High-precision mapping of nuclear pore-chromatin interactions reveals new principles of genome organization at the nuclear envelope

Swati Tyagi[1], Fei Chen[2], Jialiang Huang[2], Martin W. Hetzer[1,3,*]

[1] The Salk Institute for Biological Studies, La Jolla, CA, USA

[2] State Key Laboratory of Cellular Stress Biology, Innovation Center for Cell Signaling Network, School of Life Sciences, Xiamen University, Xiamen, Fujian 361102, China

[3] Lead contact

*Corresponding author , email: hetzer@salk.edu

SUMMARY

The role of nuclear pore complexes (NPCs) in genome organization remains poorly characterized due to technical limitations in probing genome-wide protein-DNA interactions that are specific to the nuclear envelope. Here, we developed a sensitive method, NPC-DamID, which combines *in vitro* reconstitution of nuclear import and DamID technology. This fixation-free method specifically identifies genomic DNA interactions at the NPCs in intact nuclei. We found that NPCs are preferentially associated with common and hierarchically arranged super-enhancers (SEs) across multiple cells and tissue types. We also uncovered phase-separated condensates at NPCs that compartmentalize and concentrate transcriptional coactivators and structural proteins at SE-regulated genes. Our results support the idea that NPCs are sites of anchoring SE regulatory hubs through their interaction with CTCF and also maintains the conformation of SEs through interaction with structural proteins of SEs like Med1 and PolII.

INTRODUCTION

The three-dimensional (3D) organization of the genome in a mammalian nucleus is critical for the transcriptional output of a cell which in turn is important for cell fate determination and function (Cook and Marenduzzo, 2018). Nuclear landmarks such as the nuclear envelope (NE) play important roles in orchestrating genome architecture and transcription regulation. For example, the interaction between chromatin and the lamina at the NE form hundreds of lamin-associated domains (LADs), stretches of genomic DNA that are primarily heterochromatic regions containing genes with typically low expression (Lochs et al., 2019). While the association between lamina and DNA is well understood, the precise chromatin interaction sites of nuclear pore complexes (NPCs), which collectively occupy approximately 20% of the area of NEs, have not been properly mapped.

NPCs are megadalton protein complexes, composed of 30 different subunits called nucleoporins (Nups), that regulate the traffic of cargoes between nucleus and cytoplasm (Ibarra and Hetzer, 2015). Recent findings have uncovered additional roles of NPCs and mobile Nups in gene regulation and genome organization(Buchwalter et al., 2019; D'Angelo, 2018; Ibarra and Hetzer, 2015; Pascual-Garcia and Capelson, 2021; Raices and D'Angelo, 2018). However, our functional understanding of NPC-genome interactions has been limited in terms of spatial and temporal resolution and restricted to few cell types due to technical challenges in probing the dynamic association of membrane-bound proteins with chromatin through antibody-based methods like ChIP-seq (Aughey and Southall, 2016; Aughey et al., 2019; Gatticchi et al., 2019). The efficiency of ChIP-seq is lower for membrane-embedded proteins like Nups in the structure of NPCs due to the limited accessibility of epitopes of Nups to the antibodies. As an alternative, *in vivo* DamID for Nups has been the method of choice to avoid the use of antibodies for studying NPC-DNA interactions. The *in vivo* DamID-seq method is independent of antibody usage because transgenic cell lines are used in which Dam fused proteins are expressed at low levels (Steensel and Henikoff, 2000). Dam is an enzyme from *E. Coli* which when fused to the protein of interest positions an orthogonal methylation modification to adenines at GATC of interacting genomic DNA(Steensel and

Henikoff, 2000). The methylated DNA is then sequenced to identify the genomic DNA associated with the protein of interest (Southall et al., 2013). In this way, DamID-seq has been successfully applied to many membrane-embedded proteins including Nups (Franks et al., 2016; Ibarra et al., 2016; Jacinto et al., 2015). However, the expression of Dam fused Nups for DamID-seq experiments in cell lines poses difficulties like over-expression artifacts, the toxicity of Dam fused protein, and loss of temporal resolution due to cumulative DNA methylation and the need for a long time of 24-36 h for expression of Dam fused proteins (Aughey and Southall, 2016). Moreover, Nups like Nup153, Nup98, which have been previously targeted to study NPC-DNA association are mobile Nups (i.e., they can bind chromatin at sites that are not associated with the NE) (Capelson et al., 2010; Franks et al., 2017; Ibarra et al., 2016; Jacinto et al., 2015). Hence, it is challenging to distinguish genomic DNA interactions at the NE from the ones in the nucleoplasm in these studies.

To overcome these limitations of conventional DamID-seq experiments, we have developed a method, referred to as NPC-DamID, specifically designed to dissect NPC-genome interactions specifically at the NE. In NPC-DamID, *in vitro* reconstitution of nuclear transport assay (semi-permeabilized cell assay) is combined with DamID-seq to target the Dam fused nuclear transport receptor-Importin beta (Impβ) to NPCs (Figure 1a). In this way, the method remains independent of the use of an antibody and over-expression artifacts of Dam fused protein. Since there is no requirement for the generation of transgenic cells, we were able to apply this method to multiple unfixed cells and tissue types from mice and humans, which helped us understand the general features and mechanisms behind NPCs as genome organizers.

Recently, we identified the unexpected link between Nups with super-enhancers (SEs)(Ibarra et al., 2016). SEs are the cluster of enhancers that are in close genomic proximity with unusually high levels of transcription factors and activators (Pott and Lieb, 2015). SEs regulate the activity of genes important for cell-type specification (Pott and Lieb, 2015). Since individual Nups are also present inside the nucleus able to interact with DNA (Ibarra and Hetzer, 2015) no detailed maps exist that specifically and reliably

identify NPC-genome interactions at the NE. This results in a lack of understanding of the molecular mechanisms behind these associations and their consequences on cell fate determination and genome organization.

In this study, we mapped NPC-SE interactions across multiple cell types and found unique features for SEs specifically located at the nuclear periphery, furthering our understanding of the importance and role of NPCs in the function and organization of NPC-SEs. We found that the NPCs associate with the SEs that are active in multiple cell types (common SEs) compared to other SEs that are not associated with NPCs. The common SEs have been implicated in establishing the early loop contacts during genome refolding during cell division (Ryu et al., 2019). In addition to the relation of the conserved nature of SEs to their function, the structural organization of SEs can also determine their function. About 25-60% SEs in a cell can have hierarchical structure organization, meaning they have both hub and non-hub enhancers (Huang et al., 2018). Hub enhancers are distinctly associated with CTCF and cohesin binding sites and disease-associated genetic variants. Interestingly, we found that the NPC-SE are enriched in hub enhancers which are known to be the primary contributors in functional and structural organizations of SEs (Huang et al., 2018). Next, we sought to determine the mechanism underlying the maintenance of the organization of SEs at NPCs. Through in vivo experiments, we uncovered a critical role of intrinsically disordered regions (IDRs) of Nups in the maintenance of the structure of NPC-associated SEs through phase separation. We found that phase-separated droplets of Nups can be enriched in PolII and Med1 IDRs which have been previously implicated in the maintenance of the structural organization of SEs (Sabari et al., 2018). Taken together, these results indicate the role of NPCs in orchestrating the foundational genome architecture.

RESULTS

NPC-DamID combines *in vitro* nuclear import assay with DamID to probe the genome interactions at NPCs

To overcome the challenges associated with *in vivo* DamID-seq and probe the specific DNA association of NPCs at the nuclear periphery, we have developed a method referred to as NPC-DamID. NPC-DamID uses targeting of recombinantly expressed and purified Dam fused nuclear transport receptor-Importin beta (Impβ) to NPCs in a reconstituted nuclear import assay (Figure 1a, b, Figure S1 a-,d). Since Impβ has a high affinity to NPCs, we predicted that Dam fused Impβ (Dam-Impβ) stably binds to NPCs embedded in the NE when the plasma membrane of cells is permeabilized with low concentration digitonin (Figure 1a) and should result in specific methylation of DNA in the vicinity of NPCs.

Before reconstituting the nuclear import assay for NPC-DamID, we checked the *in vitro* methylation activity of recombinantly expressed and purified Dam and Dam-Impβ. We incubated unmethylated lambda DNA with the proteins at room temperature (RT) for 30 min and then digested the DNA with DpnII. We found that purified Dam and Dam-Impβ can efficiently methylate lambda DNA in 30 min, similar to commercially available Dam (NEB Dam) (Figure 1c, Figure S1e).

Next, we confirmed the proper localization of Dam-Impβ at the nuclear periphery through imaging of the fluorescently labeled protein. We observed nuclear rim signal for Dam-Impβ labeled with Alexa 647 similar to recombinantly expressed purified and fluorescently labeled wt Impβ after performing NPC-DamID assay (Figure 1d, Figure S1f). However, in presence of wheat germ agglutin (WGA), a small molecule that blocks NPC, Dam-Impβ was not able to bind to NPCs and thus serving as a negative control (Figure 1d). We also confirmed the active import activity of Dam-Impβ, which is dependent on RanGTP through localization of NLS bearing cargo inside the nucleus (Figure S1g).

Next, we wanted to test if Dam-Impβ can specifically label genomic DNA in the vicinity of the NPCs and if localized DNA methylation can be visualized microscopically. To this end, we exploited a known fragment of the Dpn1 enzyme, which binds to GA^{me}TC residue without cleaving the DNA (Mierzejewska et al., 2014)(Figure 1e, Figure S1h-l). Previously, a similar fragment has been used as a tracer for *in vivo* DamID experiments for imaging of lamina associated domains (LADs), which are large megabase pair long stretches of DNA sequences and thus relatively easy to visualize (Kind et al., 2013). We expressed and purified the recombinant fragment (Dpn8) fused with either flag or GFP tag (Figure S1h, i). We added Flag-Dpn8 to the semi-permeabilized cells after washing off the reagents from the NPC-DamID assay. Subsequently, through confocal imaging, we observed that Flag-Dpn8 labels the entire volume of the nucleus when cells were incubated with Dam alone, indicating that Dam diffuses freely through NPCs during the assay and places methylation mark across the entire genome (Figure S2l). When Flag-Dpn8 was used to label cells treated with Dam-Impβ, the only nuclear peripheral signal was observed, similar to the cell line expressing lamin B-receptor fused to Dam (Dam-LBR, Figure S2l). This shows the selective methylation of genomic DNA associated with NPCs through Dam-Impβ. To determine the working distance range of the NPC-DamID assay, we performed dual-color super-resolution microscopy (dSTORM). We were able to visualize individual NPCs labeled with Nup153 antibody-Alexa 488 and associated genomic DNA labeled with Dpn8-Alexa647 which confirms very selective targeting of nuclear peripheral genomic DNA through the NPC-DamID method (Figure 1f). We observed that all the NPCs (~ 92%) uniformly had a signal for associated genomic DNA, and methylation pattern ranges from 200-600 nm from the nuclear periphery. These results show that NPC-DamID provides a method for high-resolution imaging of NPC-associated genomic DNA at the nuclear periphery.

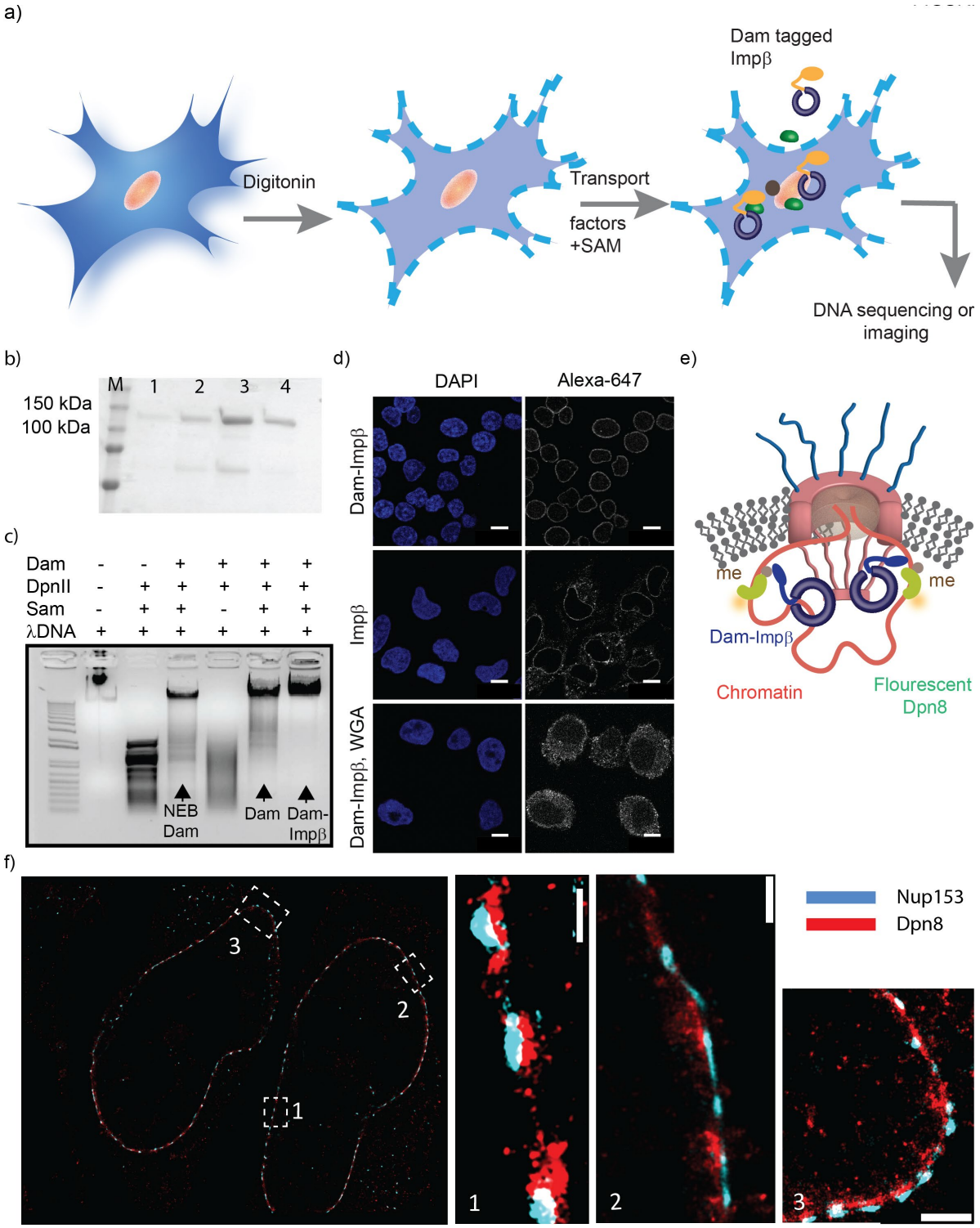

Figure 1. Development of NPC-DamID, a new method to probe NPC-genome interactions a) Schematic representation of *in vitro* DamID. The cells are semi-

permeabilized with a low concentration of digitonin (0.001-0.0025% w/v) so that the plasma membrane gets preferentially permeabilized, leaving the nuclear membrane intact. Dam fused nuclear import factor, Dam-Impβ is supplied to the cells along with the other transport factors and SAM (s-adenosyl methionine) for efficient targeting of Dam-Impβ to the NPCs and subsequent methylation of NPC associated chromatin. In the next steps, cells were lysed to extract genomic DNA to prepare the DamID-seq libraries. b) Coomassie gel of recombinantly expressed and purified Dam-Impβ (molecular weight 132 kDa) through size exclusion chromatography (lane 1-4), lane M is protein marker. c) EtBr-stained agarose gel showing the enzymatic activity of Dam-Impβ. The unmethylated lambda DNA (1 µg) was incubated with 2 µM of Dam, Dam-Impβ or 1 unit of commercial Dam (NEB Dam) in the presence or absence of 7.5 µM SAM in nuclear transport buffer at room temperature. After 20 min, DpnII was added to digest unmethylated DNA, detecting lambda DNA protection through methylation by Dam-Impβ. d) Targeting of fluorescently labeled Dam-Impβ and wt Impβ labeled with Alexa 647 to NPC in HeLa cells after semi-permeabilization and NPC-DamID assay. WGA was added to the assay to block the NPCs. The lack of nuclear rim signal for labeled Dam-Impβ shows the intact nuclear envelope and specific localization of Dam-Impβ at the periphery through NPCs. Scale bar is 10 µm. e) Schematic representation methylated DNA imaging in the vicinity of NPC through Dam-Impβ, using the fluorescently labeled Dpn1 fragment, Dpn8. f) dSTORM images of IMR90 cells treated with Dpn8 after NPC-DamID assay. The cells show co-staining of Dpn8 (staining for NPC interacting genomic DNA) and Nup153 (marking NPCs). Scale bar is 500 nm.

NPC-DamID sequencing allows identification of the genomic DNA associated with NPCs and eliminates the nucleoplasmic signal from mobile Nups

To understand the NPC-mediated genome organization at the NE, we must determine the specific genomic DNA sequences associated with the NPCs. Similar to *in vivo* DamID, we enrich for GA^{me}TC methylated DNA fragments from the genomic DNA to prepare libraries for DNA sequencing after performing NPC-DamID assay (see Methods section). We used Dam alone for background correction of the sequencing data for Dam-Impβ. For this purpose, we have determined the active enzymatic units for purified Dam and Dam-Impβ through an *in vitro* lambda DNA protection assay against the DpnII enzyme (see Methods). One enzyme unit is defined as the concentration of Dam or Dam-Impβ required to completely methylate 1 µg of lambda DNA at room temperature (RT) in 30 minutes, protecting the DNA from DpnII mediated cleavage (Figure S2a-b). In the NPC-DamID assay, the same active enzyme units for Dam and Dam-Impβ were used to ensure unbiased normalization of the data between

experiments. The kinetic assay for both Dam and Dam-Impβ revealed that 5 units of enzyme protected 2 μg of lambda DNA in 20 min (Figure S2c-d). Thus, we used 5 units of Dam or Dam-Impβ in the nuclear transport assay to target proteins in semi-permeabilized cells for 20 minutes at RT. Our results from sequencing experiments show clear peaks for Dam-Impβ, which are significantly enriched over the Dam alone and wt Impβ (Figure 2a). Dam-Impβ peaks are also found to be significantly non-overlapping with the LADs, which indicates the specificity of the assay (Figure 2a, Figure S2e). To further assess the efficiency of the assay, we added the NPC blocking small molecule WGA and found no detectable peaks over the background control of Dam alone (Figure 2a, Figure S2f). This shows that the semi-permeabilization left the NE intact, and hence the peaks for Dam-Impβ are selective for Dam-Impβ localized at NPCs.

Impβ is known to transiently move from NPCs at the nucleoplasmic side while releasing cargo, a step that is dependent on the RanGTP binding property of Impβ (Lowe et al., 2015). This raises the possibility that Dam-Impβ might be released from the NPCs during the reaction, and methylate chromatin regions in the nucleoplasm. To reduce the impact of such diffusive molecules of Dam-Impβ on the sequencing results, we expressed a dominant mutant of Impβ fused to Dam, Dam-ΔN44 Impβ, which stably binds to the nuclear face of NPCs as compared to wt Impβ since it has a low affinity to RanGTP and was shown to block nuclear import activity (Kutay et al., 1997). When sequencing data were compared for Dam-Impβ and Dam-ΔN44 Impβ, we found similar peaks (78% overlap) in both data sets (Figure 2b, Figure S2f). We observed that the presence of RanGTP also does not significantly influence either Dam-ΔN44 Impβ or Dam-Impβ peaks (Figure 2b). The addition of RanGTP improves the signal to noise for Dam-ΔN44 Impβ, which is likely because RanGTP can increase the permeability of transport receptors to the central plug of NPCs (Lowe et al., 2015). For all the assays in this study, we have used Dam-ΔN44 Impβ for all the experiments due to lower background, unless otherwise mentioned.

Next, we compared NPC-DamID with *in vivo* DamID-seq data previously obtained using Dam-Nup153 expressing IMR90 cell line. While there is significant overlap between the two data sets, we detected many additional peaks in Dam-Nup153 data which do not overlap with NPC-DamID data (Figure 2c). These additional peaks are likely the result of the mobile nucleoplasmic fraction of Nup153 binding to chromatin. To test this idea, we performed 3D genome modeling using Chrom3D (Paulsen et al., 2017) to simulate either the position of NPC-DamID peaks or Dam-Nup153 peaks relative to known LADs sequences at the nuclear periphery (Figure S2g-h). The results show that the distribution of the distance of Dam-Nup153 peaks from the nuclear periphery is indeed bimodal. The fraction of peaks closer to the nuclear periphery overlaps with NPC-DamID peaks, while the other fraction is further away from the NE. In all, NPC-DamID is a robust method when tested with Impβ either with a different affinity to NPCs or import efficiency and specifically reports NPC-associated genomic regions.

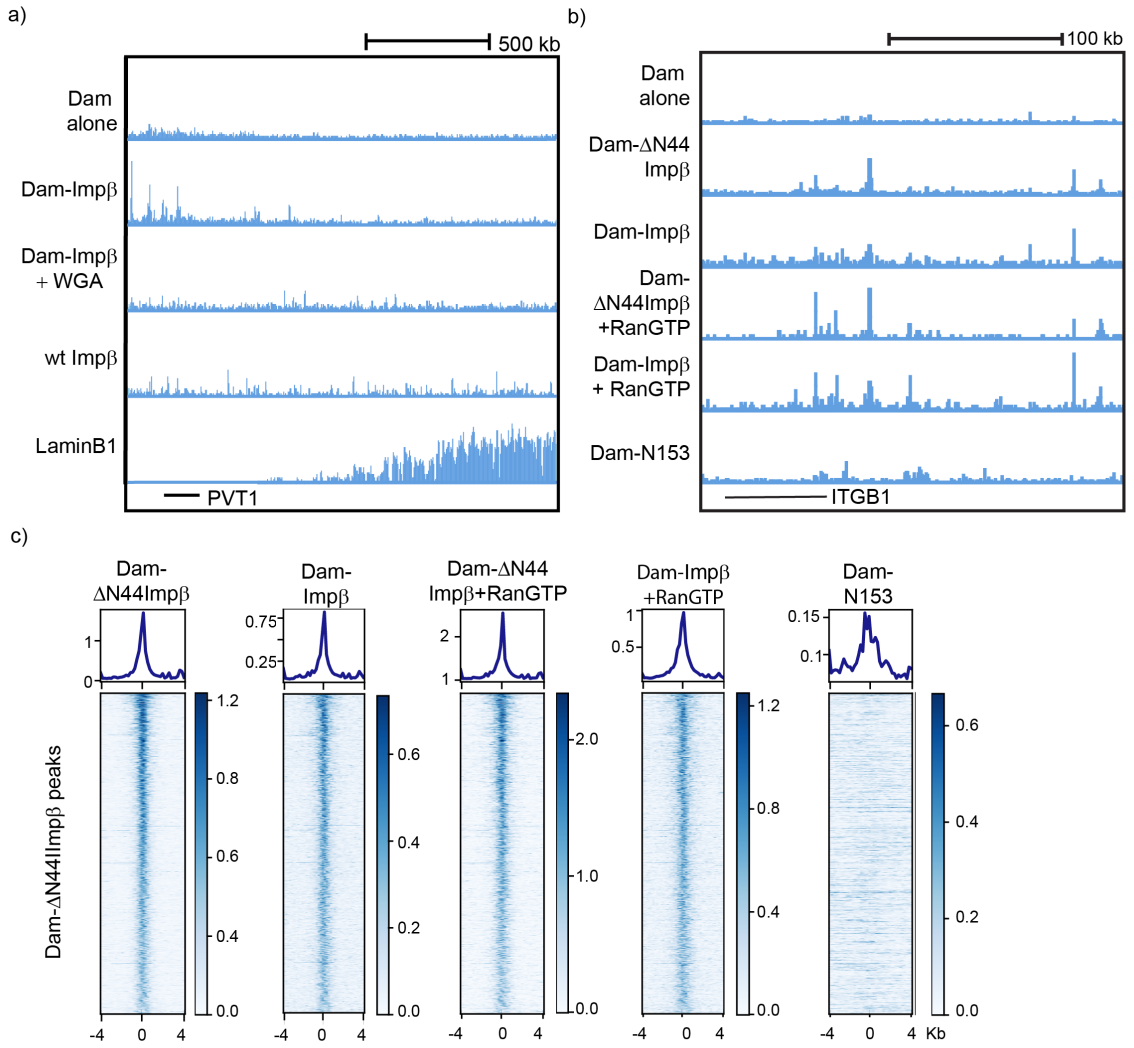

Figure 2. NPC-DamID is a robust and reproducible method to study NPC interaction with chromatin. a) Normalized NPC-DamID DNA sequencing tracks for Dam alone, Dam-Impβ, Dam-Impβ with WGA, wt Impβ, and Lamin B1 for the genomic region around the PVT1 gene. b) Normalized NPC-DamID DNA sequencing tracks for Dam alone, Dam-ΔN44Impβ, Dam-Impβ, Dam-ΔN44Impβ with RanGTP, Dam-Impβ with RanGTP and Dam-Nup153 for the genomic region around the ITGB1 gene. c) Sequencing reads/tags distribution centered around Dam-ΔN44Impβ peaks for Dam-ΔN44Impβ, Dam-Impβ, Dam-ΔN44Impβ+ RanGTP, Dam-Impβ+RanGTP, and Dam-Nup153.

NPC associated genomic regions are enriched in enhancer regulatory elements and actively express genes

After having successfully established and characterized the NPC-DamID method, we were interested in understanding the nature of the DNA associated with NPCs and incubated Dam-Impβ with different semi-permeabilized (i.e unfixed) cells, including HeLa, IMR90, and mouse C2C12 myoblasts. When we analyzed the distribution of different histone modification ChIP-seq reads centered at NPC-DamID peaks, we found that the NPC-DamID peaks preferentially enrich at open chromatin marks like H3K27Ac and H3K4Me3 as opposed to repressive mark like H3K9Me3 (Figure 3a, Figure S3). We further tested for the functional features of these genomic regions by comparing them to the 18 states of the ChromHMM model for IMR90 (Ernst and Kellis, 2017). Consistent with the enrichment of enhancer and promoter histone marks, we observed the biggest overlap with active enhancers among 18 different chromatin states for NPC DamID peaks (Figure 3b). The next significant overlap was found with weak enhancers and weak transcription regions followed by TSS. To check the transcriptional output of NPC-DamID peaks for IMR90, we compared the transcription of NPC-associated genes with a randomly chosen equal number of other genes. We have found that the average transcriptional output from NPC-genes is higher than the random genes (Figure 3c). This suggests that NPC-associated genomic regions are mostly open and active chromatin regions. This feature of NPC interacting chromatin was found to be consistent in other cell lines like HeLa, pancreatic islets, and C2C12 (Figure S3).

Since NPC-DamID does not involve generating genetically modified cells, we were also able to map chromatin associations with NPCs in unfixed human pancreatic islets tissues (which would have been challenging otherwise with *in vivo* DamID or ChIP-seq). After dissociating islets into single-cell suspension, we performed the assay in about 50,000 cells. We observed that NPC interacting DNA was enriched in histone marks for enhancer (H3K27Ac) and promoter (H3K4Me3) regions (Figure 3d), similar to other cell types. This way, NPC-DamID allows us to expand our understanding of the genomic associations at the periphery from cells to tissues.

Taken together, across multiple cell types, NPCs interact with open chromatin regions enriched in enhancers and promoters, and the NPC-associated genes show high transcriptional output.

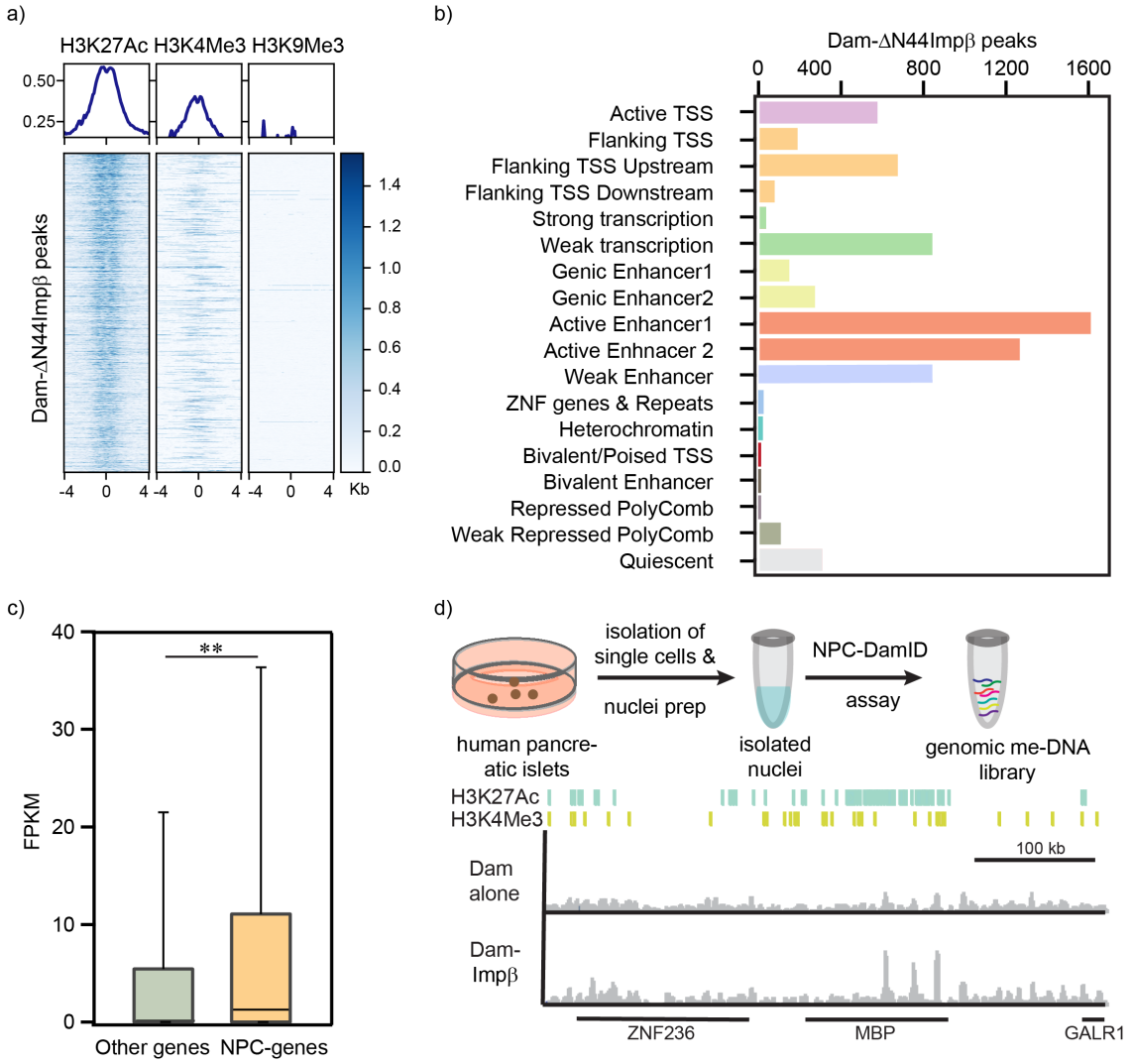

Figure 3. Genomic DNA interacting with NPCs contains mainly regulatory elements. a) Distribution of the tags/sequencing reads of ChIP-seq for H3K27Ac, H3K4Me3, H3K9Me3 centered around NPC-DamID peaks (Dam-ΔN44Impβ) in IMR90 cells. b) Percentage overlap of NPC-DamID peaks with the genome-wide annotation of IMR90 from the 18-state ChromHMM model. c) Transcriptional output comparison for genes associated with NPCs (NPC-genes, 1270) versus the same number of randomly selected other genes. d) Schematic representation of the NPC-DamID assay for live human pancreatic islet tissues. The normalized profile for Dam alone and Dam-Impβ

along with the tracks for histone marks- H3K27Ac and H3K4Me3 for human pancreatic islets are shown.

NPC association with super-enhancer regions in various cell types

While the association between Nups like Nup153, Nup93, Nup133, and Nup98 and super-enhancers (SEs) has been established(Ibarra et al., 2016; Liu et al., 2017; Scholz et al., 2019), we wanted to generate high-precision maps of these interactions in multiple cell and tissue types to identify SEs that are specifically associated with NPCs at the NE. To do this we performed NPC-DamID sequencinf for HeLa, IMR90, CC2C12 myotubes and human pancreatic islets tissue. We found a consistent and significant overlap (25-65%) between SEs and NPC-DamID peaks for all the cell and tissue types (Figure 4a). To further understand the role and importance of these interactions, we aimed to study how SEs associated with NPCs (NPC-SEs) are different from SEs that do not associate with NPCs. Although SEs are known to regulate genes that are crucial for cell-fate determination, being highly specific to a cell type, recently, it has been shown that about 48% SEs are at least shared between two or more cell types. These common SEs were also shown to have relatively high gene expression compared to unique and non-unique enhancers, and the associated genes were found to be enriched in basic cellular functions, such as transcriptional regulation, cell motility, and regulation of cellular proliferation (Ryu et al., 2019). Interestingly, common SEs are also strongly associated with fast recovering chromatin loop after sequential cohesin depletion and restoration (Ryu et al., 2019). We wanted to know if NPC-SEs are specifically enriched in common SEs as compared to the other non-NPC-SEs. We generated the cell type consensus score for NPC-SEs, non-NPC-SEs, and all SEs identified in IMR90 & Hela cells across 10 different cell types. Our results show that NPC-SEs have a slightly higher consensus score than non-NPC-SEs, indicating that NPC-SEs have higher enrichment for common super-enhancers (Figure 4b, Figure S4a). Concordantly, the gene expression for NPC-SEs was higher as compared to non-NPC-SEs (Figure 4c, Figure S4b). These results indicate that SEs associated with NPCs are relatively more shared between different cell types, and the genes regulated by these SEs are responsible for basic cellular functions needed for cell differentiation. This indicates that

NPC-linked SEs play a role in establishing chromatin organization by facilitating the formation of early high order chromatin loops that later become scaffolds for cell-type-specific loop formation.

The enrichment in common SEs at NPC-SEs indicates a possible role for NPCs in genome organization mediated through interactions with genome architectural proteins like CTCF, cohesins. A recent study revealed that a subset of SEs is hierarchically organized, where hub enhancers are enriched with CTCF and cohesin binding and play an important role in SE functional and structural organization(Huang et al., 2018). Based on this knowledge, we analyzed the structural organization of NPC-SEs in IMR90 cells and intriguingly found that a significant fraction of NPC-SEs (~70.0%) have a hierarchical organization (Figure 4d). Within hierarchical NPC-SEs, we observed NPC preferentially bind on hub enhancers compared with non-hub enhancers (2.1-fold, *P* = 3.4E-21, Fisher's exact test, Figure 4e). Interestingly, we further observed the co-occupancy of NPC and structural proteins like CTCF, BRD4, p300 and PolII are markedly elevated at hub enhancers compared with non-hub enhancers (5.1-fold, 2.9-fold, 2.2-fold, and 2.2-fold) (*P = 3.4E-08, P = 3.8E-21, P = 4.8E-13, and P = 2.4E-11*, Fisher's exact test, Figure S4c-e). This result suggests that NPCs might have functionally related roles on hub enhancer elements in mediating structural and functional organization of SEs through the recruitment of structural protein complexes (Figure S4f).

In all, these results suggest that NPC-SEs are important for genome organization and folding, possibly through interaction or regulation of chromatin architectural proteins.

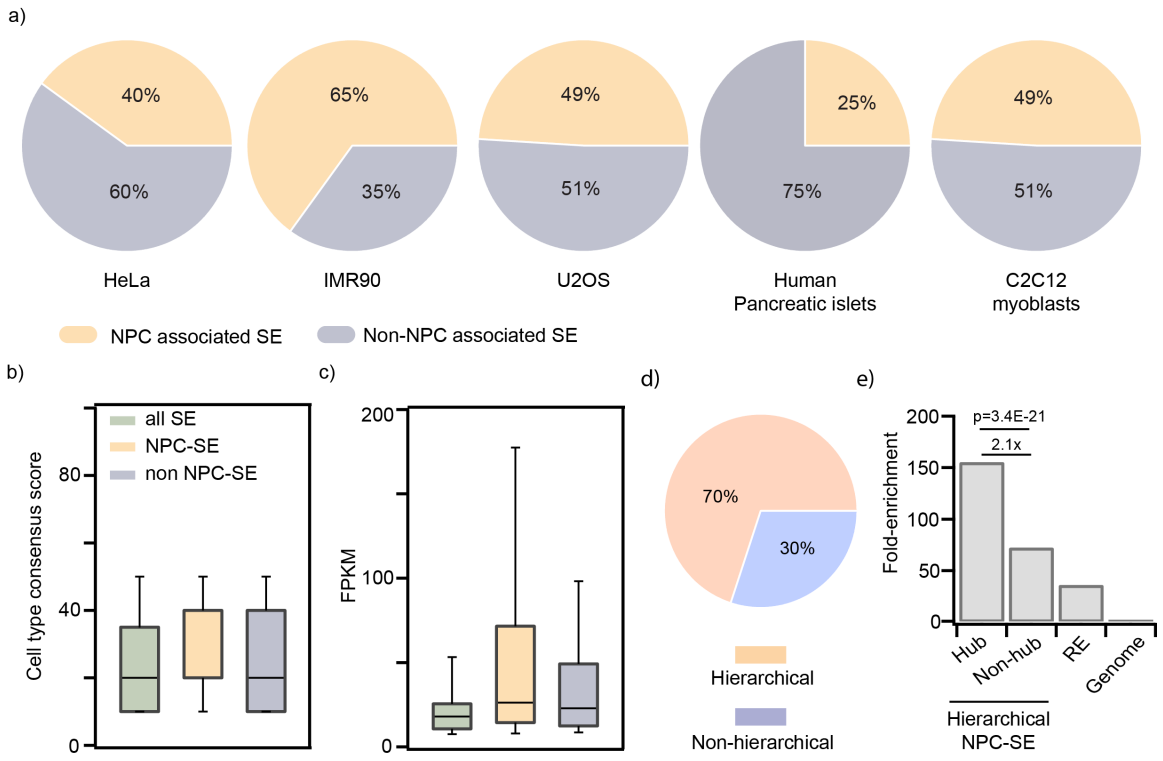

Figure 4. NPCs are consistently interacting with SEs across multiple cell and tissue types. a) Pie charts show the overlap of NPC-DamID peaks with SEs and non-SE regions for HeLa, IMR90, U2OS, human Pancreatic islets, and C2C12 myotubes. b) Bar plots for cell type consensus score of all SEs, NPC associated SEs (NPC-SE), SEs not associated with NPCs (non NPC-SE) in IMR90 cells. c) Transcriptional output of all SEs, NPC-associated SEs (NPC-SE), SEs not associated with NPCs (non NPC-SE) in IMR90 cells. d) Pie chart showing that 70% of NPC-SE have hierarchical organization while 30% have non-hierarchical. e) The occupancy of NPC in four enhancer groups: hub (n=1,050) and non-hub (n=7,874) enhancers within hierarchical NPC-SEs, regular enhancers in IMR90 and randomly selected genome regions as control. *P* values were calculated using Fisher's exact test. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$, n.s. not significant.

NPCs are hubs for chromatin structural proteins

As described above, we found NPCs to be consistently associated with SEs in several cell/tissue types and enriched in common SEs. Next, we sought to understand the mechanism through which NPCs engage and regulate the activity of SEs. SEs are known to have high occupancy of chromatin structural proteins like CTCF, mediator proteins, coactivator proteins like p300, BRD4, cohesins, and PolII (Whyte et al., 2013).

Our finding that NPCs are associated with hub enhancers (Figure 4d-e) encouraged us to propose a model where NPCs interact with chromatin structural proteins and participate in maintaining the structure of SE domains at the nuclear envelope (Figure S4f). To test this possibility, we used proximity ligation assays (PLA) to assess the interaction between Nups and structural proteins. PLA allows identifying in-situ protein interactions (at a distance <40 nm) with high specificity and sensitivity. We found that several SE-associated proteins, including CTCF, PolII, Med1 (Figure 5a), BRD4, p300 and SMC3 (Figure S5), show enriched PLA signal at the nuclear rim. We used NPC and Importin beta interaction as a positive control in these experiments to assess the specificity of the signal. Further, we also confirmed the overlap of NPC-associated chromatin with the DNA occupied by the structural proteins. To do this, we analyzed the distribution of the ChIP-seq reads for structural proteins like CTCF, BRD4, p300, and PolII around the center of NPC-DamID peaks. We found a specific enrichment for these proteins at NPC-linked chromatin regions (Figure 5b). Together, the data suggest that NPCs are structural hubs for SEs, which may be maintained through interaction with chromatin structural proteins.

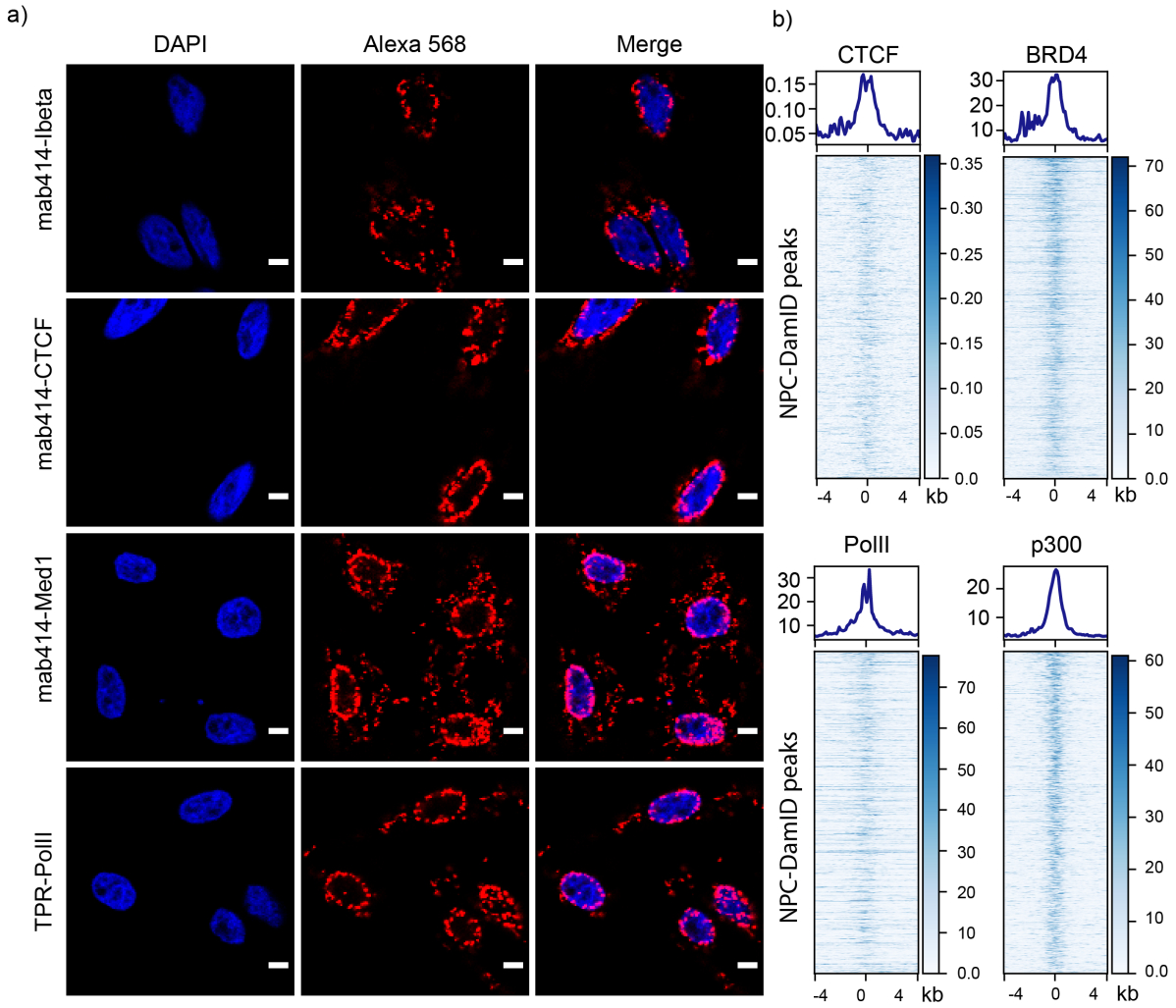

Figure 5. NPCs are hubs for chromatin structural proteins. a) Confocal images of IMR90 cells for proximity ligation assay (PLA) between the FG-nucleoporin marker mab414 and Impβ, mab414 and CTCF, mab414 and Med1, mab414 and PolII. Scale bar is 1 μm. b) Distribution of the reads/tags of CTCF, BRD4, Med1, PolII ChIP-seq centered around NPC-DamID peaks in IMR90 cells.

Nups exhibit phase separation properties which are important for their interaction with chromatin structural proteins

The observed association of SEs at NPCs raised the interesting question of how these interactions are established and maintained. Chromatin organization of SE regions has recently been shown to be maintained through multivalent cooperative

interactions between the intrinsically disordered regions (IDRs) of the mediator and transcription coactivator proteins (Sabari et al., 2018). The disordered domains of protein like Med1, BRD4, and PolII, when present in high concentrations at SEs, can form multimeric intermolecular interactions leading to the formation of phase-separated condensates (Sabari et al., 2018). NPCs are also highly enriched in completely or partially intrinsically disordered proteins. Similar to Med1, when Nucleoporins (Nups) like Nup153, Nup98, Nup93, and TPR were tested through a disorder prediction computational tool called PONDR (Xue et al., 2010)(Figure 6a, Figure S6a), they showed large IDRs (45-85%).

These results encouraged us to test whether peripheral Nups in NPCs interact with chromatin structural proteins at SE condensate hubs through their IDRs (Figure 4f). The *in vitro* liquid droplet formation property of FG-containing Nups has been previously studied in the nuclear transport process (Frey et al., 2006; Labokha et al., 2012)). However, its significance in chromatin organization has not been explored yet. We set out to compare the liquid droplet properties of Nups IDR with Med1 IDR, which has been previously shown to be involved in chromatin organization of SE regions through phase separation. We transfected the IDRs of Med1, Nup153 and Nup98 fused to GFP into cells and then observed the formation of droplets. We saw that both Nup153 and Nup98 IDRs form droplets that were disintegrated in the presence of hexanediol (HD), similar to Med1 (Figure 6b). The FRAP recovery time (τ) for Nup153 and Nup98 IDRs in a region of 1.5 μm^2 was about 19 s compared to 11 s for Med1 IDR (Figure S6c). The diffusion coefficient for Med1 IDR was 0.13 $\mu m^2/s$ as also determined by a previous study (Sabari et al., 2018). The diffusion coefficients for Nup153 and Nup98 FG were found to be 0.076 and 0.077 $\mu m^2/s$ which is roughly half of Med1 and similar to previously described components of liquid-like condensates (Nott et al., 2015; Sabari et al., 2018). From these results, we can conclude that Nups liquid-like droplets are sensitive to hexanediol disruption and are slightly less dynamic in diffusion than Med1.

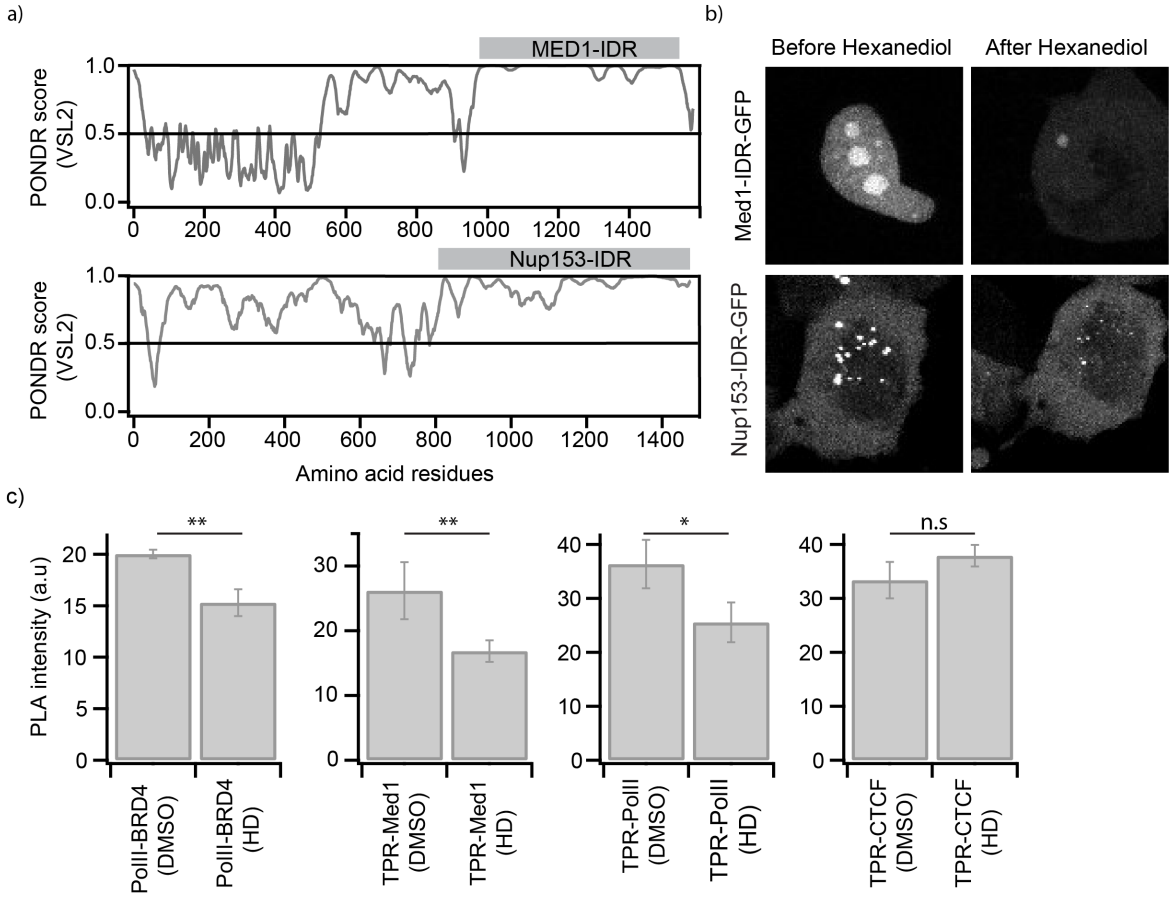

Figure 6. Phase separation properties of Nups are important for their interaction with the other proteins of SE hubs. a) PONDR prediction for intrinsically disordered regions (IDR) in the structure of Med1 and Nup153. A score over 0.5 signifies the lack of structure. b) Confocal images of HeLa cells transfected with GFP fused with either Med1-IDR and Nup153-IDR before and after 30 s of treatment with 1.5% hexanediol (HD). Scale bar is 1 μm. c) FRAP recovery plots for 1.5 μm^2 bleach spot in HeLa nuclei transfected with GFP fused to Med1-IDR or Nup153-FG or Nup98-FG. The solid black line represents the curve fit, and tau is the diffusion time calculated from the curve fits (more details in the Methods section). d) Bar plot showing the fluorescence intensity of PLA (proximity ligation assay) spots for interactions between PolII-BRD4, TPR-BRD4, TPR-PolII, or TPR-CTCF in HeLa cells after treating with either DMSO or 1.5% HD for 20 min at RT.

To further test if the Nups interaction with other SE proteins like BRD4 and PolII results in phase-separated condensates, we performed PLA experiments in the absence and presence of HD. We treated the cells either with DMSO or HD for 20 min

and probed for interactions through PLA. Upon quantification of the intensity of PLA spots, we found that TPR-BRD4 and TPR-PolII interaction decreased after HD treatment in a similar manner to the interaction between known co-phase separating condensate proteins, BRD4 and PolII (Figure 6c). However, PLA spots intensities remained unchanged for TPR-CTCF, probably because CTCF does not have IDRs to participate in phase separation with other coactivator proteins in SE regions (Figure 6c). Altogether, our results demonstrate the involvement of Nups in the organization of SE through interactions with chromatin structural proteins in phase-separated liquid droplets.

Synthetic LacO arrays mediate the formation of Nup IDR interaction hubs that can recruit other chromatin structural proteins

Several studies have used *in vitro* purified IDR proteins at high concentrations to form hydrogel or liquid-liquid droplets to understand the biophysical and biochemical properties of interactions between IDRs (Frey et al., 2006; Sabari et al., 2018) . FG repeats containing Nups like Nup98 and Nup153 have also been previously studied for *in vitro* hydrogel formation and its importance for selective permeability of nuclear transport receptors (Labokha et al., 2012). There is a significant difference between *in vivo* physiological conditions and those used for *in vitro* phase separation experiments (Chong et al., 2018). Experimental conditions like temperature, concentration, pH, purity, and the microenvironment are detrimental for interaction between IDRs to form liquid droplets. Hence, we decided to use *in vivo* lac operator (LacO) arrays to probe the interaction between IDRs of Nups and structural proteins of SE hubs. The U2OS cells containing the LacO array (harboring ~50,000 LacO repeats) [31,32] are transfected with fluorescently tagged IDRs fused to LacI (mCherry-IDR-LacI), which leads to recruitment of a large number of molecules of mCherry-IDR-LacI at LacO site (Figure 7a,b). In this way, concentrated *in vivo* interaction hubs can be formed inside the nucleus. We used this system with Nup153 IDR by fusing Nup153 IDR with mCherry, NLS, and LacI. Upon transfection of mCherry-N153-NLS-LacI to the cells with LacO array, we could observe a bright fluorescent spot with more intensity and larger size than mCherry-NLS-LacI

suggesting that the copy number of N153-LacI at LacO associated hubs is higher than LacI alone (Figure 7b). A similar trend was observed for Med1-IDR-LacI spots. Additionally, LacI fused to N153 and Med1 could form multiple fluorescent puncta throughout the nucleus compared to LacI alone, which leads to a single fluorescent spot. These results indicate likely cooperative and multivalent intermolecular interactions for N153-LacI and Med1-LacI inside the nucleus, critical for the formation of phase-separated droplets.

Next, to confirm the multivalent and cooperative interaction between LacI-IDRs at LacO hubs, we performed FRAP experiments. We found that the dynamics of LacI alone spots were different than LacI fused to IDR of either Nup153 or Med1. Upon fitting the FRAP curves with the reaction-dominant model (Chong et al., 2018), we found that Nup153 IDR and Med1 IDR fused to LacI had about 80% reduction in the dissociation constant (k_2) as compared to LacI alone (Figure S7). This suggests that the locally high concentration of Nup153 or Med1 IDR hubs at the LacO site are driven by IDR-IDR interactions via multivalent contacts, which would result in stabilizing the association with cognate genomic DNA and other partner proteins. This result is important in the context of SE hubs which are formed through multiple homotypic and heterotypic interactions between IDRs of transcription factors and chromatin structural proteins (Sabari et al., 2018).

After confirming homotypic, multivalent interactions for IDR of Nup153 and Med1 fused to LacI, we investigated the potential of Nup153 IDR to interact with the known SE resident IDRs -PolII and Med1. We again took advantage of the LacO-LacI system to form a hub for IDR of Nup153 to observe the recruitment of PolII and Med1. We co-expressed PolII CTD52 repeats of PolII or Med1 IDR fused with GFP in cells with mCherry-Nup153IDR-LacI. We observed a colocalized fluorescent signal of GFP and mCherry for both PolIICTD52 and Med1 IDR, while LacI alone cells showed only mCherry signal (Figure 7c). Similar colocalization was also observed when cells expressing mCherry-Med1 IDR-LacI were transfected with GFP tagged PolII CTD52. These results suggest that the Nup153 IDR domain, when present in high

concentrations at the LacO site, can recruit other partner proteins of SE hubs like PolII and Med1. It indicates that Nup153 functions similarly to Med1, previously shown to form phase-separated droplets at SEs sites to bind other structural proteins of SE. Altogether, these results support a model that peripheral Nups like Nup153 form phase-separated condensates at NPCs that compartmentalize and concentrate transcriptional coactivators and structural proteins at SE-regulated genes.

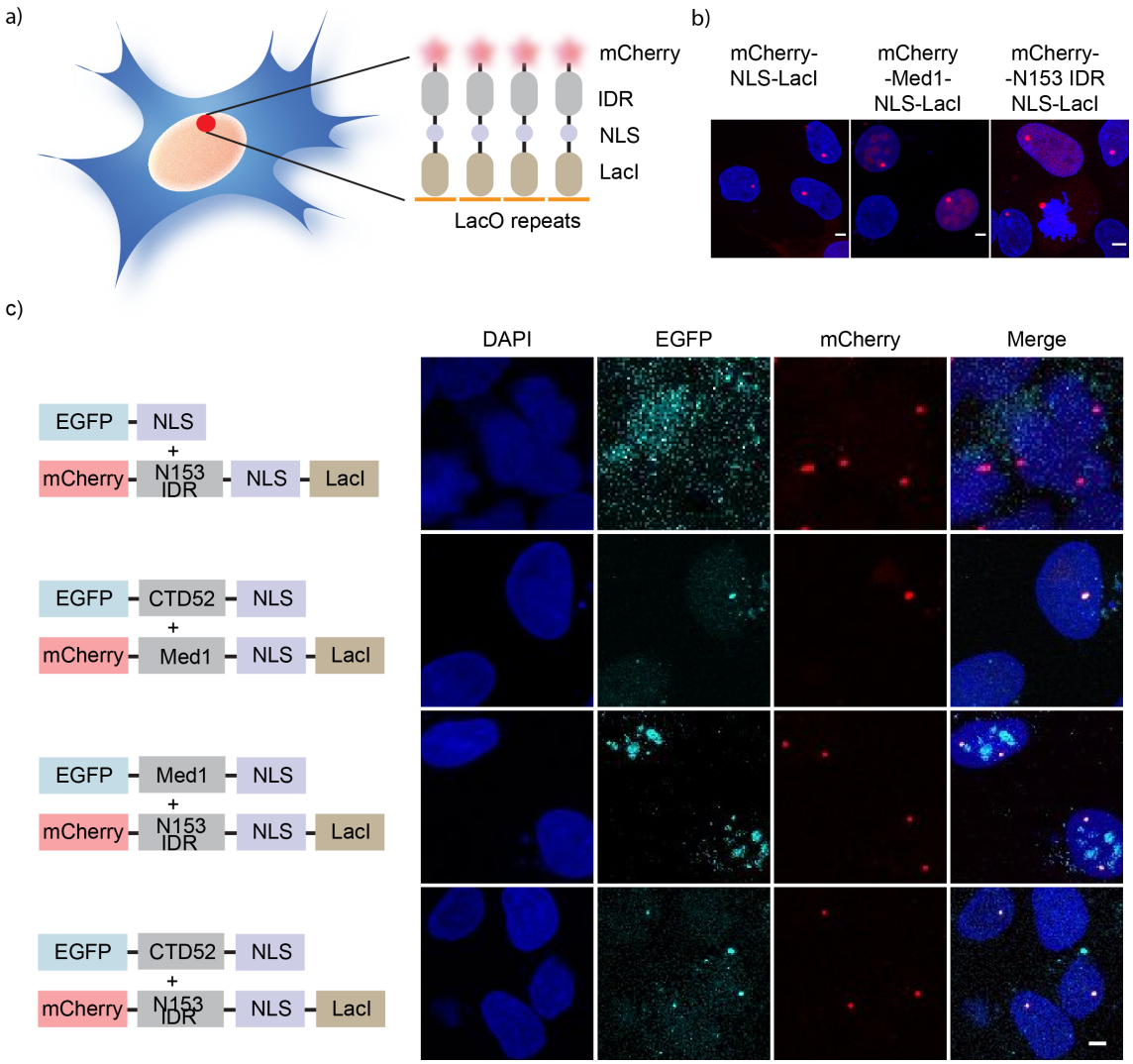

Figure 7. Synthetic LacO arrays mediated Nup IDR hubs can recruit other chromatin structural proteins. a) Scheme of a cell containing an array of LacO repeats (~50,000) which can be decorated with the binding of LacI fused to a fluorescent protein, mCherry, IDRs, and nuclear localization signal (NLS). b) Confocal representative images of U2OS cells containing LacO repeats transfected with LacI fused to mCherry alone (mCherry-NLS-LacI), mCherry-Med1(IDR)-NLS-LacI or mCherry-Nup153 (IDR)-NLS-LacI. c) Confocal images of U2OS cells harboring LacO array and co-transfected with EGFP-NLS and mCherry-Nup153-FG-LacI or EGFP-PolII CTD52-NLS and mCherry-Med1 IDR-NLS-LacI or EGFP-Med1-NLS and mCherry-Nup153FG-NLS-LacI or EGFP-PolII CTD52-NLS and mCherry-Nup153FG-NLS-LacI. Scale bar is 1 μm.

DISCUSSION

In this study, we present a new method, NPC-DamID, to study NPC-mediated genomic interactions. NPC-DamID has several advantages over existing methods to probe genome-wide interactions at the nuclear periphery. First, NPC-DamID effectively eliminates any contribution of genomic interactions with Nups that are not at the NPCs. Previously, it had been challenging to discern the difference between genomic interactions of Nups in the nucleoplasm and at NPCs due to the mobile nature of Nups used in other methods. Since NPC-DamID uses recombinantly expressed and purified Importinβ fused to Dam, which specifically targets NPCs in semi-permeability cells, it reports interactions exclusively at the nuclear periphery (Figure 1a, d,e). We have shown through experimental and modeling data that NPC-DamID better reflects peripheral NPC-genome interactions than previously published datasets using Dam-Nup153 (Figure 2a,b). Besides, we have demonstrated that the dominant mutant of Impβ (Dam-ΔN44 Impβ), which has a higher affinity to NPCs compared to wt Impβ (Kutay et al., 1997), shows NPC-genome interactions similar to Dam-Impβ. In the future, NPC-DamID-seq can be expanded to map the NPC-associated proteome by fusing Impβ to APEX2 (Hung et al., 2016), which will be useful to discover novel interacting partners of NPCs involved in genome regulation. Second, our method allows us to probe NPC-genomic DNA interactions in any cell or tissue type since the method does not rely on transgenic cell lines. To this end, we have mapped the interactions of NPC with genomic DNA in four cell lines and live human pancreatic islet tissue samples (Figure 3d). In the future, NPC-DamID will be a valuable tool to understand better the influence of disease conditions like diabetes or cancer on genome architecture at the nuclear periphery. This is possible also since NPC-DamID can be done using as few as 100,000-200,000 cells. Third, NPC-DamID is not only independent of antibody and fixation as conventional DamID but also free from over-expression artifacts of Dam fused proteins. The concentration of Dam-Importinβ and the timing of the assay can be precisely controlled to avoid any artifacts. Fourth, the temporal resolution of NPC-DamID is higher than conventional DamID since there is no requirement of waiting 2-3 days to get enough expression of Dam fused protein. Through NPC-DamID, it is

possible to obtain a snapshot of NPC-genomic DNA interactions at any time point, for example, during the time course of differentiation (data not shown). Fifth, NPC-DamID sequencing can be combined with super-resolution imaging which is difficult with conventional DamID (Kind et al., 2013). In conventional DamID, Dpn7 (a DNA binding fragment of Dpn1) fused to a fluorescent protein is co-expressed with the Dam tagged protein of interest to allow imaging of the methylated DNA. This works well in cases like Dam-lamin, when megabase pair long stretches of DNA methylation are expected. However, for sparse methylation patterns, as one would expect for NPC-associated DNA, the signal strength is too weak to perform high-resolution imaging, which requires high signal density. In NPC-DamID, we take advantage of the *in vitro* nature of the assay and use purified GFP or Flag-tagged Dpn8 labeled it with a succinimide synthetic fluorescent dye like Alexa 647 (Figure 1e,f). This results in an increased number of dye molecules per Dpn8 molecule and gives us the flexibility of using synthetic dyes, which perform better in super-resolution methods like dSTORM. Thus, we can probe NPC association with genomic DNA through imaging at a single NPC resolution. A previously described model suggested that the heterogeneous composition of NPCs in a single cell could lead to differences in its interactions with chromatin (Gozalo et al., 2019; Ori et al., 2013; Raices and D'Angelo, 2012). Our method combined with imaging will be useful in future studies to explore such models and understand the role of heterogeneity in NPC structure in genome organization. Furthermore, the *in vitro* nature of this method can allow biochemical manipulations to study the more specific nature of NPC-DNA interactions. For example, in the future, we could map the distance-dependent interaction of the genome with NPCs by using Dam-Impβ with increasing linker lengths.

The association of Nups with genomic DNA has been previously shown to be important for gene regulation. Recent evidence has indicated that the NPCs also mediate the structural organization of the genome(Kadota et al., 2020; Pascual-Garcia and Capelson, 2019; Pascual-Garcia et al., 2017) . However, a comprehensive understanding of the molecular mechanism underpinning such roles of Nups or NPCs remains unclear. Previous studies used specific Nup-targeted ChIP-seq or DamID-seq, and hence it was difficult to assess whether the genomic associations shown

represented binding to a mobile nucleoplasmic pool of Nups or NPCs at the periphery (Ibarra et al., 2016; Jacinto et al., 2015; Kadota et al., 2020). These studies rely on DNA-FISH to test the localization of genomic interactions with Nups, which can only be tested for few loci due to low throughput. With our new method NPC-DamID, we were able to probe genome-wide interactions of NPCs that happen specifically at the periphery in multiple cells and tissue types to understand the common and unique nature of these interactions. We have found that, in various cell lines, the NPC-associated genomic DNA is primarily enriched in intergenic regions bound to H3K27Ac, followed by H3K4Me3 histone marks Figure 3a, Figure S3). This indicates that NPCs associate with regulatory enhancer and promoter regions, which falls in line with previous studies in human, metazoan, yeast, and mouse cells (Franks et al., 2017; Ibarra et al., 2016; Kadota et al., 2020; Liang et al., 2013; Light et al., 2013; Pascual-Garcia et al., 2017). Interestingly, in recent studies in mouse embryonic stem cells, Drosophila, and yeast, Nups like Nup93, Nup107, and Nup153 have also been shown to associate with repressed/facultative heterochromatin regions of genomic DNA (Gozalo et al., 2019; Iglesias et al., 2019; Jacinto et al., 2015). Our study found a higher association of actively expressed genes with NPCs, and only about 7% of NPC-DamID peaks overlap with Polycomb-associated repressed or quiescent genes (Figure 3b). This result suggests that NPCs can associate with both active and repressed regions, probably through different Nups and at different locations in the nucleus. Future, NPC-DamID assay for yeast and drosophila cells will help to discern the contribution of Nups at NPCs in NE from a nucleoplasmic pool of Nups in the association with different chromatin features. It also incites the intriguing hypothesis that NPCs can vary in their composition and structure inside a nucleus, which leads to their association with different classes of genomic DNA and regulatory proteins. In the future, such models can be tested by combing NPC-DamID with super-resolution dual-color imaging to visualize the correlation between the structure and genomic associations of individual NPCs.

Another common genomic feature that associates with NPCs across multiple cells and tissue types is super-enhancers. This is consistent with previous work from

our lab in which we have shown SE association with Nup153 and Nup93 through DamID-seq(Ibarra et al., 2016). In the present study, we further explored the nature of NPC-associated super-enhancers (NPC-SE) as opposed to other SE regions. We found that genomic regions that function as super-enhancers across several cell types (common SEs) tend to show more association with NPCs than genomic regions that function as super-enhancers in a cell-type-specific manner (Figure 4b). In a recent study, the common super-enhancers were shown to be important for establishing the early genome contact points and loops in chromatin refolding during the cell cycle (Ryu et al., 2019). This suggests a possible role for NPCs in establishing the foundational organization of the genome. Moreover, we found that NPC-SEs have hierarchical structural organization and preferentially enrich for hub enhancers (Figure 4d,e). Hub enhancers are enriched in chromatin structural proteins like CTCF and cohesins. In this way, they play an important role in dictating the chromatin landscape and regulating gene expression. These findings imply that the interaction of NPCs with SEs is important to the structural and functional organization of genomic DNA at the nuclear periphery through interactions with chromatin architectural proteins.

To explore the features and mechanisms for NPC-mediated genome organization, we tested the interaction of chromatin structural proteins, specifically those associated with SEs. We found through proximity ligation experiments that NPCs can interact with SE proteins like CTCF, Med1, Brd4, PolII, and cohesins (Figure 5a,b). There is also significant enrichment of these proteins around the NPC-DamID peaks. Together, these results strengthen the possibility of NPCs mediating the organization of SEs at the periphery through interactions with chromatin structural proteins. These results also corroborate recent studies showing the interaction of Nups like Nup153 or Nup93 with CTCF and cohesin (Kadota et al., 2020; Labade et al., 2019; Sachani et al., 2018). The knockdown of Nup153 influences the genome-wide binding of CTCF and cohesin (Kadota et al., 2020). Hence, it's likely that the interactions reported in these studies are also from a nucleoplasmic pool of Nup153. It is also unclear in these studies how Nup153 may influence the binding of CTCF and cohesin to the genomic DNA.

We also explored the mechanism through which Nups at NPCs might be interacting with resident proteins of SE hubs (Figure 5,6). SEs are clusters of multiple enhancers which recruit transcription factors and coactivators containing intrinsically disordered regions (IDRs) in high concentration. The high concentration of IDRs allows multivalent interactions, leading to the formation of phase-separated condensates. This mechanism is important to ensure robust and high gene expression of cell identity genes. The peripheral Nups at NPCs like Nup153 also have a long IDR, which we show in this study forms phase-separated droplets *in vivo*. Nup153 is known to have at least 32 copies at individual NPCs (Ori et al., 2013), which likely facilitates high local concentration for phase separation. Since individual NPCs have been predicted to have different compositions within a single cell and are also known to vary in their composition from one cell type to the other (Ori et al., 2013), SE-interacting NPCs may have more than 32 copies of Nup153 to facilitate the interaction with SEs. In the future, high-resolution imaging could be done to test this hypothesis. We show that phase-separated Nup153 droplets can co-partition with IDRs of other SE resident proteins like Med1 and PolII. This observation suggests that NPCs can be structurally and functionally part of SE hubs. At the same time, it also raises multiple questions for the future regarding the importance of NPCs in regulating this subset of SEs. Do NPCs act as a scaffold to organize SE hubs? How different are the organization and function of NPC-associated SEs compared to others? Why do certain SEs tend to associate with NPCs? In this regard, the gene-gating model, in which positioning if a gene at NPCs is required for rapid transport of mRNA to the cytoplasm (Blobel, 1985), is more probable. Recent evidence from a study of an NPC-associated oncogenic super-enhancer for the myc gene showed that SE positioning at the periphery facilitated the export of mRNA transcripts in colon cancer cells in response to WNT signaling (Scholz et al., 2019). It will be interesting to determine if gene-gating holds for all NPC-associated SEs and non-cancer cells. Also, interestingly nuclear basket protein, TPR is known to interact with multiple kinases like ERK2 and CDK1 (McCloskey et al., 2018; Rajanala et al., 2014)and previously, Nup-SEs were shown to enrich the motif for AP1 like transcription factors (Ibarra et al., 2016). It will be interesting to study if kinase interaction with TPR mediates the phosphorylation of AP1 transcription factor to facilitate the gene

expression of NPC-SEs. Another functional distinction for NPC-SEs could be their role in mediating early steps of genome refolding after cell division. This could be particularly applicable to the SEs common between different cells since recent evidence suggests that they promote early loop formation during genome refolding. The NPC association of SEs may be critical for both gene-gating and genome organization roles since NPCs can associate with SEs that are unique to specific cell types too.

In conclusion, NPC-DamID has allowed us to better understand the role of NPCs as genome organizers and gene regulators. We have learned the common and unique features of NPC-interacting genome regions across multiple cells and tissue types. Further, we have shown the role of disordered domains in Nups in the structural organization of SE hubs through phase separation. This study has also raised new questions about the contribution of individual NPCs in genome organization. Moreover, further studies will be needed to better understand other interaction partners (protein or RNA) of NPCs that mediate the association of NPCs with genomic DNA.

AUTHOR CONTRIBUTIONS

S.T: conceived, designed all of the experiments, collected all the data, performed analysis, and wrote the paper. F.C and J.J : analysis for the hierarchical organization of NPC-DamID peaks and reviewed the paper. M.W.H: conceived, supervised and reviewed the paper.

STAR Methods

Recombinant expression and purification of the proteins

The plasmids for nuclear transport factors (Impβ, Impα, NTF2, Ran, NLS-MBP-GFP cargo) were received from Prof. Edward Lemke's laboratory. The plasmid contained intein-his tags for Ni-NTA column purification and later tag was removed by incubating protein in 100 mM beta-mercaptoethanol for 12-14 h at RT in presence of Complete protease inhibitor cocktail. The details of plasmids are also listed here (Paci et al., 2020). The purification steps for the proteins were followed according to Milles et. al. (Milles and Lemke, 2011) Dam was cloned in pet28 vector and expressed for 10 h at 18˚C in HMS174(DE3) E.Coli cells. Dam-Impβ and Dam-ΔN44Impβ were cloned pTXB3 plasmid with intein-His tag. It was also expressed in HMS174(DE3) cells for 10 h at 18˚C. Ni-NTA purification was done as per the manufacturer's guidelines. After purification through Ni-NTA column, intein-His tag was removed through beta-mercaptoethanol and the second Ni-NTA column purification was done. The protein was then concentrated and purified through size exclusion column chromatography.

For loading GDP to Ran we incubated Ran with 50 X molar ratio to protein with GDP in 20 mM HEPES (pH 8.0), 200 mM NaCl, 2 mM DTT and 5 mM $MgCl_2$ for 1 h at 4 C. After that we quenched the reaction with 25 mM $MgCl_2$ and overnight dialysis was against 20 mM HEPES, 200 mM NaCl, 2 mM DTT and 5 mM $MgCl_2$, 2 mM GDP.

We received ER295 and ER295(DE3) E. Coli strain and pet28-DpnI winged-helix domain (Dpn8, residues 183-254 aa) from Prof. Michael Bochtler. We cloned flag tag and also GFP tag into the plasmid Pet28-Dpn8. The protein was expressed in ER2925 DE3 E. Coli cells and purified following the step from Mierzejewska et. al. (Mierzejewska et al., 2014)

In vitro lambda DNA protection assay (enzyme assay for Dam proteins)

Unmethylated lambda DNA was used for protection assay to test the activity of Dam and Dam fused proteins. 1 µg of ʎDNA was incubated with different concentration of Dam and Dam fused protein along with 7.5 µM SAM (s-adenosyl methionine) in transport buffer (20 mM HEPES, 110 mM KOAc, 5 mM NaOAc, 2 mM MgOAc, 2 mM DTT, pH 7.3) for 30 min at RT. The enzyme was then deactivated by heating samples at 65°C for 10 min. Then 1X DpnII NEB buffer was added along with 1 µl of DpnII. The samples were incubated at 37°C for 30 min. The samples were heat-inactivated and mixed with agarose gel loading dye. The samples were loaded on 0.8% agarose gel with EtBr to assess the activity of the proteins.

Protein labeling through succinimide ester dyes

We labeled wt Impβ, Dam-Impβ, and Dpn8 with succinimide Alexa dyes with free -NH2 group in the proteins. The dye stock solution was made in anhydrous DMF solution and stored in lyophilized form at -20°C. The labeling was done in 5X molar excess of dye in sodium phosphate buffer, 2 mM DTT, pH 7.5 for 2h at RT. Later the reaction was dialyzed or purified through a size exclusion column to remove excess dye. The success of the reaction was checked by run samples on SDS-PAGE and scanning the gel on a fluorescent gel scanner.

Nuclear import assay

The buffers used were PBS, permeabilization buffer (50 mM HEPES, 50 mM KOAc, 8 mM $MgCl_2$, pH 7.3), and transport buffer (20 mM HEPES, 110 mM KOAc, 5 mM NaOAc, 2 mM MgOAc, 2 mM DTT, pH 7.3). The HeLa cells were washed for 3 × 2 min with PBS, followed by a 2-min wash with permeabilization buffer, followed by a 5-min permeabilization with digitonin at a concentration of 50 µg/ml. The digitonin was subsequently removed by washing for 3 × 3 min with transport buffer. After the final wash, HeLa cells were treated with an active nuclear import reaction mix containing a

fluorescent import cargo probe (NLS-MBP-GFP), importin-β (1 μM), Importinα (1 μM) RanGDP (5 μM), NTF2 (4 μM), and an energy regenerating system (2 mM GTP, 0.1 mM ATP, 4 mM creatine phosphate, and 20 μ/ml creatine kinase) in transport buffer. Import reactions proceeded at room temperature for 20 min before the cells were fixed with a 4% PFA solution for 15 min and washed 3 × 2 min with PBS. Cells were then imaged using a Zeiss LSM 700 confocal laser scanning microscope.

NPC-DamID assay

For NPC-DamID assay, cells were grown in 6 well plates with 70-80% confluency. The media was washed 3X with 1X PBS at RT. After washing, cells were incubated with permeabilization buffer for 3 mins followed by digitonin permeabilization for 5 min at RT. Digitonin concentration ranged from 20-30 μg/ml for semi-permeabilization. The concentration was optimized for different cells through checking immunofluorescence for antibodies against lamin in semi-permeabilized cells. The digitonin was then washed by 3X wash with transport buffer for 3 min each. Then transport buffer with cocktail of proteins and SAM required for the assay was added, 5 units of Dam or Dam-fused proteins, 4 μM NTF2, Importinα (1 μM), 7.5 μM SAM. For assays with RanGDP/GTP, RanGDP (5 μM), and an energy regenerating system (2 mM GTP, 0.1 mM ATP, 4 mM creatine phosphate, and 20 μ/ml creatine kinase) was also added. The cells were incubated with transport mix for 20 min at RT with occasional gentle rocking to mix the reagents. After 20 min, the cells were washed with 1X PBS for 3X, 3 min each.

NPC-DamID imaging through Dpn8

For super-resolution and confocal imaging, cells were grown on glass coverslips or ibidi glass-bottom chambers (for dSTORM imaging) with #1.5 thickness. After performing NPC-DamID assay, the reagents of the assay were thoroughly washed with 1X PBS at RT. Then cells were fixed with 2.4% PFA for 10 min and cells were permeabilized with 0.5% triton for 3 min. The cells were then blocked with 5% BSA for 1

h and incubated with GFP-Dpn8/Flag-Dpn8 solution made in 5% BSA (2 μM stock of the protein was diluted 1:100) for 30 min at RT. The cells were washed 3 times with 1 X PBS and costained with DAPI or Nup153 antibody for NPC marker. The cells were then post-fixed with 2.4% PFA for 5 min and mounted in vectasheild for confocal imaging. For super-resolution imaging, cells were exchanged in fresh STORM photoswitching buffer (1X PBS, 100 μM MEA cysteamine hydrochloride), 10% glucose, 5 μl catalase, 0.6 mg/ml glucose oxidase, pH 8). dSTORM single color and dual color was performed on ONI nanoimager (oxford nanoimaging instrument) with 640 nm and 488 nm excitation laser and 100X 1.49 NA Olympus objective. Up to 30,000 frames were collected with a frame rate of 50-75 Hz.

Cell culture

HeLa and U2OS cells were cultured in DMEM media supplemented with 10% FBS. IMR90 cells were cultured in 15% FBS, 1X NEAA (non-essential amino acids) in DMEM media. C2C12 cells were cultured in DMEM media supplemented with 20% FBS. To convert C2C12 myoblast to myotubes, C2C12 myoblasts were allowed to become confluent and then DMEM medium was supplemented with 2% horse serum instead of FBS. Myotubes were collected for the assay at 14 days. Pancreatic islets were bought from Prodo Laboratories Inc. Islets were maintained in cell culture at 37 °C, 5% CO_2 in PIM(S) Complete media for up to 7 days. Islets for NPC-DamID assay were disrupted into single-cells using TryplE just before the assay. The assay was performed in suspension in tubes.

NPC-DamID sequencing

Cells after NPC-DamID assay were scraped from 6 well plates and used for genomic DNA isolation. Qiagen gDNA isolation kit was used using the manufacturer's protocol. DNA was eluted in 500 μl of water or elution buffer. The DNA was then concentrated using a rotatory evaporator if eluted in water otherwise precipitated with ethanol to dissolve in 50 μl of water so that concentration is about 500 ng/μl. The DNA

was then prepared for library preparation following the published protocol for DamID-seq(Marshall et al., 2016). The library for DNA sequencing was prepared using Roche Kapa Hyper-prep DNA kit following the manufacturer's protocol. The quality of the DNA libraries were checked using Tapestation bioanalyzer and Quibit. The libraries were usually sequenced with single-end 50 bp sequencing.

Super enhancer consensus scoring

For super-enhancer consensus scoring, we downloaded the annotated SE sequences for 10 different cell lines from the super-enhancer database (SEA) (Chen et al., 2019). Then all SE sequences were merged and used as the total number of SEs. Then each of them was tested for co-occurrence in 10 cell types and scored (s) using formula $s = \frac{\Sigma_{i=1}^{n}(1-\hat{x}_i)}{n-1}$; $\hat{x} = \frac{x_i}{\max(x_i)}$ where x indicates cell/tissue type specificity and n denotes the number of cell/tissue types (Ryu et al., 2019).

Super enhancer hierarchy determination

Super-enhancer hierarchical score determination using NPC-DamID peaks and Hi-C already published data for IMR90 was done using the methodology described previously by Huang et.al. (Huang et al., 2018) For correlation with the RNA expression, we used RNA seq data for IMR90 and HeLa cells and selected the closest gene in distance from the super-enhancer as the related gene for calculating the expression.

PLA experiment

We performed PLA experiments using the Duolink PLA kit from Sigma-Aldrich. We followed the protocol from the manufacturer for the experiments. For imaging, we used Leica SP8 confocal microscope. For quantification, a custom script in Igor pro was used to calculate the fluorescence intensity in nuclei after background correction and segmentation in the maximum projected image from the z stack collected for each PLA experiment.

FRAP experiment

FRAP experiments were performed on Leica SP8 using the module for FRAP. To measure the FRAP dynamics of GFP fused to Nup-IDRs and Med1-IDR we acquired 90 frames every 1.3 s with 5 frames before bleaching. For m-Cherry labeled LCD-LacI or LacI, we acquired 200 frames every 2 s and 5 frames before bleaching. The average of the first 5 frames before bleaching was used to calculate the background fluorescence for the baseline. We chose a circular spot of radius 1.5-2 μm for photobleaching.

We followed the steps of analysis from Chong et. al. (Chong et al., 2018)We corrected the XY-drift of the bleached spot using the ImageJ plugin “Template matching and slice alignment”. We normalized the mean intensity of bleach spot and whole nucleus at time t to pre-bleaching baseline intensity. After this normalization, we normalized relative bleach spot intensity to nuclear intensity. Then we calculated the difference between normalized FRAP intensity before and after photobleaching. Finally, we normalized this difference in intensity to 100% to plot the curves. The reaction-dominated fitting to LacI curves was also done as previously described(Chong et al., 2018).

REFERENCES

Aughey, G.N., and Southall, T.D. (2016). Dam it's good! DamID profiling of protein-DNA interactions. Wiley Interdiscip Rev Dev Biology *5*, 25–37.

Aughey, G.N., Cheetham, S.W., and Southall, T.D. (2019). DamID as a versatile tool for understanding gene regulation. Development *146*, dev173666.

Blobel, G. (1985). Gene gating: a hypothesis. Proc National Acad Sci *82*, 8527–8529.

Buchwalter, A., Kaneshiro, J.M., and Hetzer, M.W. (2019). Coaching from the sidelines: the nuclear periphery in genome regulation. Nat Rev Genetics *20*, 39–50.

Capelson, M., Liang, Y., Schulte, R., Mair, W., Wagner, U., and Hetzer, M.W. (2010). Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes. Cell *140*, 372–383.

Chen, C., Zhou, D., Gu, Y., Wang, C., Zhang, M., Lin, X., Xing, J., Wang, H., and Zhang, Y. (2019). SEA version 3.0: a comprehensive extension and update of the Super-Enhancer archive. Nucleic Acids Res *48*, D198–D203.

Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., et al. (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science *361*, eaar2555.

Cook, P.R., and Marenduzzo, D. (2018). Transcription-driven genome organization: a model for chromosome structure and the regulation of gene expression tested through simulations. Nucleic Acids Res *46*, gky763.

D'Angelo, M.A. (2018). Nuclear pore complexes as hubs for gene regulation. Nucleus *9*, 00–00.

Ernst, J., and Kellis, M. (2017). Chromatin-state discovery and genome annotation with ChromHMM. Nat Protoc *12*, 2478–2492.

Franks, T.M., Benner, C., Narvaiza, I., Marchetto, M.C.N., Young, J.M., Malik, H.S., Gage, F.H., and Hetzer, M.W. (2016). Evolution of a transcriptional regulator from a transmembrane nucleoporin. Gene Dev *30*, 1155–1171.

Franks, T.M., McCloskey, A., Shokirev, M., Benner, C., Rathore, A., and Hetzer, M.W. (2017). Nup98 recruits the Wdr82–Set1A/COMPASS complex to promoters to regulate H3K4 trimethylation in hematopoietic progenitor cells. Gene Dev *31*, 2222–2234.

Frey, S., Richter, R.P., and Görlich, D. (2006). FG-Rich Repeats of Nuclear Pore Proteins Form a Three-Dimensional Meshwork with Hydrogel-Like Properties. Science *314*, 815–817.

Gatticchi, L., Heras, J.I. de las, Roberti, R., and Schirmer, E.C. (2019). Optimization of DamID for use in primary cultures of mouse hepatocytes. Methods *157*, 88–99.

Gozalo, A., Duke, A., Lan, Y., Pascual-Garcia, P., Talamas, J.A., Nguyen, S.C., Shah, P.P., Jain, R., Joyce, E.F., and Capelson, M. (2019). Core Components of the Nuclear Pore Bind Distinct States of Chromatin and Contribute to Polycomb Repression. Mol Cell.

Huang, J., Li, K., Cai, W., Liu, X., Zhang, Y., Orkin, S.H., Xu, J., and Yuan, G.-C. (2018). Dissecting super-enhancer hierarchy based on chromatin interactions. Nat Commun *9*, 943.

Hung, V., Udeshi, N.D., Lam, S.S., Loh, K.H., Cox, K.J., Pedram, K., Carr, S.A., and Ting, A.Y. (2016). Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. Nat Protoc *11*, 456–475.

Ibarra, A., and Hetzer, M.W. (2015). Nuclear pore proteins and the control of genome functions. Gene Dev *29*, 337–349.

Ibarra, A., Benner, C., Tyagi, S., Cool, J., and Hetzer, M.W. (2016). Nucleoporin-mediated regulation of cell identity genes. Gene Dev *30*, 2253–2258.

Iglesias, N., Paulo, J.A., Tatarakis, A., Wang, X., Edwards, A.L., Bhanu, N.V., Garcia, B.A., Haas, W., Gygi, S.P., and Moazed, D. (2019). Native Chromatin Proteomics Reveals a Role for Specific Nucleoporins in Heterochromatin Organization and Maintenance. Mol Cell.

Jacinto, F.V., Benner, C., and Hetzer, M.W. (2015). The nucleoporin Nup153 regulates embryonic stem cell pluripotency through gene silencing. Gene Dev *29*, 1224–1238.

Kadota, S., Ou, J., Shi, Y., Lee, J.T., Sun, J., and Yildirim, E. (2020). Nucleoporin 153 links nuclear pore complex to chromatin architecture by mediating CTCF and cohesin binding. Nat Commun *11*, 2606.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-Cell Dynamics of Genome-Nuclear Lamina Interactions. Cell *153*, 178–192.

Kutay, U., Izaurralde, E., Bischoff, F.R., Mattaj, I.W., and Görlich, D. (1997). Dominant-negative mutants of importin-β block multiple pathways of import and export through the nuclear pore complex. Embo J *16*, 1153–1163.

Labade, A.S., Salvi, A., Karmodiya, K., and Sengupta, K. (2019). Nup93 and CTCF co-modulate spatiotemporal dynamics and function of the HOXA gene cluster during differentiation. Biorxiv 646224.

Labokha, A.A., Gradmann, S., Frey, S., Hülsmann, B.B., Urlaub, H., Baldus, M., and Görlich, D. (2012). Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. Embo J *32*, 204–218.

Liang, Y., Franks, T.M., Marchetto, M.C., Gage, F.H., and Hetzer, M.W. (2013). Dynamic association of NUP98 with the human genome. Plos Genet *9*, e1003308.

Light, W.H., Freaney, J., Sood, V., Thompson, A., D'Urso, A., Horvath, C.M., and Brickner, J.H. (2013). A conserved role for human Nup98 in altering chromatin structure and promoting epigenetic transcriptional memory. Plos Biol *11*, e1001524.

Liu, X., Zhang, Y., Chen, Y., Li, M., Zhou, F., Li, K., Cao, H., Ni, M., Liu, Y., Gu, Z., et al. (2017). In Situ Capture of Chromatin Interactions by Biotinylated dCas9. Cell *170*, 1028-1043.e19.

Lochs, S.J.A., Kefalopoulou, S., and Kind, J. (2019). Lamina Associated Domains and Gene Regulation in Development and Cancer. Cells *8*, 271.

Lowe, A.R., Tang, J.H., Yassif, J., Graf, M., Huang, W.Y.C., Groves, J.T., Weis, K., and Liphardt, J.T. (2015). Importin-β modulates the permeability of the nuclear pore complex in a Ran-dependent manner. Elife *4*, e04052.

Marshall, O.J., Southall, T.D., Cheetham, S.W., and Brand, A.H. (2016). Cell-type-specific profiling of protein–DNA interactions without cell isolation using targeted DamID with next-generation sequencing. Nat Protoc *11*, 1586–1598.

McCloskey, A., Ibarra, A., and Hetzer, M.W. (2018). Tpr regulates the total number of nuclear pore complexes per cell nucleus. Gene Dev *32*, 1321–1331.

Mierzejewska, K., Siwek, W., Czapinska, H., Kaus-Drobek, M., Radlinska, M., Skowronek, K., Bujnicki, J.M., Dadlez, M., and Bochtler, M. (2014). Structural basis of the methylation specificity of R.DpnI. Nucleic Acids Res *42*, 8745–8754.

Milles, S., and Lemke, E.A. (2011). Single Molecule Study of the Intrinsically Disordered FG-Repeat Nucleoporin 153. Biophys J *101*, 1710–1719.

Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., et al. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. Mol Cell *57*, 936–947.

Ori, A., Banterle, N., Iskar, M., Andrés-Pons, A., Escher, C., Bui, H.K., Sparks, L., Solis-Mezarino, V., Rinner, O., Bork, P., et al. (2013). Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. Mol Syst Biol *9*, 648.

Paci, G., Zheng, T., Caria, J., Zilman, A., and Lemke, E.A. (2020). Molecular determinants of large cargo transport into the nucleus. Elife *9*, e55963.

Pascual-Garcia, P., and Capelson, M. (2019). Nuclear pores in genome architecture and enhancer function. Curr Opin Cell Biol *58*, 126–133.

Pascual-Garcia, P., and Capelson, M. (2021). The nuclear pore complex and the genome: organizing and regulatory principles. Curr Opin Genet Dev *67*, 142–150.

Pascual-Garcia, P., Debo, B., Aleman, J.R., Talamas, J.A., Lan, Y., Nguyen, N.H., Won, K.J., and Capelson, M. (2017). Metazoan Nuclear Pores Provide a Scaffold for Poised Genes and Mediate Induced Enhancer-Promoter Contacts. Mol Cell *66*, 63-76.e6.

Paulsen, J., Sekelja, M., Oldenburg, A.R., Barateau, A., Briand, N., Delbarre, E., Shah, A., Sørensen, A.L., Vigouroux, C., Buendia, B., et al. (2017). Chrom3D: three-dimensional genome modeling from Hi-C and nuclear lamin-genome contacts. Genome Biol *18*, 21.

Pott, S., and Lieb, J.D. (2015). What are super-enhancers? Nat Genet *47*, 8–12.

Raices, M., and D'Angelo, M.A. (2012). Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. Nat Rev Mol Cell Bio *13*, 687–699.

Raices, M., and D'Angelo, M.A. (2018). Nuclear Pore Complexes in Genome Organization, Function and Maintenance. 159–182.

Rajanala, K., Sarkar, A., Jhingan, G.D., Priyadarshini, R., Jalan, M., Sengupta, S., and Nandicoori, V.K. (2014). Phosphorylation of nucleoporin Tpr governs its differential localization and is required for its mitotic function. J Cell Sci *127*, 3505–3520.

Ryu, J., Kim, H., Yang, D., Lee, A.J., and Jung, I. (2019). A new class of constitutively active super-enhancers is associated with fast recovery of 3D chromatin loops. Bmc Bioinformatics *20*, 127.

Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*, eaar3958.

Sachani, S.S., Landschoot, L.S., Zhang, L., White, C.R., MacDonald, W.A., Golding, M.C., and Mann, M.R.W. (2018). Nucleoporin 107, 62 and 153 mediate Kcnq1ot1 imprinted domain regulation in extraembryonic endoderm stem cells. Nat Commun *9*, 2795.

Scholz, B.A., Sumida, N., Lima, C.D.M. de, Chachoua, I., Martino, M., Tzelepis, I., Nikoshkov, A., Zhao, H., Mehmood, R., Sifakis, E.G., et al. (2019). WNT signaling and AHCTF1 promote oncogenic MYC expression through super-enhancer-mediated gene gating. Nat Genet *51*, 1723–1731.

Southall, T.D., Gold, K.S., Egger, B., Davidson, C.M., Caygill, E.E., Marshall, O.J., and Brand, A.H. (2013). Cell-Type-Specific Profiling of Gene Expression and Chromatin Binding without Cell Isolation: Assaying RNA Pol II Occupancy in Neural Stem Cells. Dev Cell *26*, 101–112.

Steensel, B. van, and Henikoff, S. (2000). Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase. Nat Biotechnol *18*, 424–428.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell *153*, 307–319.

Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., and Uversky, V.N. (2010). PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. Biochimica Et Biophysica Acta Bba - Proteins Proteom *1804*, 996–1010.