 ms exposures. Prior to tracking, images were pre-processed with a rolling ball background subtraction and histogram equalization contrast enhancement using ImageJ. Single molecules were then tracked using the TrackMate plugin for ImageJ[54]. To account for variation in protein expression levels between cells and avoid potential tracking artifacts due to different densities of fluorescent molecules, Pandas software library for python[55] was used to register single particle tracking datasets such that the number particles within a rolling 100 window was consistent both within and between conditions. We controlled for photobleaching and phototoxicity by confirming that mean molecular speeds within a single cell did not vary substantially throughout the course of the imaging experiment. Finally, molecular trajectories were fit to a two-state kinetic model using Spot-On[22] to estimate the mean diffusion coefficients and fraction of molecular populations for both the slow-diffusing/bound state and rapidly diffusing/free state.

In Situ Hi-C

In situ Hi-C was performed exactly as described by Rao *et al*[56]. Five million of cells were crosslinked in 1% formaldehyde for ten minutes with stirring and quenched by adding 2.5M glycine to a final concentration of 0.2M for 5 min with rocking. Cells were pelleted by spinning at 300g for 5 min at 4°C. The pellet was washed with cold PBS and spun again prior to freezing in liquid nitrogen. Cells were lysed with 10mM Tris-HCl pH8.0, 10mM NaCl, 0.2% Igepal CA630 and protease inhibitors (Sigma, P8340) for 15 min on ice. Cells were pelleted and washed once more using the same buffer. Pellets were resuspended in 50μl of 0.5% SDS and incubates for 7 min at 62°C. Next, reactions were quenched with 145μl of water and 25μl of 10% Triton X-100 (Sigma, 93443) at 37°C for 15 min. Chromatin was digested overnight with 25μl of 10X NEBuffer2 and 100U of MboI at 37°C with rotation. Reactions were incubated at 62°C for 20 min to inactivate MboI and then cooled to room temperature. Fragment overhangs were repaired by adding 37.5 μl of 0.4mM biotin-14-dATP, 1.5 μl of 10mM dCTP, 1.5 μl of 10mM dGTP, 1.5 μl of 10mM dTTP, and 8 μl of 5U/ul DNA Polymerase I, Large (Klenow) Fragment and incubating at 37°C for 1.5 hr. Ligation was performed by adding 667 μl of water, 120 μl of 10X NEB T4 DNA ligase buffer, 100 μl of 10% Triton X-100, 12 μl of 10 mg/ml BSA, and 1 μl of 2000 U/ul T4 DNA Ligase and incubating at room temperature for 4 hr with slow rotation. Samples were pelleted at 2500g and resuspended in 432 μl water, 18 μl 20 mg/ml proteinase K, 50 μl 10% SDS, 46 μl 5M NaCl and incubated for 30 min at 55°C. The temperature was raised to 68°C and incubated overnight. Samples were cooled to room temperature. 874 μl of pure ethanol and 55 μl of 3M sodium acetate pH 5.2 were added to each tube which were subsequently

incubated for 15 min at −80°C. Tubes were spun at max speed at 2°C for 15 min and washed twice with 70% ethanol. The resulting pellet was resuspended in 130 μl of 10mM Tric-HCl, pH8 and incubated at 37°C for 15 min. DNA was sheared using an LE220 Covaris Focused-ultrasonicator to a fragment size of 300–500 bp. Sheared DNA was size selected using AMPure XP beads. 110 μl of beads were added to each reaction and incubated for 5 min. Using a magnetic stand supernatant was removed and added to a fresh tube. 30μl of fresh AMPure XP beads were added and incubated for 5 min. Beads were separated on a magnet and washed twice with 700 μl of 70% ethanol without mixing. Beads were left to dry and then sample was eluted using 300 μl of 10 mM Tris-HCl, pH 8. 150 of 10 mg/ml Dynabeads MyOne Streptavidin T1 beads were washed resuspended in 300 μl of 10 mM Tris HCl, pH 7.5. This solution was added to the samples and incubated for 15 min at room temperature. Beads were washed twice with 600μl Tween Washing Buffer (TWB; 250 μl Tris-HCl, pH 7.5, 50 μl 0.5 M EDTA, 10 mL 5M NaCl, 25 μl Tween-20, and 39.675 mL water) at 55°C for 2 min with shaking. Sheared ends were repaired by adding 88 μl 1× NEB T4 DNA ligase buffer with 1mM ATP, 2 μl of 25 mM dNTP mix, 5 μl of 10U/μl NEB T4 PNK, 4ul of 3U/μl NEB T4 DNA polymerase I, 1μl of 5U/μl NEB DNA polymerase I, Large (Klenow) Fragment and incubating at room temperature for 30 min. Beads were washed two more times with TWB for 2 min at 55°C with shaking. Beads were washed once with 100 μl 1× NEBuffer 2 and resuspended in 90 μl of 1X NEBuffer 2, 5 μl of 10 mM dATP, 5μl of 5U/μl NEB Klenow exo minus, and incubated at 37°C for 30 min. Beads were washed two more times with TWB for 2 min at 55°C with shaking. Beads were washed once in 50 μl 1× Quick Ligation reaction buffer and resuspended in 50 μl 1X Quick Ligation reaction buffer. 2 μl of NEB DNA Quick ligase and 3 μl of an illumina indexed adaptor were added and the solution was incubated for 15 min at room temperature. Beads were reclaimed using the magnet and washed two more times with TWB for 2 min at 55°C with shaking. Beads were washed once in 100 μl 10 mM Tris-HCl, pH 8 and resuspended in 50 μl 10 mM Tris-HCl, pH8. Hi-C libraries were amplified for 7–12 cycles in 5 μl PCR primer cocktail, 20 μl of Enhanced PCR mix, and 25 μl of DNA on beads. The PCR settings included 3 min of 95°C followed by 7–12 cycles of 20 s at 98°C, 15 s at 60°C, and 30 s at 72°C. Samples were then held at 72°C for 5 min before lowering to 4°C until samples were collected. Amplified samples were brought to 250 μl with 10 mM Tris-HCl, pH 8. Samples were separated on a magnet and supernatant was transferred to a new tube. 175 μl of AMPure XP beads were added to each sample and incubated for 5 min. Beads were separated on a magnet and washed once with 700 μl of 70% ethanol. Supernatant was discarded. 100 μl of 10 mM Tris-HCl and 70 μl of fresh AMPure XP beads were added and the solution was incubated for 5 min at room temperature. Beads were separated with a magnet and washed twice with 700 μl 70% ethanol. Beads were left to dry until cracking started to be observed and eluted in 25 μl of Tris HCl, pH 8.0. The resulting libraries were next quantified by Qubit and Bioanalyzer. A low depth sequencing was performed first using the MiniSeq sequencer system (Illumina) and analyzed using the Juicer pipeline[57] to assess quality control prior to deep sequencing (NovaSeq S4). Each Hi-C library was assessed in biological and technical duplicate achieving a total of 3 billion reads per cell line.

Hi-C Data Processing and Analysis

In situ Hi-C datasets were processed using the Juicer Hi-C pipeline with default parameters as described in Durand *et al.*[57]. MboI was used as the restriction enzyme, and reads were aligned to the hg19 human reference genome with bwa (version 0.7.17). Data was processed for 3,058,370,530 Hi-C read pairs in N-IDR$_{WT}$/A9 cells, yielding 1,791,818,927 Hi-C contacts (58.59%) and 2,914,343,903 Hi-C read pairs in N-IDR$_{FS}$/A9 cells, yielding 1,708,441,327 Hi-C contacts (58.62%). Hi-C matrices were constructed for each individual replicate for downstream analysis. A Hi-C mega map was constructed by combining all replicates for each condition (i.e. N-IDR$_{WT}$/A9 or N-IDR$_{FS}$/A9). For visualization, the resulting Hi-C contact matrices were normalized with a matrix balancing algorithm according to Knight *et al.*[58] ("KR") to adjust for regional background differences in chromatin accessibility.

Loops were detected using HiCCUPS from the Juicer tools software (version 1.11.09) as described in Rao *et al.*[56] via the following command: "hiccups -m 2048 -c 2 -r 5000,10000,25000 -k KR -f 0.1,0.1,0.1 -p 4,2,1 -i 8,6,4 -t 0.2,1.5,1.5,1.75 -d 30000,30000,60000". 4,788 loops were identified in N-IDR$_{WT}$/A9 and 2,826 loops were identified in N-IDR$_{FS}$/A9 for a total of 7,616 loops at 10 Kb resolution. After filtering out redundant loops, 6,615 combined loops remained. Unnormalized loop counts were extracted using the straw api[57] for all loops in each replicate (8 total). Differential loops between N-IDR$_{WT}$/A9 and N-IDR$_{FS}$/A9 were determined using DESeq2[59], including biological replicate and condition as covariates in the model. 232 N-IDR$_{WT}$/A9-specific loops and 52 N-IDR$_{FS}$/A9-specific loops were considered significantly differential at a Benjamini-Hochberg adjusted p-value 0.01.

Aggregate peak analysis (APA) of N-IDR/A9 binding site interactions was conducted in R using straw. All unique, paired interactions between the 500 strongest N-IDR$_{WT}$/A9 ChIP-seq binding sites were categorized into 1) inter-chromosomal (n=95959), 2) long (>=2Mb) intra-chromosomal (n=6298), or 3) short (<2Mb) intra-chromosomal (n=574) interactions. Short interactions were filtered out such that the corner of the APA plot would not intersect the diagonal, reducing them from n=574 to n=309. Unnormalized pixel values +/− 10 surrounding pixels were extracted from N-IDR$_{WT}$/A9 and N-IDR$_{FS}$/A9 Hi-C files at 10Kb resolution for each interaction pair. Resulting 21×21, 10Kb pixel matrices were aggregated and normalized to the number of binding site pairs.

Aggregate peak analysis (APA) of differential loop calls was conducted in R using straw. APA was run for all loops (n=6615), N-IDR$_{WT}$/A9-specific loops (n=232), and N-IDR$_{FS}$/A9-specific loops (n=52) using both N-IDR$_{WT}$/A9 and N-IDR$_{FS}$/A9 Hi-C. Short interactions were filtered out as described above, reducing the number of interactions to n=3427, n=121, and n=24 for all, N-IDR$_{WT}$/A9-specific and N-IDR$_{FS}$/A9-specific loops, respectively. Unnormalized pixels were extracted with straw producing a 21×21 pixel matrix at 10Kb resolution that was aggregated and normalized by the number of loops per group.

All loops were partitioned as either N-IDR$_{WT}$/A9-specific loops (WT loops) or N-IDR$_{FS}$/A9-specific loops (FS loops) based on differential loop calling (as described above) and then split into separate loop anchors. Loop anchors were then intersected (bedtoolsr) with

several features including ChIP-seq peaks for NUP98-HOXA9, CTCF, or H3K27Ac in both cell types (N-IDR$_{WT}$/A9 or N-IDR$_{FS}$/A9) and with promoter regions (defined as 1000 bp upstream of transcription start sites). Permutation testing was used to calculate p-values for each feature's intersection with loop anchors. In short, the observed percentage of each feature present at WT or FS loop anchors was calculated. The expected percentage was determined by randomly sampling an equivalent number of loop anchors from all loop anchors called, then calculating the percentage overlap with each feature. This procedure was repeated 1,000 times to create a distribution of expected values. P-values were determined by summing the number of expected values greater than (or less than if the observed value was less than the mean) the observed value for that feature.

All loops were partitioned as either N-IDR$_{WT}$/A9-specific loops (WT loops) or N-IDR$_{FS}$/A9-specific loops (FS loops) based on differential loop calling (as described above). Each loop was then intersected with 5 kb windows around the transcription start sites of genes using the bedtoolsr "pairtobed" function with either end of the loop constituting an overlap. The log2 fold-change in expression value (WT/FS) of genes overlapping either end of a WT or FS differential loop were plotted along with the expression of all genes. A Dunn's multiple comparison test following a Kruskal-Wallis test showed a statistically significant difference in expression between WT-specific gene-loops and either FS-specific gene-loops ($p = 0.015$) or all genes ($p < 0.001$), after p-value correction with the Benjamini-Hochberg procedure. In this study, WT-specific loops were present in the N-IDR$_{WT}$/A9-expressing cells and absent in N-IDR$_{FS}$/A9 cells while mutant-specific loops were absent in N-IDR$_{WT}$/A9 cells and present in N-IDR$_{FS}$/A9 cells, supporting accurate calling of differential loops.

Chromatin Conformation Capture (3C) Followed by qPCR (3C-qPCR)

Cell samples were processed and analyzed as previously described with slight modifications[36]. Briefly, 10 million of cells were fixed in 1% formaldehyde at room temperature for 10 min, followed by quenching in 0.125M glycine for 5 min. Fixed cells were washed in cold PBS and lysed in ice-cold lysis buffer (10mM Tris-Cl, pH8.0, 10mM NaCl, 0.2% NP-40, 1X complete protease inhibitor cocktail) for 1hr at 4 °C. Nuclei were collected by centrifugation at 5000rpm for 5 min and digested with 800 unit of Bgl-II enzyme, added with 0.3% of SDS and 1.8% of Triton X-100 in the molecular-grade water with respective enzyme digestion buffer (1.2X) for overnight at 37 °C. After inactivation at 65°C for 20 min with 1.6% of SDS, digested chromatin was subjected to ligation by T4 ligase (NEB) with 1% Triton X-100 for overnight at 16 °C, followed by 30 min incubation at room temperature. Ligated chromatin was treated with protease K for overnight at 65 °C and then treated with RNase for 2hr at 37 °C, followed by DNA purification with the phenol-chloroform extraction protocol. For q-PCR, the obtained DNA were diluted by 50 folds and used as template. Primers were designed for the respective genomic loci with chromatin loop as detected by Hi-C mapping experiment. All PCR products were sequenced to confirm that they are indeed correctly ligated products from two distant genomic loci where chromatin loop is expected to form between them. All the primers used for 3C-qPCR are listed in Supplementary Table 8.

Statistics and Reproducibility.

Experimental data are presented as the mean ± SD of three independent experiments unless otherwise noted. Statistical analysis was carried out with two-sided Student's t-test for comparing the two sets of data with assumed normal distribution. We used a log-rank test for the Kaplan-Meier survival curve to define statistical significance. A *P* value less than 0.05 was considered to be significant. Statistical significance levels are denoted as follows: *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001. Sample numbers are indicated in the figure legends. Results of imaging, staining, protein sample examination, and western blotting were reproducible in three experiments, with the representative ones shown in the figure.

Reporting summary

Additional information related to experimental design is available in the Nature Research Reporting Summary linked to this paper.

Code Availability

The scripts for genomic data analyses and all other data are available from the corresponding author upon request.

Data Availability

Next-generation sequencing datasets including those of ChIP-seq, RNA-seq and Hi-C used in this current study are deposited in the NCBI GEO under the accession number GSE144643. The mass spectrometry-based proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD023548 and 10.6019/PXD023548. Source data are provided with this paper.

Extended Data
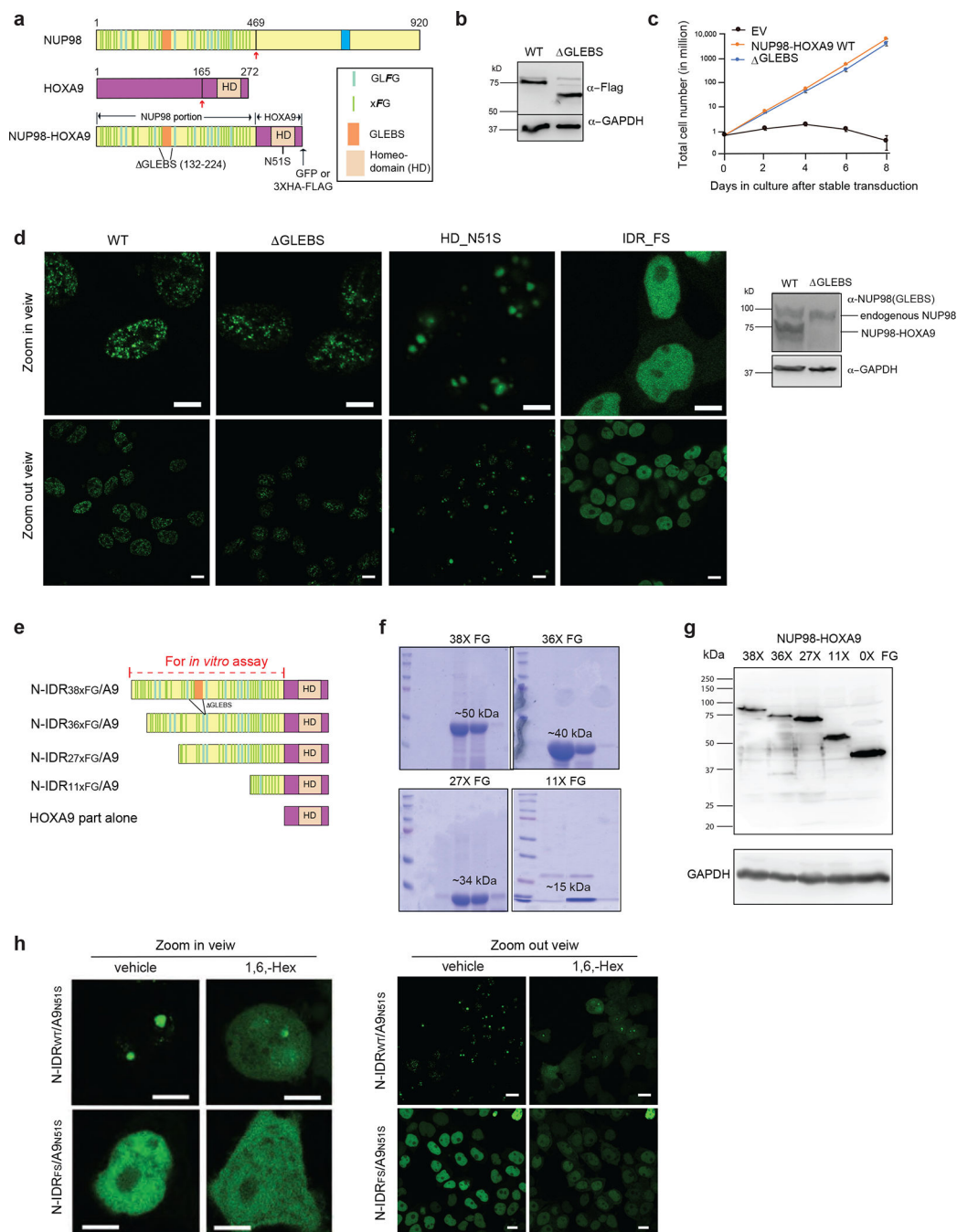
Extended Data Figure 1|. Intrinsically disordered region (IDR) retained within the leukemia-related chimeric NUP98-HOXA9 forms phase-separated condensates *in vitro* and is essential for establishing phase-separated fusion protein assemblies in the nucleus.

a, Schematic showing the domain structure of full-length NUP98 (top), full-length HOXA9 (middle) and NUP98-HOXA9 chimera (bottom; with either GFP or 3XHA-3XFLAG tag fused to C-terminus). GLFG or non-GLFG (xFG) motif contents and other important domains are shown in the box.

b, Immunoblotting of full-length (WT) or GLEBS-deleted NUP98-HOXA9, as detected by indicated antibodies, after stable transduction into primary murine hematopoietic stem/progenitor cells (HSPCs).

c, Proliferation of murine HSPCs stably transduced with full-length (WT) or GLEBS-deleted NUP98-HOXA9, relative to empty vector-infected controls (n=3 stably transduced cell cultures per group). Data are presented as mean ± standard deviation (SD).

d, Live cell fluorescence imaging (GFP; zoomed-in and zoomed-out views on the top and bottom) of 293 cells with stable transduction of GFP-tagged NUP98-HOXA9, either full-length (WT), GLEBS-deleted (also referred to as N-IDR$_{WT}$/A9), or that with a DNA-binding-disrupting mutation in homeodomain (HDN51S) or a F→S mutation at FG-repeats (IDRFS, also referred to as N-IDR$_{FS}$/A9) that substitutes Phe residues within FG-repeats to Ser. The right panel shows immunoblotting of normal NUP98 and the stably transduced NUP98-HOXA9, either full-length (WT) or GLEBS-deleted, as detected by a previously described antibody raised against GLEBS of NUP98, into 293 cells. Scale bar, 10μm.

e, Schematic of the indicated N-IDR fusion domains with a varying number of FG-repeats. The IDR portion used for *in vitro* assay in Fig 1d is indicated by a red dotted line.

f, SDS-PAGE images showing recombinant protein of N-IDR domain (see red label in panel **e**) with the indicated varying number of FG-repeats (His6×-tagged), purified with Ni-column and an additional size exclusion column purification step. The protein size is labeled above the recombinant protein.

g, Anti-GFP immunoblotting for GFP-tagged NUP98-HOXA9 chimera with the indicated varying number of FG-repeats described in panel **e** after stable transduction into 293 cells.

h, Live cell fluorescence for the N51S-mutated N-IDR/A9 (GFP-tagged) with either WT (top) or the F→S mutated IDR (bottom) in 293 stable expression lines before (left) and after (right) treatment of 10% of 1,6-hexanediol for one minute. The left panels show zoomed-in images of a representative cell from the right panels of zoomed-out images. Scale bar, 10 μm.
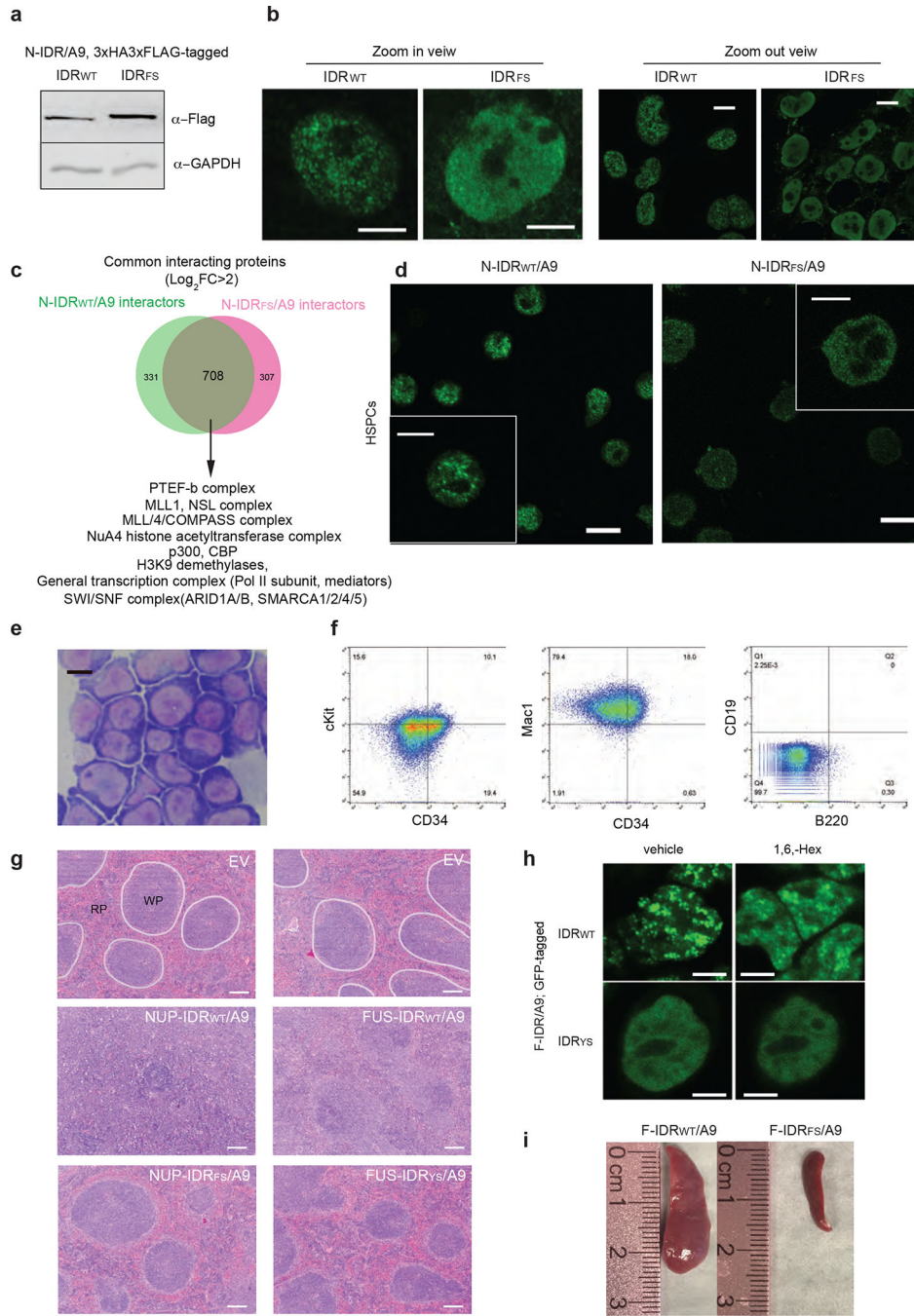
Extended Data Fig 2|. IDR harbored within the leukemia-related chimeric TF fusion is required for leukemic transformation of primary murine HSPCs.

a-b, Immunoblotting (panel **a**) and fixed cell immunostaining (panel **b**; anti-FLAG) of the LLPS-competent N-IDR$_{WT}$/A9 and LLPS-incompetent N-IDR$_{FS}$/A9 after stable transduction into 293 cells. The left side of panel b shows a zoomed-in view of the right side. Scale bar, 10μm.

c, Venn diagram shows significant overlap between N-IDR$_{WT}$/A9 and N-IDR$_{FS}$/A9 interactomes detected by BioID. Examples of the detected interacting proteins are shown below.

d-f, Immunostaining (panel **d**; anti-GFP), Wright-Giemsa staining (**e**) and FACS with the indicated surface marker (**f**) using murine HSPCs transformed by N-IDR$_{WT}$/A9 (GFP or 3xHA-3xFLAG-tagged) one month post-transduction, which reveals a typical acute myeloid leukemia cell phenotype (cKit+, Cd34+, MacI^{high}, CD19-, B220-). The insert in panel d shows a zoomed-in view of the representative cell. Scale bar, 5 μm.

g, Haematoxylin-Eosin (H&E)-stained spleen section images for the indicated cohort at 10X magnification. White Pulp (WP) is outlined with white line for the sample from mice transplanted with empty vector (EV)-infected HSPCs (Top). Note that clear demarcation between WP and Red Pulp (RP), as observed in cohorts receiving either EV or the mutant forms of fusion (bottom panels), is lost in those with N-IDR$_{WT}$/A9 and F-IDRWT/A9 (middle panels) due to an excessive expansion of transformed HPSCs that infiltrated into spleen leading to splenomegaly observed in panel **i.**

h, Live cell fluorescence (GFP) imaging of 293 cells with stable expression of an artificial HOXA9 chimera created by replacing NUP98's FG-repeats with IDR of an unrelated RNA-binding protein FUS, either WT or Y→S mutated (hereafter referred to as the F-IDRWT/A9 and F-IDRYS/A9 fusion, respectively), before and after treatment with 10% of 1,6-hexanediol for one minute. Scale bar, 10 μm.

i, Representative image of spleen from mice seven months post-transplantation of murine HPSCs stably transduced with either F-IDRWT/A9 (left) or F-IDRYS/A9 (right).
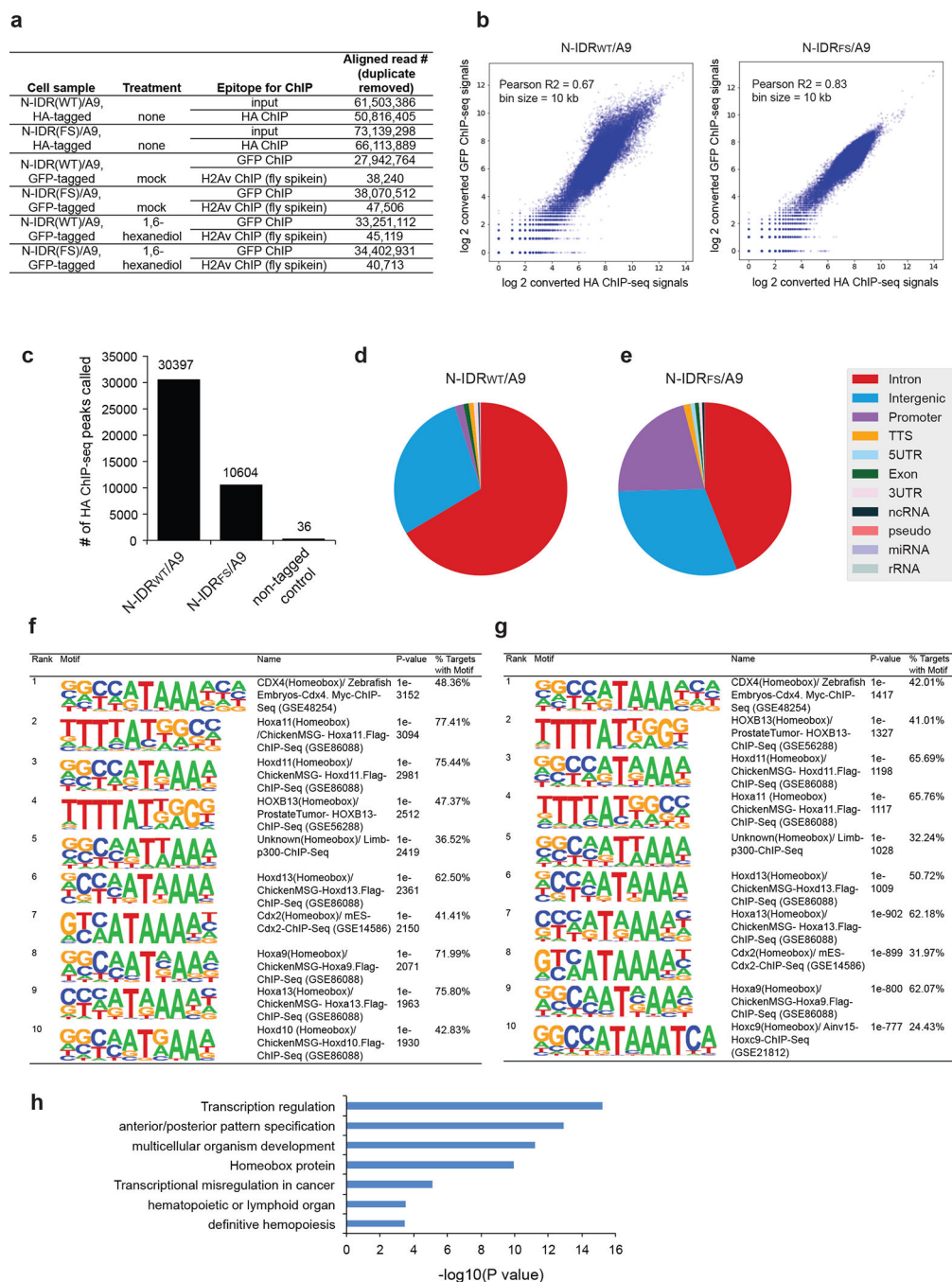
Extended Data Fig 3|. ChIP-seq reveals binding patterns of NUP98-HOXA9 that carries either WT or an F→S mutated IDR.

a, Summary of the counts of ChIP-seq read tags for the indicated samples.

b, Scatterplots showing correlation of global N-IDR$_{WT}$/A9 (left) or N-IDR$_{FS}$/A9 (right) ChIP-seq signals using either HA (x-axis) or GFP (y-axis) antibodies in two biological replicates of 293 stable cells. Coefficient of determination (R2) is determined by Pearson correlation.

c, Total number of the called HA ChIP-seq peaks in stable 293 cells expressing HA-tagged N-IDR$_{WT}$/A9 (left) or N-IDR$_{FS}$/A9 (middle) or empty vector control (right).

d-e, Pie chart showing distribution of the indicated annotation feature among the called N-IDR$_{WT}$/A9 **(d)** or N-IDR$_{FS}$/A9 **(e)** ChIP-seq peaks in 293 stable expression cells.

f-g, Summary of the most enriched motifs identified within the called N-IDR$_{WT}$/A9 **(f)** or N-IDR$_{FS}$/A9 **(g)** ChIP-Seq peaks in 293 stable expression cells.

h, Gene Ontology (GO) analysis of genes associated with broad super-enhancer-like peaks of N-IDR$_{WT}$/A9 as identified in 293 stable cells.
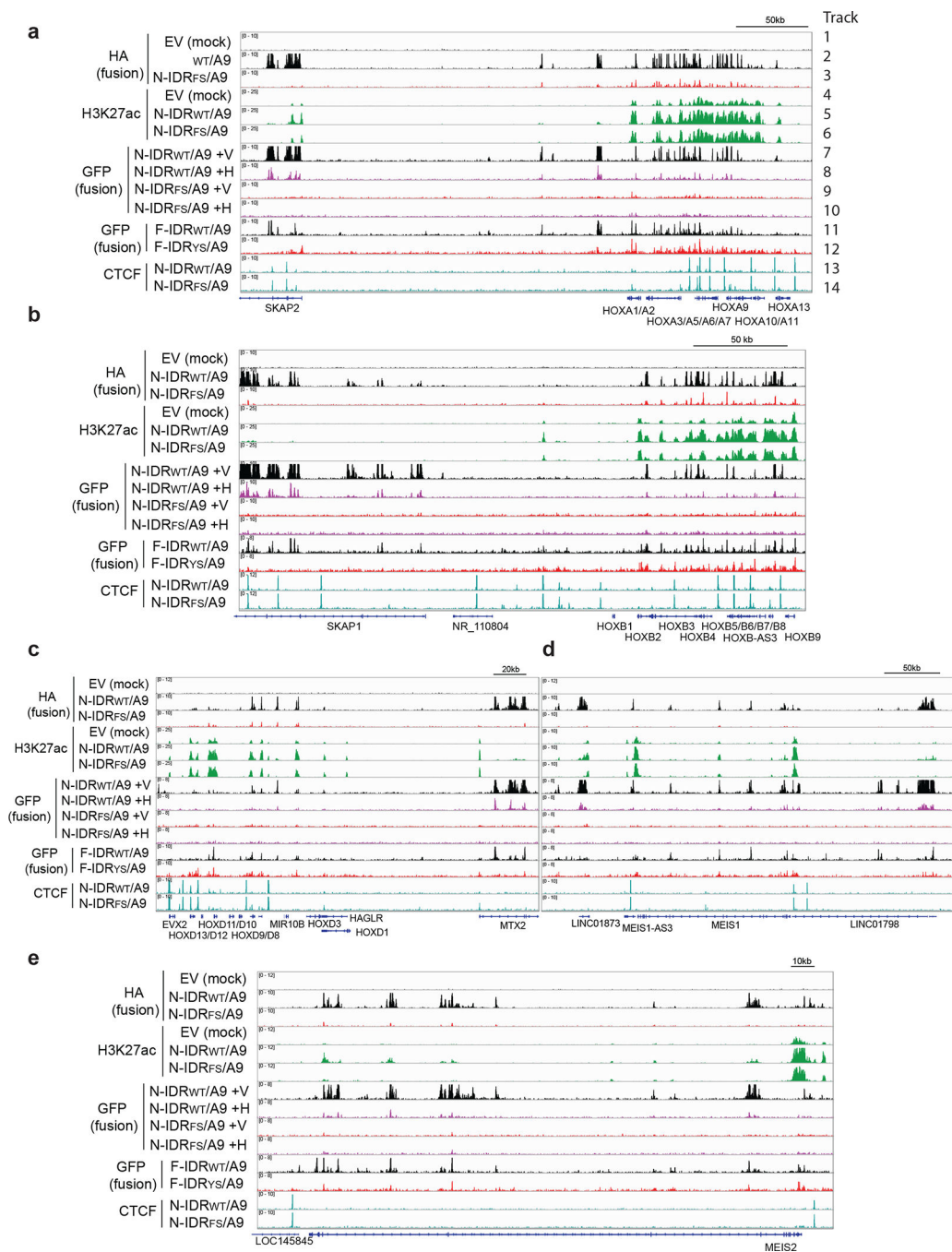
Extended Data Fig 4|. Dramatically enhanced chromatin occupancy, as well as a broad super-enhancer-like binding pattern typically seen at leukemia-related genomic loci, is characteristic for the LLPS-competent NUP98-HOXA9 (N-IDR$_{WT}$/A9) and not its LLPS-incompetent IDR mutant (N-IDR$_{FS}$/A9).

a-e, Integrative Genomics Viewer (IGV) views for the indicated ChIP-seq signal at the well-known leukemia-associated loci such as the HOXA (**a**), HOXB (**b**) and HOXD (**c**) gene clusters, MEIS1 (**d**) and MEIS2 (**e**).

Samples from top to bottom are HA (tracks 1–3) and H3K27ac (tracks 4–6) ChIP-seq signals in the 293 cells stably expressed with either empty vector (tracks 1 and 4; EV

as negative control for ChIP specificity) or the HA-tagged N-IDR$_{WT}$/A9 (tracks 2 and 5) or N-IDR$_{FS}$/A9 (tracks 3 and 6), GFP ChIP-seq signals (tracks 7–12) in the 293 cells stably expressed with GFP-tagged N-IDR$_{WT}$/A9 (tracks 7–8 represent samples post-treatment of vehicle or 1,6-hexanediol, respectively), N-IDR$_{FS}$/A9 (tracks 9–10 represent samples post-treatment of vehicle or 1,6-hexanediol, respectively), F-IDRWT/A9 (track 11) or F-IDRYS/A9 (track 12), as well as CTCF ChIP-seq in 293 cells with N-IDR$_{WT}$/A9 (track 13) or N-IDR$_{FS}$/A9 (track 14). HA and CTCF ChIP-seq signals were normalized to input signals, whereas GFP ChIP-seq, conducted in the spike-in controlled experiments, normalized to the spike-in Drosophila chromatin signals (those from antibody of a *Drosophila* specific histone, H2Av).
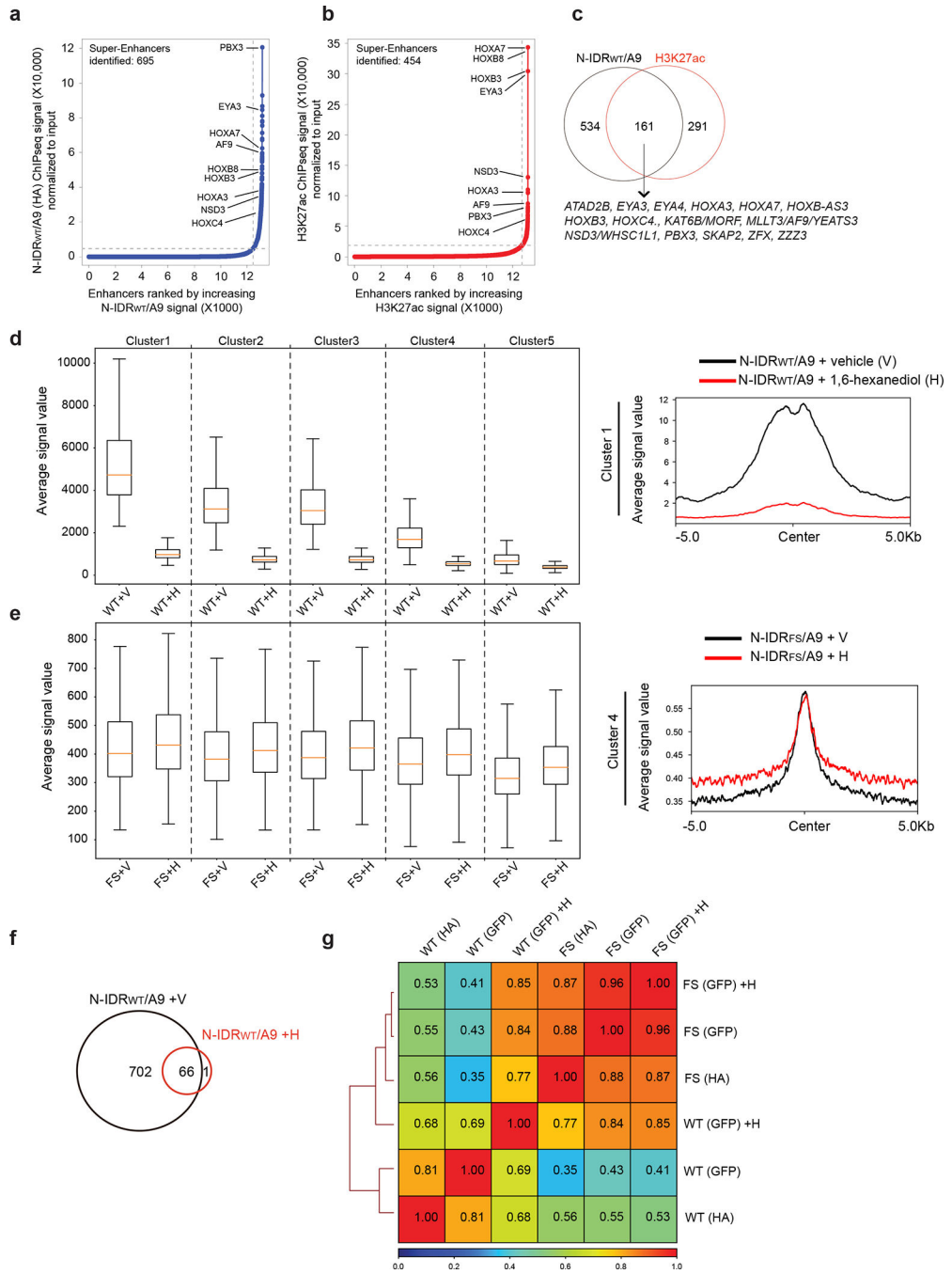
Extended Data Fig 5|. Formation of enhanced and broad super-enhancer-like binding patterns of leukemia-related chimera TFs requires an intact phase-separation-competent IDR.

a-b, Hockey-stick plot shows distribution of the input-normalized ChIP-seq signals of N-IDR$_{WT}$/A9 (**a**) or H3K27ac (**b**) across all enhancers annotated by H3K27ac peaks (TSS +/−2.5kb regions were excluded) in 293 cells. Dotted line indicates the threshold level set by the ROSE algorithm to call super-enhancers. Relative rankings of super-enhancers associated with some example genes are shown.

c, Venn diagram illustrates overlap among super-enhancers called based on N-IDR$_{WT}$/A9 and H3K27ac ChIP-seq signals.

d-e, The K-means clustered box plots of averaged ChIP-seq signals of the LLPS-competent N-IDR$_{WT}$/A9 (panel **d**; WT) show a dramatic reduction in binding post-treatment of 293 stable cells with 1,6-Hexanediol (WT+H), relative to mock (WT+V); this is particularly significant for peak clusters 1–3 shown in the main Figure 2b. In contrast, N-IDR$_{FS}$/A9 binding (panel **e**) shows insensitivity to the same treatment of 1,6-Hexanediol (FS+H) in comparison to mock (FS+V). On the right, averaged ChIP-seq signal distribution profiles are shown for N-IDR$_{WT}$/A9 and N-IDR$_{FS}$/A9 over a 10kb region in the indicated cluster as an example.

f, Venn diagram to compare genes associated with the super-enhancer-like, broad N-IDR$_{WT}$/A9 peaks after treatment of 1,6-Hexanediol (+H), relative to vehicle mock (+V).

g, Hierarchical clustered heatmaps for the pairwise correlation of ChIP-Seq signals between each of the indicated sample. The coefficients were determined by Pearson correlation. HA and GFP represent HA and GFP ChIP-seq for chimera TFs, respectively; +H and +V represent treatment of 1,6-Hexanediol and vehicle mock, respectively.

a, The K-mean clustered heatmaps of NUP98 IDR fusion (N-IDR$_{WT}$/A9 and N-IDR$_{FS}$/A9; two panels on the left) and FUS IDR fusion (F-IDRWT/A9 and F-IDRYS/A9; two panels on the right) reveal a similarly enhanced binding for the LLPS-competent chimera that carries a WT form of IDR, relative to its LLPS-incompetent IDR mutant in 293 stable expression cells. Note that, albert to a less degree, the artificially created F-IDRWT/A9

Extended Data Fig 6|. Similar to what was seen with NUP98 IDR (N-IDR) fusion, the phase-separation-promoting property harbored within an unrelated IDR of FUS (FIDR) is sufficient to induce the enhanced binding of chimeric TF.

fusion also displays a broad, super-enhancer-like binding pattern at same sites observed with N-IDR$_{WT}$/A9 fusion.

b, Pie chart showing distribution of the indicated genomic annotation features among the ChIP-Seq peaks of GFP-tagged F-IDRWT/A9 (left) or F-IDRYS/A9 (right) in the 293 stable expression cells.

c, The K-mean clustered heatmaps (left) and its averaged ChIP-seq signal distribution profiles (right) of NUP98 IDR fusion (N-IDR$_{WT}$/A9) in the transformed murine HPSCs.

d, Venn diagram shows overlap between the annotated genes found in clusters 1–3 of ChIP-seq profiles of transformed murine HPSCs and 293 stable expression cells. Examples of the shared critical oncogenes are shown below.

e, IGV views of N-IDR$_{WT}$/A9 ChIP-seq signals (GFP-tagged) at the indicated loci in murine leukemia cells transformed by this chimera.

f, ChIP-qPCR for binding of the GFP-tagged N-IDR$_{WT}$/A9 or N-IDR$_{FS}$/A9 at *CCL15* (a negative control region), *PBX3* and *HOXA9* in 293 stable cells post-treatment with 1,6-Hexanediol for one minute (+H), relative to mock (+V). Data are presented as mean ± ±SD of three replicate experiments. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$; n.s., not significant.

g, ChIP-qPCR for binding of the GFP-tagged F-IDRWT/A9 or F-IDRYS/A9 at *CCL15* (a negative control region), *PBX3* and *HOXA9* in the 293 stable cells. Data are presented as mean ±SD of three replicate experiments. $***P < 0.001$.
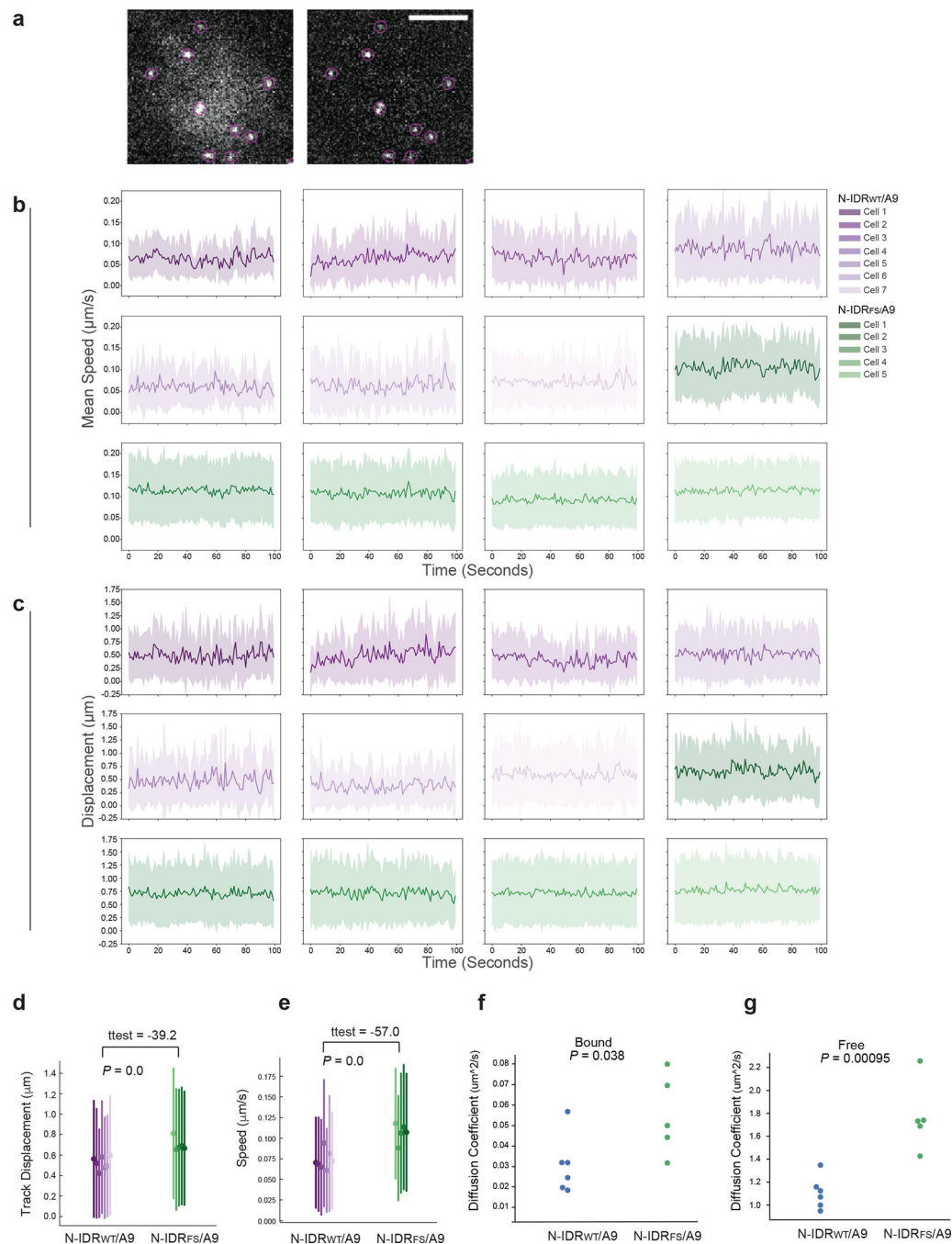
Extended Data Figure 7|. Single-molecule tracking (SMT) shows that phase-separation-competent N-IDR$_{WT}$/A9 proteins behave with less dynamic characteristics, compared to phase-separation-incompetent N-IDR$_{FS}$/A9.

a, Representative images of single-molecule particles identification in an N-IDR$_{WT}$/A9-expressing cell, either the original captured image (left) or after processing to remove background (right). Scale bar represents 5 μm.

b-c, Single particle tracks for mean speed (panel **b**) and mean displacement (**c**) of either N-IDR$_{WT}$/A9 or N-IDR$_{FS}$/A9 single molecules within the temporally registered reference frame binned into one-second intervals.

d-e, Displacement (**d**) and mean velocity (**e**) of single particle tracks indicate that N-IDR$_{WT}$/A9 with the LLPS-competent IDR (WT) is less mobile and navigates nuclear space at a slower rate than its LLPS-incompetent IDR mutant (FS). Dots indicate mean values in a single cell. Line indicates one standard deviation.

f-g, The diffusion coefficient for chromatin-bound (**f**) and freely diffusing states (**g**) of N-IDR$_{WT}$/A9 or N-IDR$_{FS}$/A9, calculated based on SMT studies of its 293 stable expression cells.
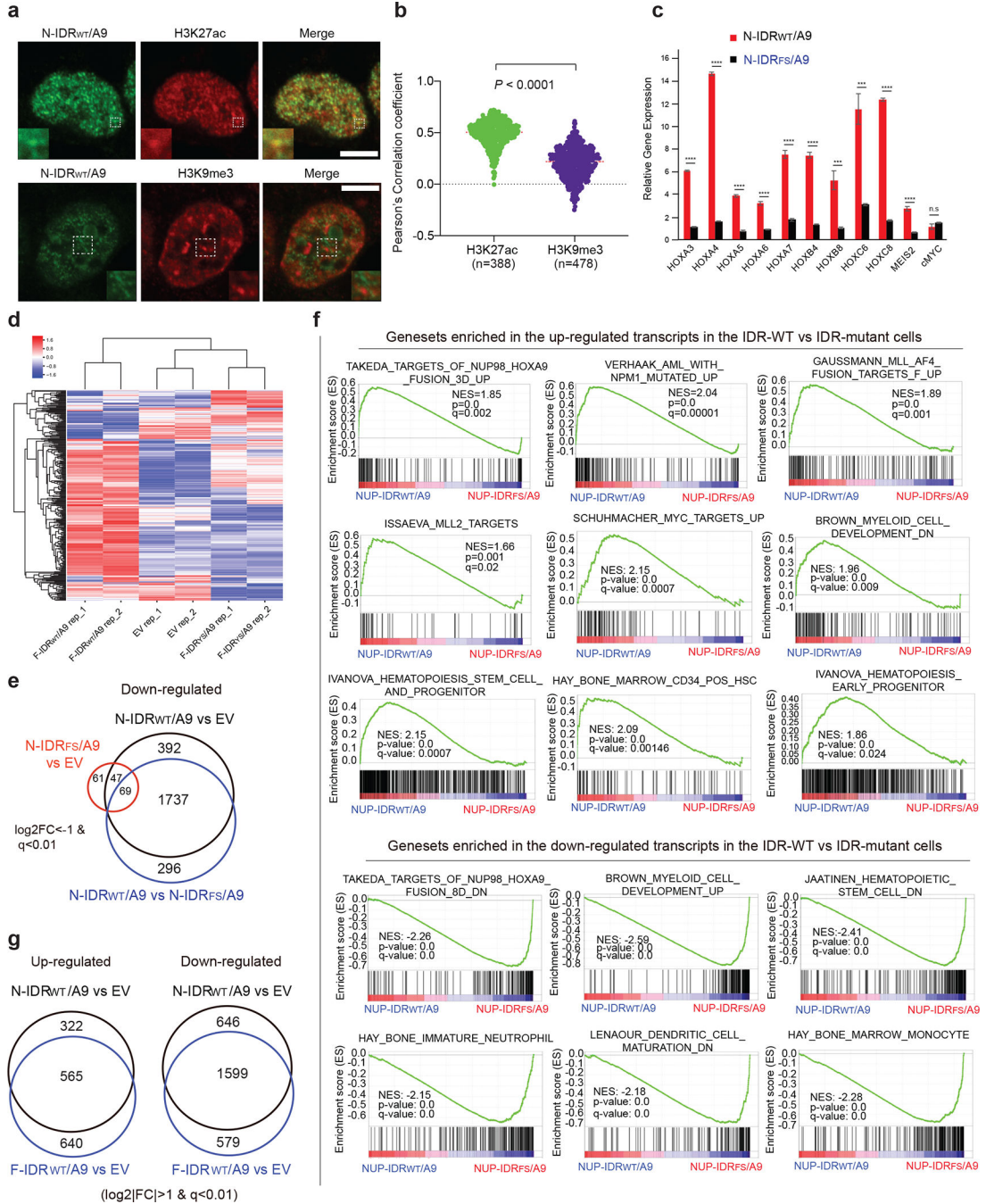
Extended Data Fig 8|. A LLPS-competent IDR harbored within leukemia-related TF chimera is essential for potentiating transcriptional activation of the downstream oncogenic gene-expression program.

a, Fixed cell immunostaining for the 3xHA-3xFLAG-tagged N-IDR$_{WT}$/A9 (left panels; anti-FLAG) and the indicated histone modification (middle panels) in 293 stable cells. Shown in the lower insert are enlarged images of an example region (white dotted box) where chimera is co-localized with H3K27ac (top) and not H3K9me3 (bottom). Scale bar, 10 μm.

b, Pearson's correlation coefficient values between N-IDR$_{WT}$/A9 and the indicated histone marks. The red dotted line indicates the calculated average value of each plot. The calculated means (red dotted lines) were compared with an independent two-tailed Student's T-test.

c, RT-qPCR to assess the impact of phase separation in target gene expression in 293 cells. All of the tested HOX and MEIS2 genes are direct targets of both N-IDR$_{WT}$/A9 and N-IDR$_{FS}$/A9. cMYC is not a direct target gene serving as a negative control. Note that only LLPS-competent N-IDR$_{WT}$/A9 induces significant upregulation of target genes whereas LLPS-incompetent N-IDR$_{FS}$/A9 shows no or significantly decreased effect. Data are presented as mean $\pm$ $\pm$SD of three replicate experiments. ***$P < 0.001$; ****$P < 0.0001$; n.s., not significant.

d, Heatmap illustrating relative expression of the 374 genes that show significant upregulation post-transduction of F-IDRWT/A9, compared to empty vector (EV) and its IDR-mutant form (F-IDRYS/A9), in 293 cells.

e, Venn diagrams show overlap of the significantly downregulated genes identified post-transduction of the indicated construct into mouse HPSCs.

f, Gene Set Enrichment Analysis (GSEA) shows that, compared to N-IDR$_{FS}$/A9, N-IDR$_{WT}$/A9 is positively correlated with the indicated leukemia- or HSPC-related genesets (top) and negatively correlated with the indicated differentiation-related genesets (bottom).

g, Venn diagrams show overlap of the significantly upregulated (left) or downregulated (right) genes identified post-transduction of the indicated construct into mouse HPSCs.
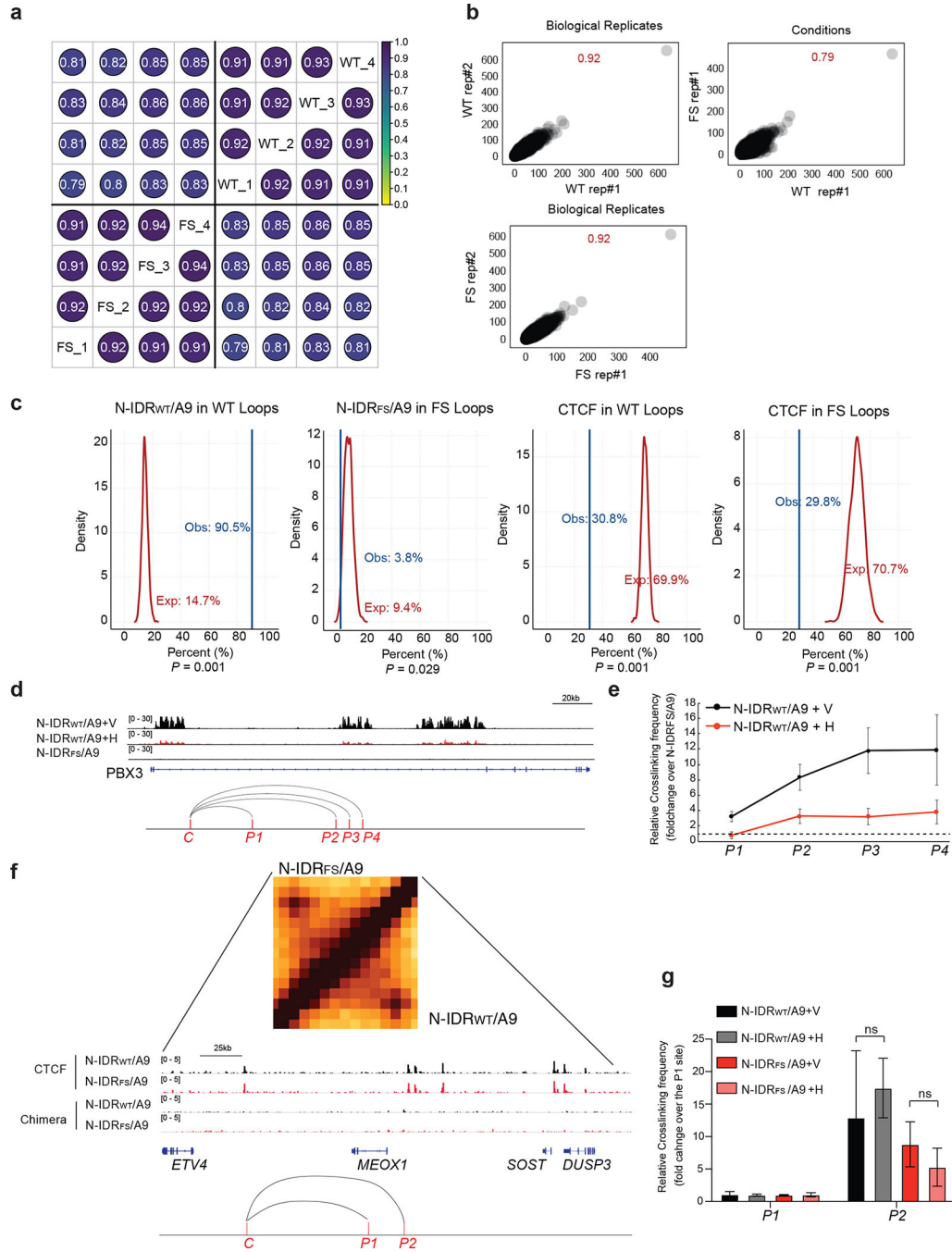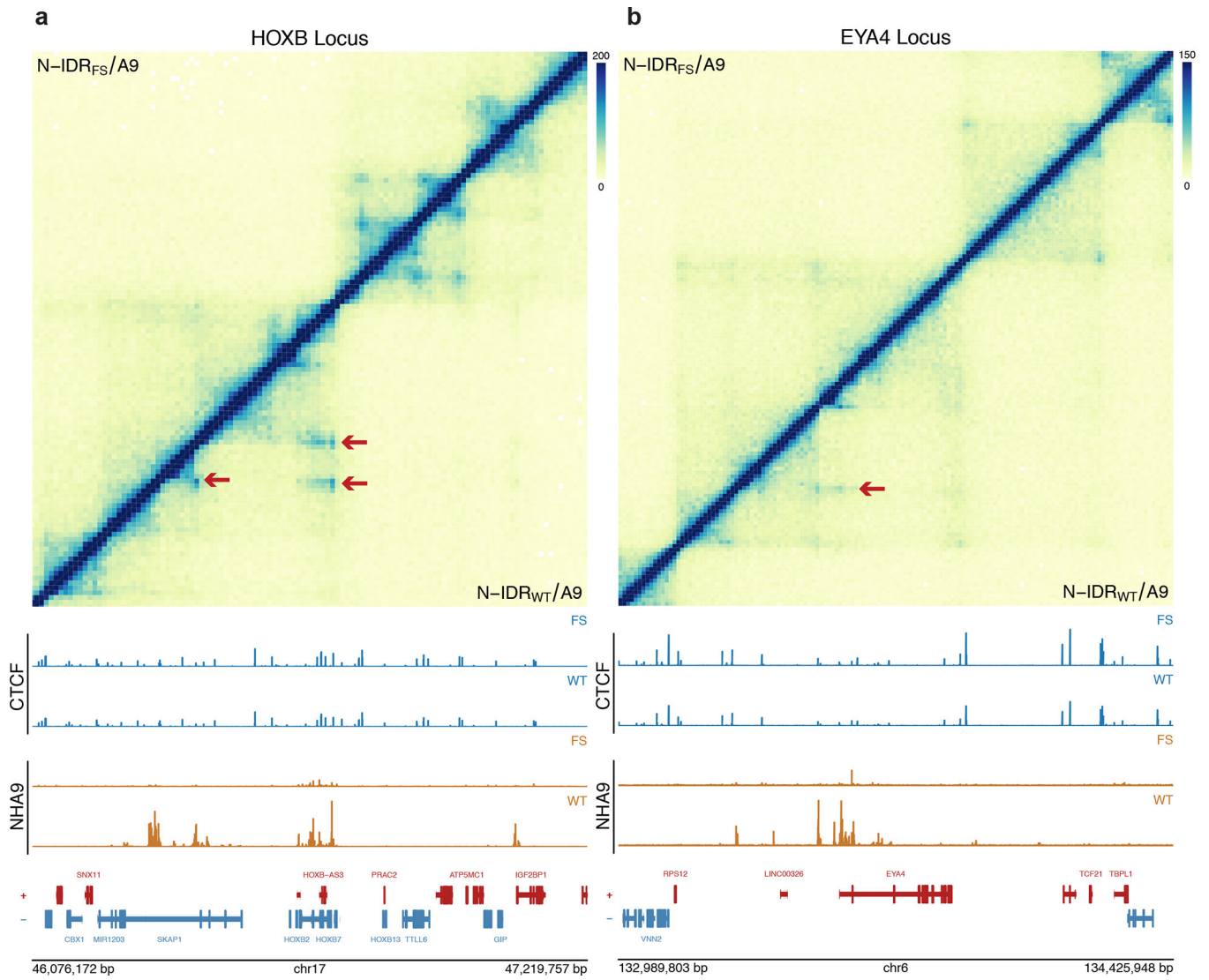
Extended Data Fig 9|. Hi-C mapping reveals that a phase-separation-competent IDR harbored within NUP98-HOXA9 is required for inducing formation of CTCF-independent chromatin loops at leukemia-relevant gene loci.

a, Matrix of Pearson correlation coefficients of loop counts among and between biological replicates of N-IDR$_{WT}$/A9 (WT; n=4 replicates) or N-IDR$_{FS}$/A9 (FS; n=4 replicates) conditions. Condition is denoted along the diagonal as WT or FS, followed by numbers indicating biological replicate for that condition.

b, Example correlation plots of loop counts between biological replicates and conditions.

c, All loops were partitioned into either WT or FS-specific loops and split into separate loop anchors. Loop anchors were then intersected with ChIP-seq peaks for N-IDR/A9 or CTCF. The percentage of observed (Obs.) overlaps for each feature is shown as a vertical blue line. The red line shows the expected (Exp.) distribution of overlaps as determined by randomly sampling loop anchors and calculating each feature's overlap 1000 times. P-values were determined by summing the number of expected values greater than (or less than if the observed value was less than the mean) the observed value for that feature.

d-g, 3C-qPCR assays measuring the change in crosslinking frequency of either an N-IDR$_{WT}$/A9-specific loop at *PBX3* locus (**d-e**) or a CTCF-dependent loop (**f-g**; at Chr17 [41604677 – 41883642]) after treatment of 293 stable cells with 10% of 1,6-hexanediol for one minute (+H) relative to mock (+V). The IGV view panels at **d,f** show the indicated ChIP-seq signals, with positions of the used 3C-PCR primers labeled under IGV tracks. PCR was performed using the same constant forward primer (C) paired with a differently numbered reverse primer (P1 to P4) at each locus tested. Panels **e,g** are plotted with signals of 3C-qPCR measuring the relative crosslinking frequency at *PBX3* (**d-e**) or a Chr17 locus with CTCF loop (**f-g**) before (V) and after (H) treatment with 1,6-hexanediol. Signals in panel **e** are normalized to those of the N-IDR$_{FS}$/A9-expressing cells (n =3 replicated experiments). ns, no significant.
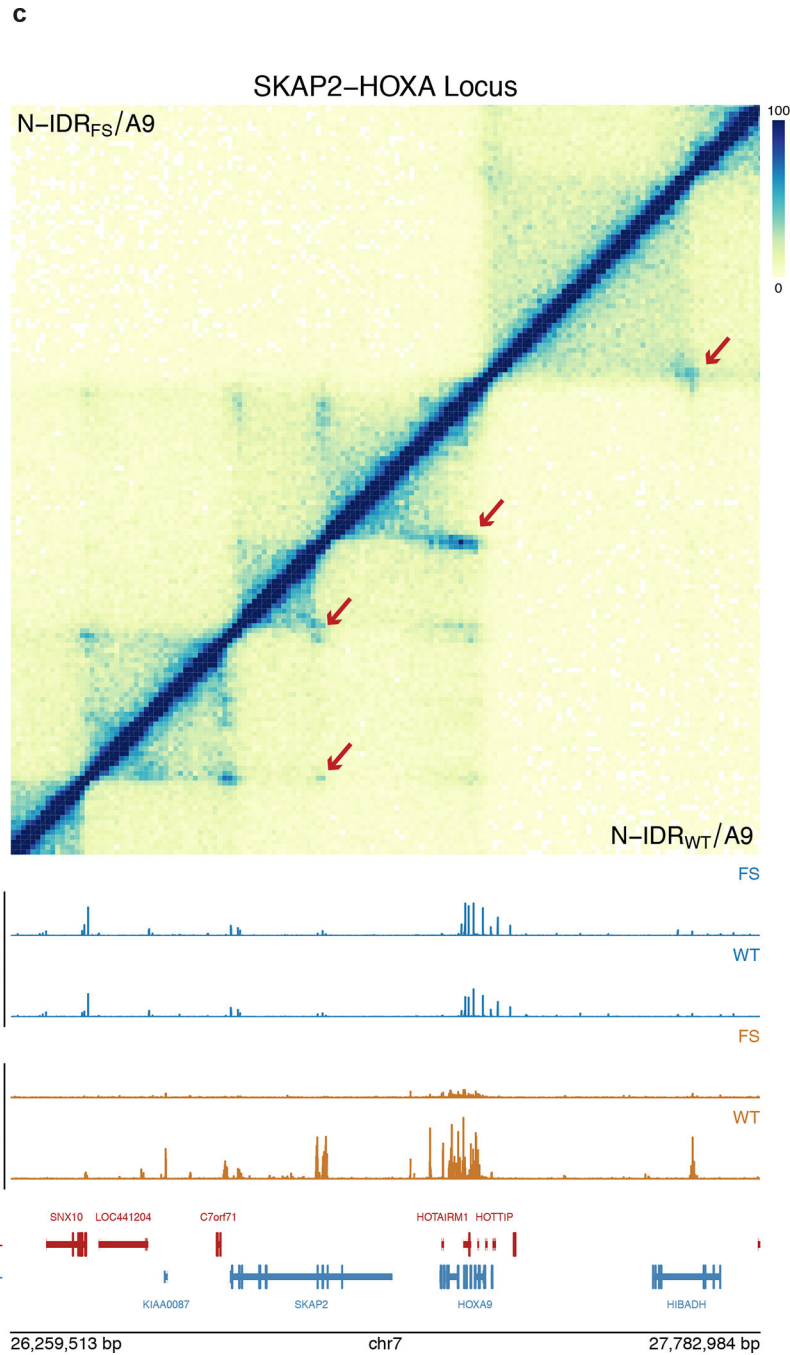
Extended Data Fig 10|. Hi-C mapping reveals chromatin loops specific to cells with the LLPS-competent NUP98-HOXA9, compared to the LLPS-competent mutant, at leukemia-relevant gene loci.

Views for Hi-C mapping, RNA-seq, and ChIP-seq for CTCF, N-IDR/A9, and H3K27ac at the *HOXB* (**a**), *EYA4* (**b**), and *SKAP2-HOXA* loci (**c**) in 293 stable cells expressing either N-IDR$_{WT}$/A9 (WT) or N-IDR$_{FS}$/A9 (FS). Hi-C mapping views (top panels) show results from the N-IDR$_{WT}$/A9 or N-IDR$_{FS}$/A9 expressing cells (lower and upper diagonal, respectively). Corresponding ChIP-seq and gene tracks are shown below each Hi-C plot. N-IDR$_{WT}$/A9 loops are indicated with red arrows.
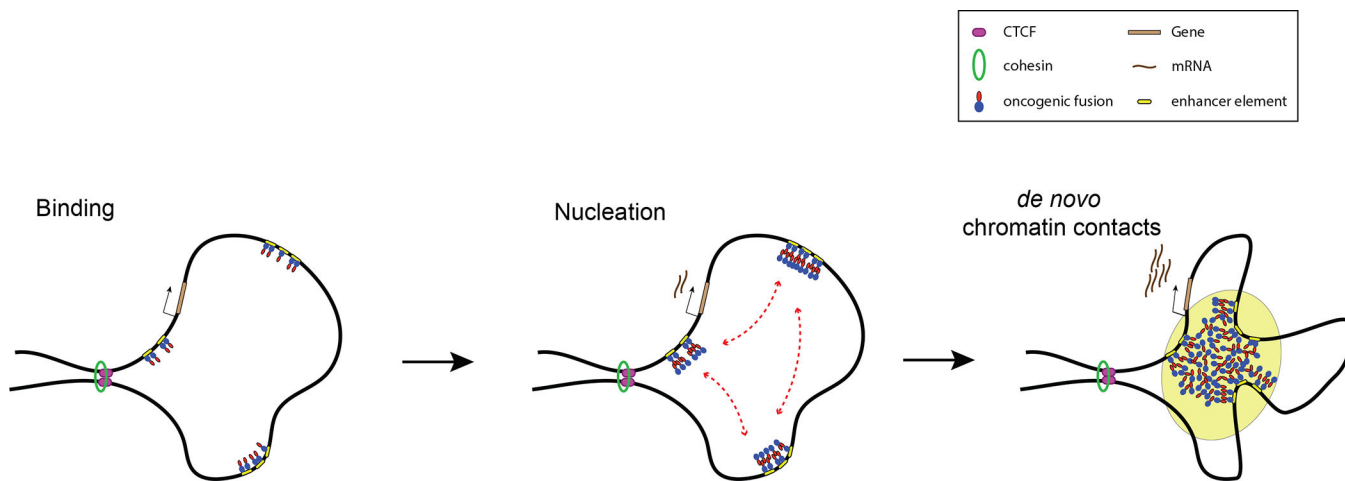
Extended Data Fig 11.
A model illustrating a critical requirement of LLPS-competent IDR harbored within NUP98-HOXA9 for enhancing chimeric TF binding to genomic targets and for promoting long-distance chromatin looping between leukemogenic gene promoter/enhancers, thereby inducing an oncogenic gene-expression program and leukemic development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

Reference

1. Gough SM, Slape CI & Aplan PD NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. Blood 118, 6247–6257, doi:10.1182/blood-2011-07-328880 (2011). [PubMed: 21948299]

2. Mendes A & Fahrenkrog B NUP214 in Leukemia: It's More than Transport. Cells 8, doi:10.3390/cells8010076 (2019).

3. Murray DT et al. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. Cell 171, 615–627.e616, doi:10.1016/j.cell.2017.08.048 (2017). [PubMed: 28942918]

4. Alberti S & Hyman AA Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. Nat Rev Mol Cell Biol 22, 196–213, doi:10.1038/s41580-020-00326-6 (2021). [PubMed: 33510441]

5. Boija A, Klein IA & Young RA Biomolecular Condensates and Cancer. Cancer cell 39, 174–192, doi:10.1016/j.ccell.2020.12.003 (2021). [PubMed: 33417833]

6. Wan L et al. Impaired cell fate through gain-of-function mutations in a chromatin reader. Nature 577, 121–126, doi:10.1038/s41586-019-1842-7 (2020). [PubMed: 31853060]

7. Kovar H, Jekyll Dr. and Hyde Mr.: The Two Faces of the FUS/EWS/TAF15 Protein Family. Sarcoma 2011, doi:10.1155/2011/837474 (2011).

8. Sabari BR et al. Coactivator condensation at super-enhancers links phase separation and gene control. Science 361, eaar3958, doi:10.1126/science.aar3958 (2018). [PubMed: 29930091]

9. Nair SJ et al. Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. Nature structural & molecular biology 26, 193–203, doi:10.1038/s41594-019-0190-5 (2019).

10. Chong S et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science 361, eaar2555, doi:10.1126/science.aar2555 (2018). [PubMed: 29930090]

11. Wang GG et al. Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. Nature 459, 847–851, doi:10.1038/nature08036 (2009). [PubMed: 19430464]

12. Jankovic D et al. Leukemogenic mechanisms and targets of a NUP98/HHEX fusion in acute myeloid leukemia. Blood 111, 5672–5682, doi:10.1182/blood-2007-09-108175 (2008). [PubMed: 18388181]

13. Pak Chi W. et al. Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. Molecular cell 63, 72–85, doi:10.1016/j.molcel.2016.05.042 (2016). [PubMed: 27392146]

14. LaRonde-LeBlanc NA & Wolberger C Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. Genes & development 17, 2060–2072, doi:10.1101/gad.1103303 (2003). [PubMed: 12923056]

15. Calvo KR, Sykes DB, Pasillas M & Kamps MP Hoxa9 immortalizes a granulocyte-macrophage colony-stimulating factor-dependent promyelocyte capable of biphenotypic differentiation to neutrophils or macrophages, independent of enforced meis expression. Molecular and cellular biology 20, 3274–3285 (2000). [PubMed: 10757811]

16. Frey S, Richter RP & Görlich D FG-Rich Repeats of Nuclear Pore Proteins Form a Three-Dimensional Meshwork with Hydrogel-Like Properties. Science 314, 815–817, doi:10.1126/science.1132516 (2006). [PubMed: 17082456]

17. Kasper LH et al. CREB Binding Protein Interacts with Nucleoporin-Specific FG Repeats That Activate Transcription and Mediate NUP98-HOXA9 Oncogenicity. Molecular and cellular biology 19, 764–776, doi:10.1128/mcb.19.1.764 (1999). [PubMed: 9858599]

18. Xu H et al. NUP98 Fusion Proteins Interact with the NSL and MLL1 Complexes to Drive Leukemogenesis. Cancer cell 30, 863–878, doi:10.1016/j.ccell.2016.10.019 (2016). [PubMed: 27889185]

19. Kroon E, Thorsteinsdottir U, Mayotte N, Nakamura T & Sauvageau G NUP98-HOXA9 expression in hemopoietic stem cells induces chronic and acute myeloid leukemias in mice. The EMBO journal 20, 350–361 (2001). [PubMed: 11157742]

20. Wang J et al. A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. Cell 174, 688–699.e616, doi:10.1016/j.cell.2018.06.006 (2018). [PubMed: 29961577]

21. Qamar S et al. FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. Cell 173, 720–734.e715, doi:10.1016/j.cell.2018.03.056 (2018). [PubMed: 29677515]

22. Hansen AS et al. Robust model-based analysis of single-particle tracking experiments with Spot-On. eLife 7, e33125, doi:10.7554/eLife.33125 (2018). [PubMed: 29300163]

23. Strom AR et al. Phase separation drives heterochromatin domain formation. Nature 547, 241–245, doi:10.1038/nature22989 (2017). [PubMed: 28636597]

24. Wang L et al. Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism. Molecular cell 76, 646–659.e646, doi:10.1016/j.molcel.2019.08.019 (2019). [PubMed: 31543422]

25. Gibson BA et al. Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell 179, 470–484.e421, doi:10.1016/j.cell.2019.08.037 (2019). [PubMed: 31543265]

26. Shin Y et al. Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. Cell 175, 1481–1491.e1413, doi:10.1016/j.cell.2018.10.057 (2018). [PubMed: 30500535]

27. Calvo KR, Sykes DB, Pasillas MP & Kamps MP 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 (2002). [PubMed: 12082612]

28. Fahrenkrog B et al. Expression of Leukemia-Associated Nup98 Fusion Proteins Generates an Aberrant Nuclear Envelope Phenotype. PLoS ONE 11, e0152321, doi:10.1371/journal.pone.0152321 (2016). [PubMed: 27031510]

29. Yu M et al. A resource for cell line authentication, annotation and quality control. Nature 520, 307–311, doi:10.1038/nature14397 (2015). [PubMed: 25877200]

30. Xu B et al. Selective inhibition of EZH2 and EZH1 enzymatic activity by a small molecule suppresses MLL-rearranged leukemia. Blood 125, 346–357, doi:10.1182/blood-2014-06-581082 (2015). [PubMed: 25395428]

31. Cai L et al. An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. Molecular cell 49, 571–582, doi:10.1016/j.molcel.2012.11.026 (2013). [PubMed: 23273982]

32. Wang GG, Cai L, Pasillas MP & Kamps MP NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. Nature cell biology 9, 804–812 (2007). [PubMed: 17589499]

33. Stauffer W, Sheng H & Lim HN EzColocalization: An ImageJ plugin for visualizing and measuring colocalization in cells and organisms. Scientific Reports 8, 15764, doi:10.1038/s41598-018-33592-8 (2018). [PubMed: 30361629]

34. Wang GG et al. Quantitative production of macrophages or neutrophils ex vivo using conditional Hoxb8. Nature methods 3, 287–293 (2006). [PubMed: 16554834]

35. Wang GG, Pasillas MP & Kamps MP Meis1 programs transcription of FLT3 and cancer stem cell character, using a mechanism that requires interaction with Pbx and a novel function of the Meis1 C-terminus. Blood 106, 254–264 (2005). [PubMed: 15755900]

36. Lu R et al. Epigenetic Perturbations by Arg882-Mutated DNMT3A Potentiate Aberrant Stem Cell Gene-Expression Program and Acute Leukemia Development. Cancer cell 30, 92–107, doi:10.1016/j.ccell.2016.05.008 (2016). [PubMed: 27344947]

37. Roux KJ, Kim DI & Burke B BioID: a screen for protein-protein interactions. Curr Protoc Protein Sci 74, 19 23 11–19 23 14, doi:10.1002/0471140864.ps1923s74 (2013). [PubMed: 24510646]

38. Roux KJ, Kim DI, Burke B & May DG BioID: A Screen for Protein-Protein Interactions. Curr Protoc Protein Sci 91, 19 23 11–19 23 15, doi:10.1002/cpps.51 (2018). [PubMed: 29516480]

39. Li J et al. ZMYND11-MBTD1 induces leukemogenesis through hijacking NuA4/TIP60 acetyltransferase complex and a PWWP-mediated chromatin association mechanism. Nat Commun 12, 1045, doi:10.1038/s41467-021-21357-3 (2021). [PubMed: 33594072]

40. Nesvizhskii AI, Keller A, Kolker E & Aebersold R A Statistical Model for Identifying Proteins by Tandem Mass Spectrometry. Analytical Chemistry 75, 4646–4658, doi:10.1021/ac0341261 (2003). [PubMed: 14632076]

41. Cai L et al. ZFX Mediates Non-canonical Oncogenic Functions of the Androgen Receptor Splice Variant 7 in Castrate-Resistant Prostate Cancer. Molecular cell 72, 341–354 e346, doi:10.1016/j.molcel.2018.08.029 (2018). [PubMed: 30270106]

42. Egan B et al. An Alternative Approach to ChIP-Seq Normalization Enables Detection of Genome-Wide Changes in Histone H3 Lysine 27 Trimethylation upon EZH2 Inhibition. PLoS ONE 11, e0166438, doi:10.1371/journal.pone.0166438 (2016). [PubMed: 27875550]

43. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21, doi:10.1093/bioinformatics/bts635 (2012). [PubMed: 23104886]

44. Zhang Y et al. Model-based Analysis of ChIP-Seq (MACS). Genome biology 9, R137, doi:10.1186/gb-2008-9-9-r137 (2008). [PubMed: 18798982]

45. Ramírez F et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic acids research 44, W160–W165, doi:10.1093/nar/gkw257 (2016). [PubMed: 27079975]

46. Lovén J et al. Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. Cell 153, 320–334, doi:10.1016/j.cell.2013.03.036 (2013). [PubMed: 23582323]

47. Ren Z et al. PHF19 promotes multiple myeloma tumorigenicity through PRC2 activation and broad H3K27me3 domain formation. Blood 134, 1176–1189, doi:10.1182/blood.2019000578 (2019). [PubMed: 31383640]

48. Wang K et al. MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. Nucleic acids research 38, e178, doi:10.1093/nar/gkq622 (2010). [PubMed: 20802226]

49. Li B & Dewey CN RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323, doi:10.1186/1471-2105-12-323 (2011). [PubMed: 21816040]

50. Anders S & Huber W Differential expression analysis for sequence count data. Genome Biol 11, R106, doi:10.1186/gb-2010-11-10-r106 (2010). [PubMed: 20979621]

51. Subramanian A et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences 102, 15545–15550, doi:10.1073/pnas.0506580102 (2005).

52. Chen B-C et al. Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. Science 346, 1257998, doi:10.1126/science.1257998 (2014). [PubMed: 25342811]

53. Grimm JB et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. Nature Methods 12, 244–250, doi:10.1038/nmeth.3256 (2015). [PubMed: 25599551]

54. Tinevez J-Y et al. TrackMate: An open and extensible platform for single-particle tracking. Methods 115, 80–90, doi:10.1016/j.ymeth.2016.09.016 (2017). [PubMed: 27713081]

55. Virtanen P et al. SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python. arXiv e-prints (2019). <https://ui.adsabs.harvard.edu/abs/2019arXiv190710121V>.

56. Rao Suhas S. P. et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159, 1665–1680, doi:10.1016/j.cell.2014.11.021 (2014). [PubMed: 25497547]

57. Durand NC et al. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Systems 3, 95–98, doi:10.1016/j.cels.2016.07.002 (2016). [PubMed: 27467249]

58. Knight PA & Ruiz D A fast algorithm for matrix balancing. IMA Journal of Numerical Analysis 33, 1029–1047, doi:10.1093/imanum/drs019 (2012).

59. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15, 550, doi:10.1186/s13059-014-0550-8 (2014). [PubMed: 25516281]

60. Ren Y, Seo H-S, Blobel G & Hoelz A Structural and functional analysis of the interaction between the nucleoporin Nup98 and the mRNA export factor Rae1. Proceedings of the National Academy of Sciences 107, 10406–10411, doi:10.1073/pnas.1005389107 (2010).

61. Yung E et al. Delineating domains and functions of NUP98 contributing to the leukemogenic activity of NUP98-HOX fusions. Leukemia research 35, 545–550, doi:10.1016/j.leukres.2010.10.006 (2011). [PubMed: 21130494]
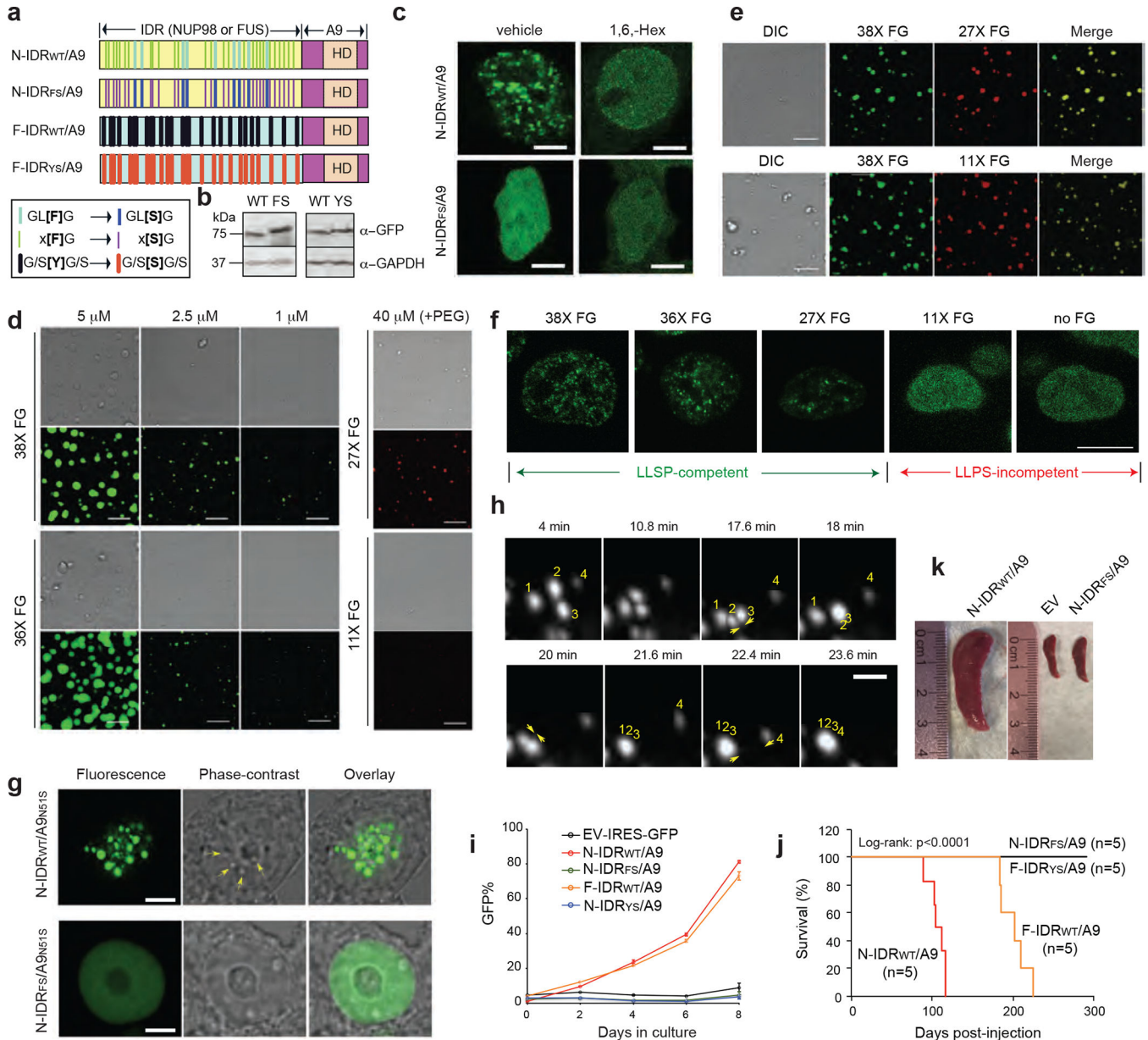
Figure 1. IDRs within chimeric TF oncoproteins establish phase-separated assemblies, inducing leukemogenesis.

a, Scheme for N-IDR/A9 and F-IDR/A9 chimera, with the F→S and Y→S mutations introduced to the NUP98 and FUS IDRs, respectively, shown in box. HD, homeodomain.

b-c, Immunoblotting (**b**) and live-cell fluorescence (**c**) for GFP-tagged chimera carrying the WT or mutant IDR in 293 cells. 1,6-hex, 1,6-hexanediol. Scale bar, 10 μm.

d-e, Differential interference contrast (DIC) and concurrent fluorescence imaging (bottom) of N-IDR recombinant proteins harboring the varying number of FG-repeats, prepared at the indicated concentration with either single protein species (**d**) or a mixture of the two (**e**). PEG, polyethylene glycol-3350. Scale bar, 10 μm.

f, Live-cell imaging of GFP-tagged N-IDR/A9 with the indicated number of FG-repeats. Scale bar, 10 μm.

g, Live-cell imaging (GFP) and concurrent phase-contrast imaging for N51S-mutated GFP-NUP98-HOXA9 with either WT (top) or F→S-mutated IDR (bottom).

Arrows indicate droplet-like structures. Scale bar, 10μm.

h, Coalescence of GFP-NUP98-HOXA9 condensates (N51S-mutated). Scale bar, 2 μm.

i, Proliferation of murine HSPCs transduced with empty vector (EV) or the indicated chimera (n=3 independent biological replicates; data presented as mean ± SD).

j, Kaplan-Meier survival plot of mice post-transplantation of HSPCs transduced with the indicated chimera (n=5 mice per group).

k, Splenomegaly associated with N-IDR$_{WT}$/A9-induced leukemias, three months post-transplantation of infected HSPCs into mice.

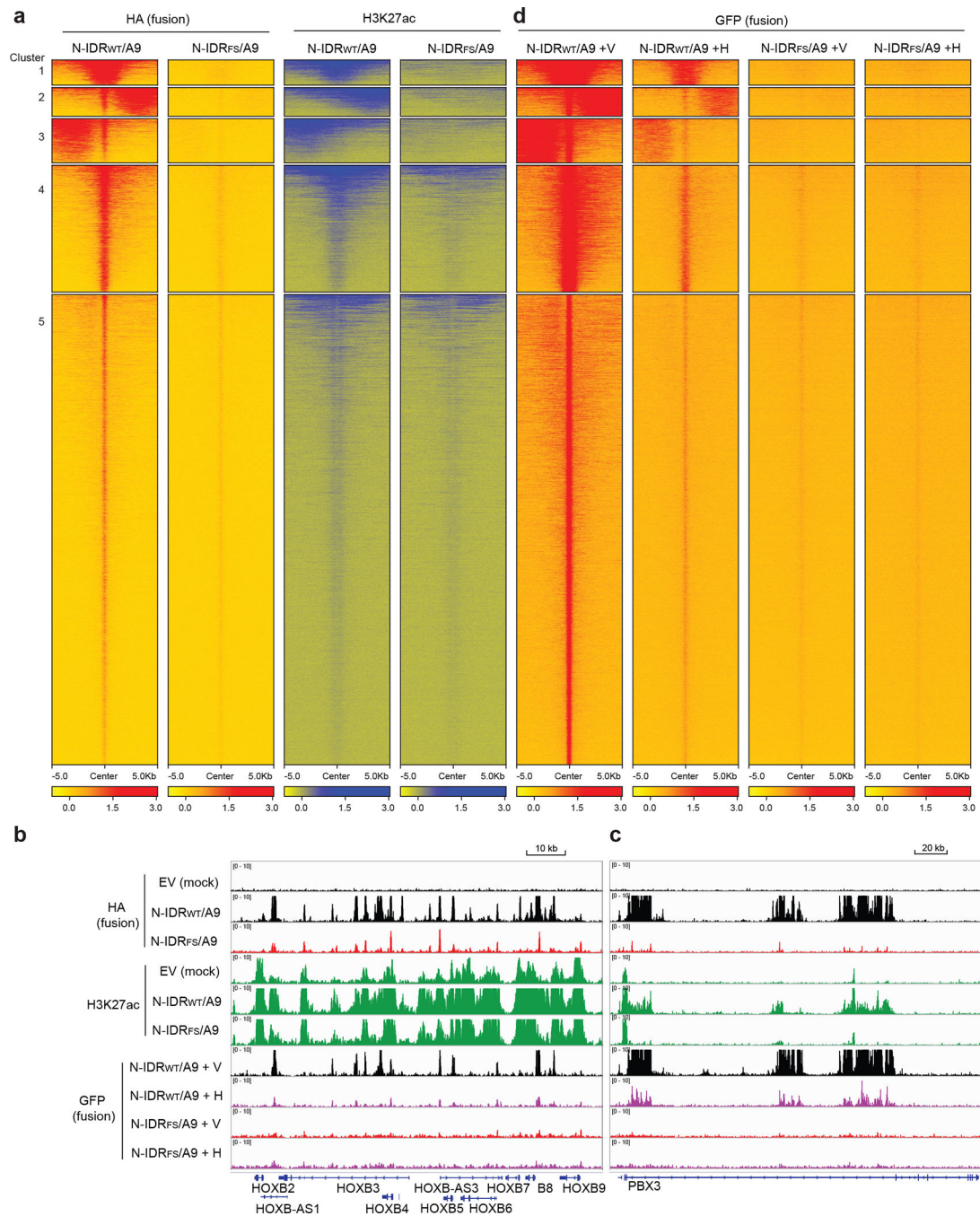**Figure 2. Phase separation dramatically enhances chromatin binding of NUP98-HOXA9, featured with broad, 'super-enhancer'-like genomic occupancy.**

a,**d,** Heatmaps for k-means clustering of ChIP-seq signals in 293 cells expressing HA-tagged (**a**; input-normalized) or GFP-tagged (**a**; spikein control normalized) N-IDR/A9 with either WT or F→S mutated IDRs. Cells in (**d**) were treated with 10% of 1,6-hexanediol (+H), compared to vehicle (+V), for one minute. Each row represents a peak called for WT samples (first column) (±5Kb from peak center).

b,c, IGV tracks of the indicated ChIP-seq signals at *HOXB* (**b**) and *PBX3* (**c**) in 293 cells. EV-transduced cells serve as a ChIP control.
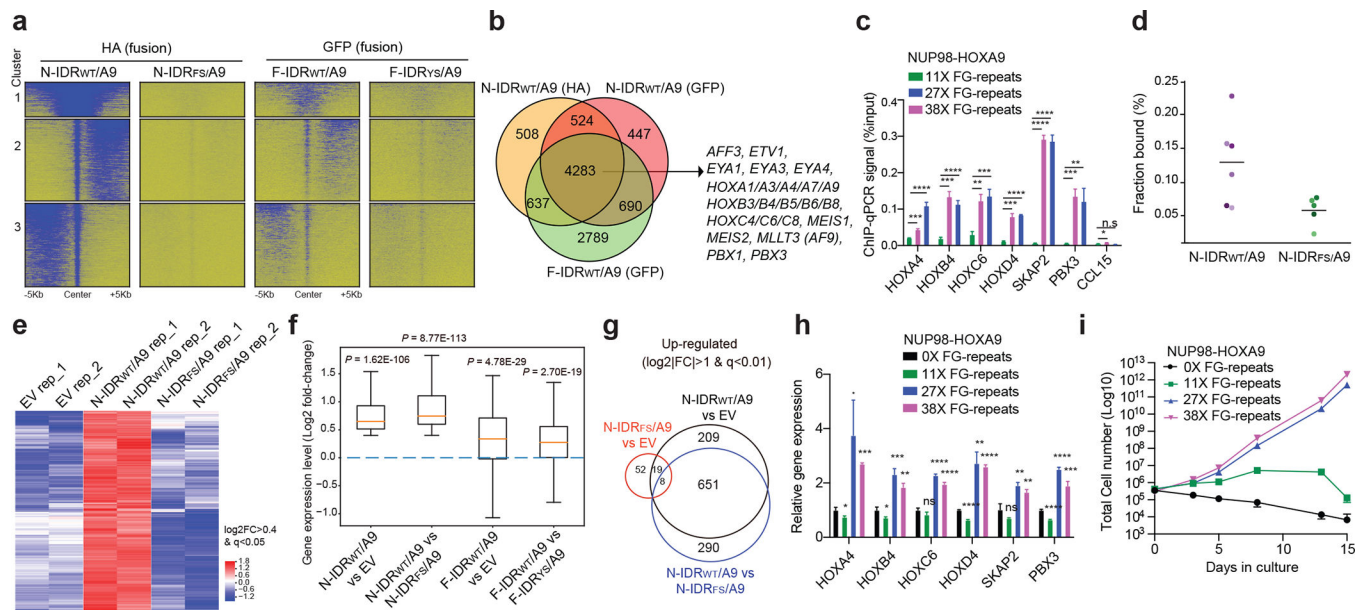
Figure 3. Fusing an unrelated LLPS-competent IDR of FUS with HOXA9's HD (F-IDR/A9), as well as altering the FG-repeats valency in NUP98-HOXA9, demonstrates a role for IDR and LLPS in promoting target oncogene activation and cancerous transformation.

a, ChIP-seq signal heatmaps showing N-IDR/A9 (HA-tagged; left) and F-IDR/A9 (GFP-tagged; right), either WT or IDR-mutated (FS or YS), in 293 cells. See also Extended Data Fig 6a.

b, Venn diagram using direct targets of N-IDR$_{WT}$/A9 or F-IDR$_{WT}$/A9 in 293 cells, with a battery of leukemia-related oncogenes highlighted.

c, ChIP-qPCR for binding of GFP-tagged N-IDR with the indicated number of FG-repeats at examined loci in 293 cells (n=3 independent samples; data presented as mean ± S.D.). CCL15 acts as a control. Statistics was performed with two-sided t-test. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001.

d, Single-molecule imaging estimated the fraction of chromatin-bound N-IDR$_{WT}$/A9 and N-IDR$_{FS}$/A9 in 293 stable cells. Presented are values based on two-state kinetic modeling (individual standard deviations <0.0003). Black bar, averaged value.

e, Heatmap of 303 genes upregulated in 293 cells post-transduction of N-IDR$_{WT}$/A9, compared to EV and N-IDR$_{FS}$/A9.

f, Boxplots showing relative expression of 303 N-IDR$_{WT}$/A9-activated genes in **e** among the indicated pairwise comparison of 293 cells. Boxes extend from the first quartile to third quartile values of dataset, with a line showing the median. The whiskers extend from the box edges to show the data range. Statistics was conducted with two-sided t-test.

g, Venn diagram using genes upregulated in mouse HPSCs post-transduction of the indicated construct.

h, RT-qPCR for oncogenes in 293 cells expressing chimera with the indicated number of FG-repeats (n=3 independent samples; data presented as mean ± S.D.). Expression was normalized to the 0×FG-repeat sample. Statistics was performed with two-sided t-test. *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001; ns, not significant.

i, Proliferation of murine HSPCs transduced with N-IDR fusion with the indicated number of FG-repeats (n= 3 independent replicates; data presented as mean ± SD).
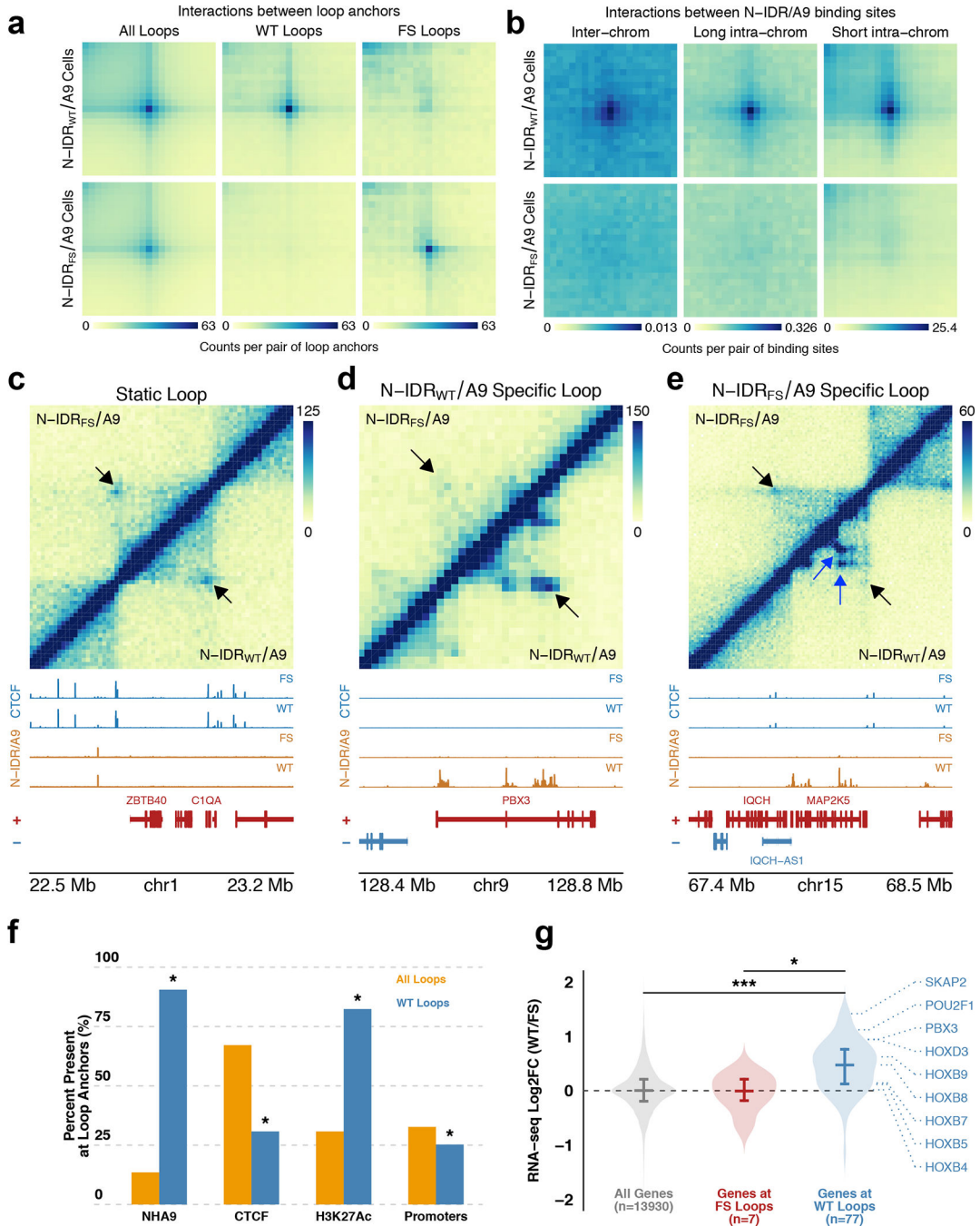
Figure 4. Phase-separation-competent IDRs harbored within NUP98-HOXA9 induce CTCF-independent looping at oncogenes.

a, Aggregate peak analysis (APA) for all loops (n=6,615), WT-specific (n=232) and FS-specific (n=52) loops defined by Hi-C in 293 cells expressing N-IDR$_{WT}$/A9 (top) or N-IDR$_{FS}$/A9 (bottom). Pixel color represents the mean interaction counts per loop, plotted on a common scale.

b, APA plots at 10 Kb resolution for interactions between the 500 strongest N-IDR/A9 binding sites in cells with N-IDR$_{WT}$/A9 (top) or N-IDR$_{FS}$/A9 (bottom). Paired interactions

were categorized as inter-chromosomal (n=95,959), long (>=2Mb) intra-chromosomal (n=6,298), or short (<2Mb) intra-chromosomal (n=574). Pixel color represents the mean interaction counts per pair of loci interrogated. Color scale in each plot is adjusted to the maximum value.

c-e, Non-differential static **(c),** N-IDR$_{WT}$/A9-specific **(d;** "Gained in WT" at PBX3**)** and N-IDR$_{FS}$/A9-specific loop **(e;** "Lost in WT") detected by Hi-C (arrowheads in top panel) with 293 cells expressing N-IDR$_{WT}$/A9 (below diagonal) or N-IDR$_{FS}$/A9 (above diagonal). Bottom panels show CTCF (blue) and N-IDR/A9 (orange) ChIP-seq signals (gene tracks shown below) in same cells. Note that CTCF but not N-IDR/A9 is present at the loop anchors.

f, Percentage of the indicated feature present either at all loops or WT-specific loops. Significance was determined by a permutation test as described in the Methods section. *$P <$ 0.001.

g, Relative expression of genes associated with WT-specific (n=77) and FS-specific loops (n=7) in 293 cells expressing N-IDR$_{WT}$/A9 versus N-IDR$_{FS}$/A9. *BH-adjusted $P<0.05$; *** Benjamini–Hochberg-adjusted $P<0.00001$.