

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


Investigating the Function of Canonical Polycomb Repressive Complex 1 Phase Separation in Epigenetic Regulation

Citation

Oei, Theresa Ann. 2024. Investigating the Function of Canonical Polycomb Repressive Complex 1 Phase Separation in Epigenetic Regulation. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Link

https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37378992

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

HARVARD

Kenneth C. Griffin

GRADUATE SCHOOL
OF ARTS AND SCIENCES

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of

have examined a dissertation entitled

presented by

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature _____

Typed name: Prof.

Signature _____

Typed name: Prof.

Signature _____

Typed name: Prof.

Signature Richard Young

Typed name: Prof.

Signature Robert E Kingston

Typed name: Prof.

Date:

Investigating the Function of Canonical Polycomb Repressive Complex 1 Phase Separation in Epigenetic Regulation

A dissertation presented by

Theresa Ann Oei

to

The Committee on Higher Degrees in Chemical Biology

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the subject of

Chemical Biology

Harvard University

Cambridge, Massachusetts

May 2024

© 2024 Theresa Ann Oei
All rights reserved.

Dissertation Advisor: Robert E. Kingston Theresa A. Oei

Investigating the Function of Canonical Polycomb Repressive Complex 1 Phase Separation in Epigenetic Regulation

Abstract

Polycomb group (PcG) proteins are key developmental regulators involved in gene repression and the maintenance of cell identity. They are a dynamic part of facultative chromatin found clustered in the nucleus as Polycomb bodies. Canonical Polycomb Repressive Complex 1 (cPRC1) is hypothesized to facilitate the silencing of chromatin by compaction and phase separation. The CBX subunit recruits cPRC1 to chromatin and has been shown to contribute to PRC1 condensate formation. The goal of my thesis work was to characterize the phase separation of PRC1 driven by paralogs of the CBX family of proteins and to demonstrate the function of phase separation in PRC1 epigenetic regulation.

In Chapter 1, I will introduce PRC1 and its role in gene repression and development. I will also introduce the concept of liquid liquid phase separation and its proposed role in transcriptional control. Additionally, I will describe the many ways in which PcG proteins, such as the CBXs, are misregulated in cancer.

In Chapter 2, I will describe my characterization of the CBX4 and CBX8 proteins and their phase separation activity both *in vitro* and in cells. I will show how positively charged amino acids in the intrinsically disordered regions of these proteins contribute to their condensate behavior and compare the dynamics of CBX4 and CBX8 to CBX2 in cells.

In Chapter 3, I will share an adaptation of my submitted manuscript demonstrating that

heterologous phase separation can rescue PRC1 function of a CBX2 phase separation mutant in a cancer model system.

In Chapter 4, I will summarize the conclusions from my thesis research and discuss the ways in which this work advances the field of epigenetic regulation. I will also explore potential future directions for further understanding the mechanisms of CBX phase separation in Polycomb-mediated gene repression.

Table of Contents

Title Page i
Copyright ii
Abstract iii
Table of Contents v
List of Figures vii
Dedication viii
Acknowledgements ix
Chapter 1 : Introduction 1
1.1 Polycomb Group Proteins and their role in mammalian development 2
1.2 Polycomb bodies and liquid liquid phase separation in transcriptional control 9
1.3 PcG proteins in cancer 15
References 19
Chapter 2 : Characterizing the phase separation behavior of CBX4 and CBX8 28
2.1 Introduction to CBX4 and CBX8 29
2.2 CBX4 and CBX8 form droplets in vitro and condensates in cells 34
2.3 CBX4 and CBX8 drive condensate formation in cells 36
2.4 Conclusions and future directions 39
2.5 Materials and methods 45
References 50
Chapter 3 : Phase separation drives Polycomb-mediated gene repression 56
3.1 Summary and acknowledgements 57
3.2 Introduction to CBX2 in Polycomb epigenetic regulation 58
3.3 Heterologous IDRs induce orthogonal phase separation of Polycomb condensates 59
3.4 Overexpression of $CBX2_{KRA}$ fails to repress neurite elongation 66
3.5 The FUS^{IDR} rescues $CBX2_{KRA}$ in a neural differentiation phenotype 69
3.6 The FUS^{IDR} rescues transcriptional repression of neurogenesis genes 72
3.7 PRC1 target loci are mostly identical in the wildtype, mutant and rescue cell lines 78
3.8 FUS^{IDR} fusion rescues gene repression of bound genes 79
3.9 FUS^{IDR} fusion rescues a CBX2-dependent breast cancer phenotype 83
3.10 Conclusions 85
3.11 Materials and methods 88
References 98

Chapter 4 : Conclusions and future directions 107
***4.1 CBX2, CBX4, and CBX8 drive phase separation of Polycomb Repressive Complex 1* 108**
4.2 Model for Polycomb body phase separation 114
4.3 Mechanism of Polycomb-mediated gene repression by phase separation 117
4.4 Polycomb condensates in maintenance and memory 125
4.5 Polycomb condensates in disease 130
References 133

List of Figures

Figure 1.1 Schematic of cPRC1 5

Figure 1.2 The CBX family of proteins 8

Figure 1.3 Phase separation in the nucleus 10

Figure 1.4 CBX2 compaction and phase separation mutant 14

Figure 2.1 Predicted protein structure for CBX2, CBX4, and CBX8 31

Figure 2.2 CBX4 and CBX8 in vitro assays 35

Figure 2.3 CBX4 and CBX8 form condensates in cells 38

Figure 3.1 Heterologous IDRs can induce Polycomb bodies 63

Figure 3.2 Puncta quantification, immunofluorescence, and neurite assay development ... 65

Figure 3.3 Heterologous IDRs can rescue a CBX2 mutant phenotype 68

Figure 3.4 Control experiments for the neurite assay 71

Figure 3.5 Heterologous phase separation rescues gene repression 76

Figure 3.6 The FUS^{IDR} and $CBX2^{IDR}$ both silence transcription 77

Figure 3.7 The FUS^{IDR} rescues bound Polycomb target genes 81

Figure 3.8 Genome-wide H3K27me3 and H2AK119ub1 binding 82

Figure 3.9 The FUS^{IDR} can rescue a CBX2 breast cancer phenotype 84

Figure 4.1 Model for CBX phase separation 110

Figure 4.2 Model for the FUS^{IDR} rescue of the $CBX2_{KRA}$ mutant 119

Figure 4.3 Model for CBX2 phase separation in epigenetic memory 128

Dedicated to my parents, Ellen McHugh and Paul Oei, the ones who first sparked my curiosity for scientific discovery.

Acknowledgements

This dissertation would not have been possible without the following people. First, I would like to thank Bob Kingston for his support and guidance throughout my time in the lab. I am grateful for his mentorship and influence in thinking critically about the mechanisms of biology. The opportunity to work on problems I was excited about and learn by doing (and failing!) has been a tremendous gift. I would also like to thank the entire Kingston lab, past and present, for their support and friendship. From Bay 2 coffee to sharing protocols, from pandemic shift schedules to practice talks, from troubleshooting experiments to group runs along the Esplanade, I feel so grateful to have learned from each of you along the way. I would particularly like to thank Philipp Schneider and Stefan Niekamp for their collaborative discussions, Elizabeth Jaensch and MacKenzie Mauger for sharing the highs and lows of graduate school, my fellow bay mates - Jongmin Kim and Christos Tsokos - for their inestimable wisdom, and Sharon Marr for always keeping the lab running smoothly.

I would like to thank my Dissertation Advisory Committee: Brian Liau, Stuart Schreiber, and Mitzi Kuroda, for their thoughtful advice throughout this work. I would also like to thank Rick Young for joining my thesis defense committee. I have often been inspired by his pioneering work on condensates in biology.

I would like to specially thank a few of the many mentors I have had throughout my studies and scientific training. As an undergraduate, I learned so much from Joan Steitz as a research assistant in her lab about the relentless pursuit of scientific questions. Working together with Eric Guo in Joan's lab, was my first real foray into academic research and I am thankful for his

mentorship and support. I am thankful for the opportunity to have worked with Sara Jones in Feng Zhang's lab at MIT. From Sara and Feng, I learned the importance and challenge of working on high-risk scientific projects. I would like to thank Robert Shulman, who deepened my appreciation for epigenetic mechanisms and the strong connection between philosophy and scientific discovery.

I would also like to thank the Chemical Biology Program including Jason Millberg, Dan Kahne, and Suzanne Walker. I joined the Chem Bio program because I felt the concern and care each of you felt towards the students in the program and I have deeply appreciated your encouragement.

I am thankful for my Chem Bio cohort and the graduate school friends I have made because of my time at Harvard. It would have been so much less fun without all of you.

To my roommates throughout the years and college friends, your love and emotional support has been a light throughout the inevitable challenges of graduate school. I am so grateful for your friendship today and always.

Finally, I would like to thank my family for their constant love and support. To my parents, your sacrifices throughout my entire education have made this day possible. I would not be so excited about science without the years of homeschooling in the basement or science fair projects in the garage. To my siblings, Bridget, Maura, and John Paul, growing up with you has been an incredible journey of discovery and I cannot wait to see the worlds we will explore in the future.

And last, but certainly not least, I would like to thank my fiancé, Ben. From our earliest days in Organic Chemistry, a shared love for science has been a foundational part of our relationship. Thank you for your belief in this dream of mine and your loving presence every step along the way.

Chapter 1 : Introduction

1.1 Polycomb Group Proteins and their role in mammalian development

The term epigenetics was first coined by Conrad Waddington in 1948 to describe the "whole complex of developmental processes" connecting genotype to phenotype during embryogenesis.[1] In the decades since, this concept has been further developed to include the regulation of gene expression during cellular differentiation at a molecular level.[1,2] The development of an organism from a single cell with a single genome requires the proper spatial and temporal regulation of gene expression.[3] This requires both activation of lineage specific genes and the maintenance of repression at non-lineage genes.[3,4] The Polycomb Group (PcG) proteins have important roles in silencing non-lineage specific genes during differentiation and maintaining cell identity throughout the organism over many generations of cell division.[5] A year before Waddington defined epigenetics, Pam Lewis discovered the first PcG protein in *Drosophila melanogaster*.[6] Mutations in the *Pc* gene (a homolog of the mammalian CBX proteins) resulted in disrupted body segmentation of flies.[7] Consistent with their important role in development, PcG proteins are functionally conserved from *Drosophila melanogaster* to *Homo sapiens*.[8,9] Mammalian PcG complexes have diversified with the expansion of this family of proteins reflecting the evolved complexity of mammalian development.[8,9] PcG proteins are often essential; gene ablation leads to failures in gastrulation or to early embryonic or post-natal lethality.[10-14] However, there is known to be some redundancy amongst paralogs, suggesting some overlapping function.[9] Overall, the PcG proteins play a vital role in gene repression and cellular memory in mammalian development. Additionally, the diversity of mammalian PcG complexes maintaining silent chromatin states enables the tunability of the epigenetic landscape as a fundamental mechanism of developmental biology.[8,9]

Polycomb-mediated gene repression in mammals is accomplished through the interplay of two major complexes: PRC1 and PRC2.[15] In the current working model, unmethylated CpG-enriched DNA near gene promoters, known as CpG Islands (CGIs), is recognized by the KDM2B protein or Polycomb Like proteins (PCLs), which recruit noncanonical PRC1 (ncPRC1).[16,17] Other Polycomb recruitment methods include RNA dependent targeting.[3] For example, X chromosome inactivation involves the recruitment of ncPRC1 by the lncRNA, Xist, to initiate Polycomb silencing of the inactive X.[3] The RING1 subunit of ncPRC1 is a ubiquitin E3 ligase responsible for depositing a ubiquitin group at Histone 2A lysine 119 (H2AK119ub1).[3] This histone mark can recruit PRC2, whose EZH subunit is a histone methyltransferase that deposits mono, di, and trimethylation at Histone 3 lysine 27 (H3K27) in broad domains. Canonical PRC1 (cPRC1_ binds to H3K27me3 and compacts chromatin to maintain a silent, repressed state and prevent transcription machinery from accessing target loci.[3,4,10] Changes to genome structure upon cPRC1 binding occur at multiple levels. There is local compaction of the promoter region alongside the establishment of Polycomb Associated Domains (PADs), which can span many kilobases to form higher orders of chromatin structure.[18-20] These self-interacting H3K27me3 and cPRC1-defined domains can change throughout the process of differentiation to affect gene regulation.[18-20] The hierarchical model of Polycomb recruitment is nuanced by the feedback provided between complexes.[3] PRC2 can spread along the chromatin by recognizing its own H3K27me3 mark through the EED subunit and further stimulating the catalytic EZH subunit.[21] Canonical PRC1 can interact with the EED subunit, potentially recruiting PRC2 to stimulate more H3K27me3 deposition.[22] The interplay of Polycomb complexes reinforces gene repression and facilitates epigenetic memory. The repression of Polycomb target genes is heritable: the epigenetic memory of repression passes on to daughter cells to maintain cell identity.[3,23]

Although the mechanisms of epigenetic memory are still being worked out, the necessity of maintaining gene repression over many cellular divisions is undeniable for the maintenance of healthy tissues.

The focus of my dissertation is canonical PRC1, which is composed of four main subunits: PHC, RING1, PCGF, and CBX (**Fig. 1.1**).[10,24] Each subunit exists as a family of proteins with several paralogs, which can be included in the complex diversifying the possible complexes.[10,24] Different paralogs are expressed at different levels in distinct tissues and at different times during development. For example, CBX7 is highly expressed in embryonic stem cells, but as cells differentiate CBX7 expression goes down and CBX2/4/8 expression goes up.[10] The PHC protein has a SAM domain that has important roles in oligomerization to generate the long-range chromatin contacts found in PADs.[20,25] Mutations in the SAM domain disrupt higher order chromatin structure *in vivo*.[20,25] The RING protein is an E3 ubiquitin ligase, but it is known to be more enzymatically active when it is part of the noncanonical form of PRC1, therefore canonical PRC1 has very little ubiquitylation activity.[26] PCGF2/4 are the main PCGF proteins found in cPRC1 and structurally form the heart of the complex bringing the other subunits together.[24] The CBX protein is specific to cPRC1 and includes a chromodomain that binds to the H3K27me3 mark and helps to localize cPRC1 to target genes.[24] However, the affinity for the mark varies among the different CBX paralogs and there are likely other mechanisms of DNA targeting at work.[27,28]

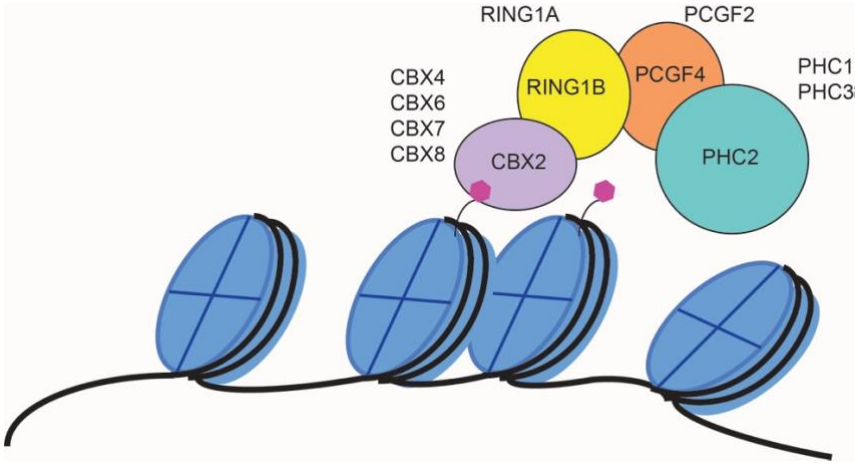


Figure 1.1 Schematic of cPRC1

Canonical PRC1 has four core subunits and binds H3K27me3 via the chromodomain of the CBX protein to compact heterochromatin.

My thesis work has focused on the CBX family of proteins, specifically CBX2, CBX4, and CBX8. Besides the chromodomain, the CBX proteins have a C box domain that is responsible for interacting with RING1B and the rest of the cPRC1 complex (**Fig 1.2A**).[23] CBX2 also has an AT-Hook motif and the other CBXs have AT-Hook like motifs with DNA binding capacity (**Fig 1.2A**).[23] CBX2, CBX4, and CBX8 each have an Intrinsically Disordered Region (IDR) at the center of the protein that is highly charged (**Fig. 1.2B**). This region in CBX2 has been shown to drive *in vitro* chromatin compaction and condensate formation in cells, thus it has been referred to as the Compaction and Phase Separation (CaPS) region.[29,30] CBX7 is a much smaller paralog lacking this disordered region (**Fig. 1.2A**). In line with these observations, CBX7 cannot compact nucleosomes or form condensates *in vitro*.[29,30] Interestingly, CBX7 is highly expressed in embryonic stem cells, where bivalent chromatin facilitates a plastic chromatin environment.[31] It is hypothesized that the transition from CBX7 to CBX2, CBX4, and CBX8 in more differentiated cell types corresponds with the repression of non-lineage specific genes in a more stable, and heritable manner.[10] Accordingly, introduction of the CaPS region of CBX2 into the CBX7 protein impaired mESC differentiation and the proper activation of lineage specific genes.[31] Phase separation of CBX7 can be induced *in vitro* with the insertion of CBX2's CAPS domain and this leads to changes in gene expression and differentiation.[31] This suggests that phase separation may be related to the biology of nucleosome compaction and gene silencing.[29-31] CBX6 is the least-studied CBX protein. It is highly expressed in embryonic stem cells and has recently been shown to be essential for pluripotency maintenance.[32] CBX6 is also highly disordered, but it can associate with both ncPRC1 and cPRC1 and has low affinity for H3K27me3.[32] Characterizing the ability of CBX4 and CBX8 to phase separate will enable

greater understanding of these proteins and their role in heterochromatin formation while further exploring the relationship between phase separation and compaction.

There are many tools for characterizing intrinsically disordered proteins based on their amino acid sequence. The plots in **Fig. 1.2B** compare disorder and charge between CBX2, CBX4, and CBX8 as predicted by the PONDR VSL2 algorithm and the MobiDB algorithm.[33-35] CBX4 and CBX8 were hypothesized to phase separate based on their ability to compact nucleosomes and their similarly disordered regions in the central domain of the protein when compared to CBX2. Therefore, this research has focused on the role of CBX2, CBX4, and CBX8 in cPRC1-mediated epigenetic regulation. In Chapter 2, I will discuss the role of the CaPS region in the phase separation behavior of the CBX proteins and compare the tunability of condensation amongst CBX2, CBX4 and CBX8.

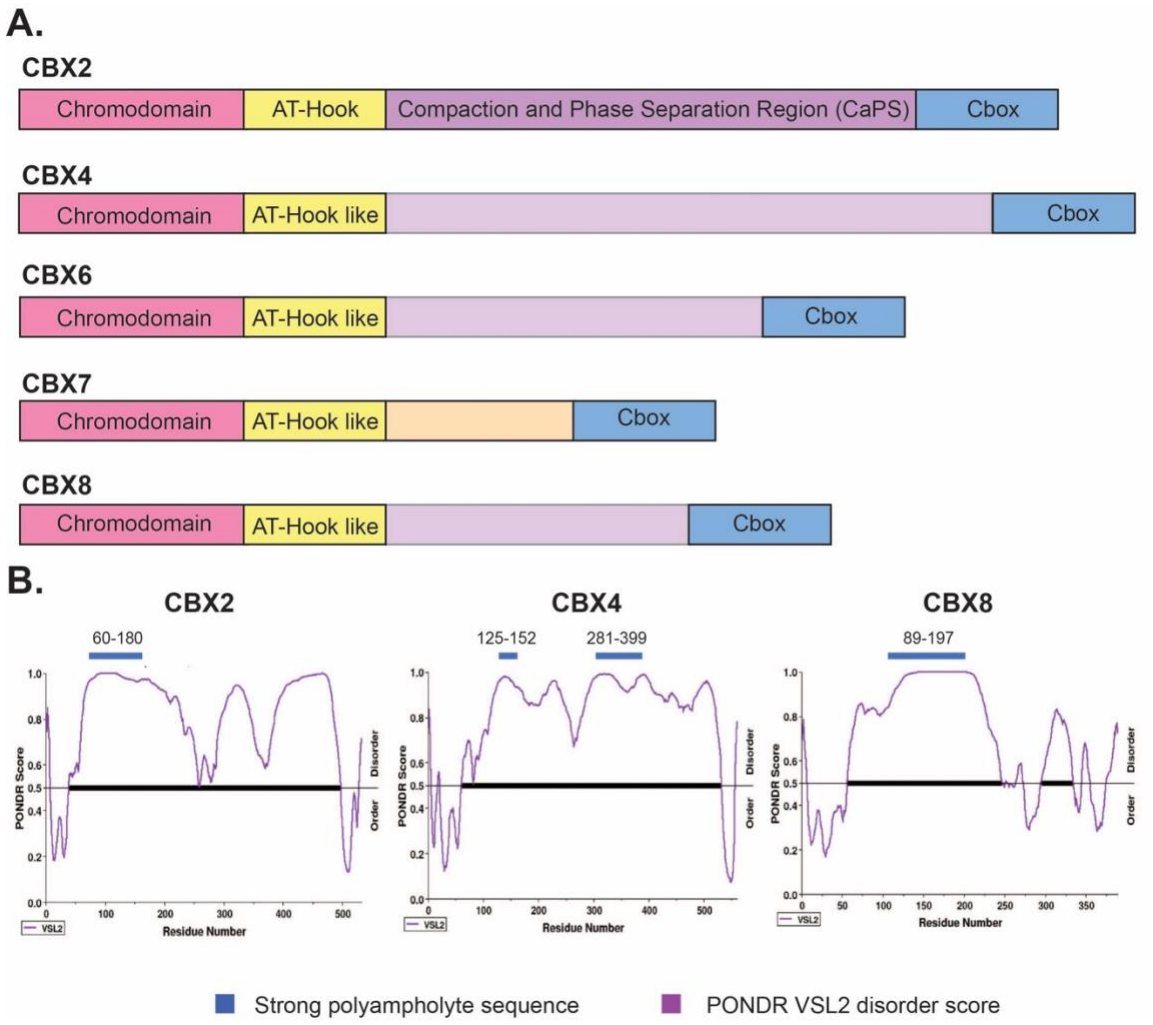


Figure 1.2 The CBX family of proteins

***The CBX family contains paralogs with diverse disordered regions*. (A)** Schematic of the protein domains in the CBX family of paralogs. **(B)** PONDR VSL2 disorder plots highlighting strong polyampholyte sequences in CBX2, CBX4 and CBX8.[33-35]

1.2 Polycomb bodies and liquid liquid phase separation in transcriptional control

Polycomb defined chromatin domains enriched with H3K27me3, H2AK119ub1 and Polycomb complexes are thought to condense within nuclear foci.[36] These foci, or Polycomb bodies, were first described in 1998 and are one type of biomolecular condensate found in the cell (**Fig. 1.3.**)[37,38] There are many membraneless compartments throughout the nucleus thought to form through liquid liquid phase separation, a physical phenomenon that occurs due to thermodynamic favorability driven by homotypic and heterotypic multivalent interactions between macromolecules.[38] Proteins and nucleic acids can phase separate from the surrounding milieu of the nucleus or cytoplasm as liquid droplets that concentrate certain factors and exclude others.[38] These membraneless organelles in the nucleus are important for compartmentalizing biological processes.[38] For example, the nucleolus is a biomolecular condensate involved in ribosome biogenesis, whereas nuclear speckles have a role in RNA splicing.[39,40] Phase separation in the cell is specific and controlled since the viscoelasticity of liquid droplets can be tuned by molecular actors such as DNA, RNA, protein concentration, post-translational modifications and more.[41]

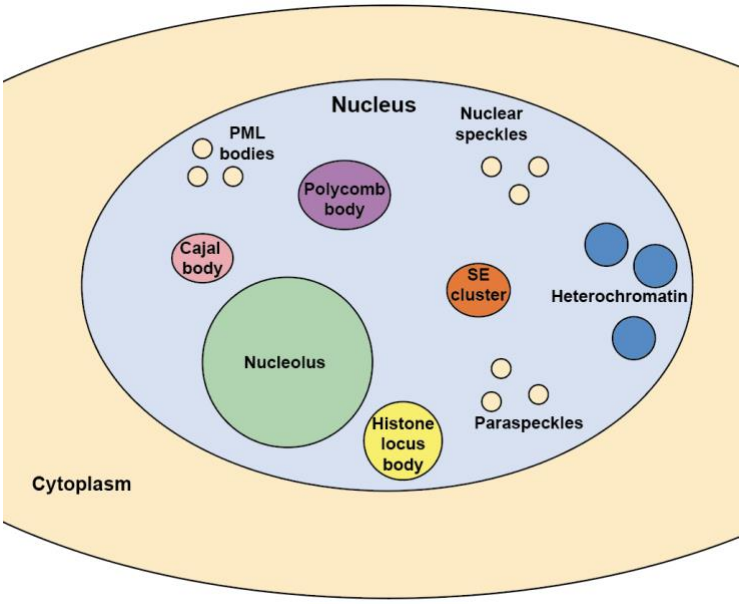


Figure 1.3 Phase separation in the nucleus

Phase separated condensates form membraneless organelles in the nucleus including Polycomb bodies. The nucleolus houses ribosomal RNA transcription and processing.[38] Cajal bodies are important for snRNP biogenesis.[38] PML bodies play a role in DNA damage and repair.[38] Nuclear speckles are involved in mRNA processing.[38] SE clusters bring together superenhancers and transcription machinery.[38] Heterochromatin maintains gene repression.[38] Paraspeckles contain noncoding RNA and proteins involved in the stress response.[38] The histone locus body is involved in histone mRNA transcription and processing.[38]

Biomolecular condensates can compartmentalize distinct cellular processes within the cell, while also facilitating the dynamics of biological interactions. In the context of chromatin, phase separation may be a mechanism for concentrating activating or silencing factors to initiate or maintain transcriptional states.[42,43] It is an attractive model for transcriptional control supported by a wide variety of recent work: RNA Polymerase II and Mediator phase separation facilitate transcriptional bursting and super enhancer condensation;[44-46] HP1α forms liquid droplets *in vitro* and stable, silent chromatin compartments *in vivo*;[47-50] nucleosomes demonstrate liquid-like properties as the foundation for chromatin dynamics;[51-53] and the chromatin remodeler, cBAF, forms condensates that mediate partner recruitment for transcriptional activation.[54] Within the Polycomb field, our group and others have furthered this paradigm demonstrating a potential role for condensation in Polycomb-driven gene silencing.[30,55] In the case of Polycomb bodies, the concentration of cPRC1 in biomolecular condensates may generate a repressed chromatin state and facilitate epigenetic memory by stabilizing H3K27me3 and chromatin compaction.[30,55] It has been shown that phase separation can be involved in the catalysis of histone modifications such as ubiquitination on H2B.[56] Similarly, Polycomb body formation may be important in driving H3K27me3 spreading, an important aspect of epigenetic inheritance. The concentration of PcG proteins in biomolecular condensates may also facilitate the reestablishment of repression after DNA replication and cell division on the inherited parental nucleosomes. However, the Polycomb phase separation model is not without controversy. First, some have argued that the number of cPRC1 complexes quantified by single molecule imaging in Polycomb bodies is lower than the proposed saturation point required for true phase separation.[57] Additionally, CBX8-cPRC1 chromatin condensates seem to be porous and accessible as measured by cryo-electron tomography contradicting the concept of stable, inaccessible Polycomb compartments in

the nucleus.[58] Recent work has described the mechanism of Polycomb repression as spatial feedback rather than phase separation or compaction due to the high amount of intermingling of Polycomb bound chromatin with A or B compartments.[59] Overall, this might suggest a view of Polycomb bodies as very dynamic rather than stable globule-type droplets. Additionally, the role of histone marks as the main encoding of repression rather than cPRC1 compaction has also been proposed.[60] These challenges to the model of cPRC1 phase separation are important to consider as our work further explores the mechanisms and function of phase separation in Polycomb-mediated gene repression.

Canonical PRC1 condensate formation is driven by two subunits of cPRC1: PHC and CBX. The *Drosophila* Ph subunit has been shown to phase separate and recruit PRC1 from cell extracts via its polymerizing SAM domain while promoting PRC1's ubiquitin ligase activity.[61] Mammalian PHC phase separation is also influenced by its polymerization activity, which enables the generation of stable chromatin condensates *in vitro*.[62] In contrast, CBX was shown to be more important for initiating condensation of cPRC1 *in vitro*.[62] Phase separation of CBX2 is driven by the CaPS region and more specifically by the positive lysine and arginine residues sprinkled through the IDR.[30] The mutation of twenty-three lysines and arginines in the CBX2 IDR to alanine ablated phase separation activity *in vitro* and greatly reduced the formation of puncta in cells.[30] These same residues are simultaneously required for cPRC1's nucleosome compaction activity *in vitro*.[29] Significantly, mutating these residues in mice causes axial patterning defects suggesting an important functional role for the CaPS region in mammalian gene expression (**Fig. 1.4**).[63] In Chapter 2, I will characterize the other CBX proteins containing a CaPS-like IDR region - CBX4 and CBX8 - and investigate the role of positive lysine and arginine residues in

their condensate formation activity. In Chapter 3, I will return to the $CBX2_{KRA}$ phase separation mutant and demonstrate functional rescue of the mutant at a phenotypic and molecular level using heterologous IDRs to interrogate phase separation activity as a mechanism of cPRC1 gene repression. The biophysical concept of phase separation offers a Kuhnian paradigm shift for how we understand chromatin and the biological mechanisms of transcriptional control. My thesis research has attempted to contribute to this paradigm by characterizing the phase separation of cPRC1 and establishing the significance of its role in Polycomb-mediated transcriptional silencing.

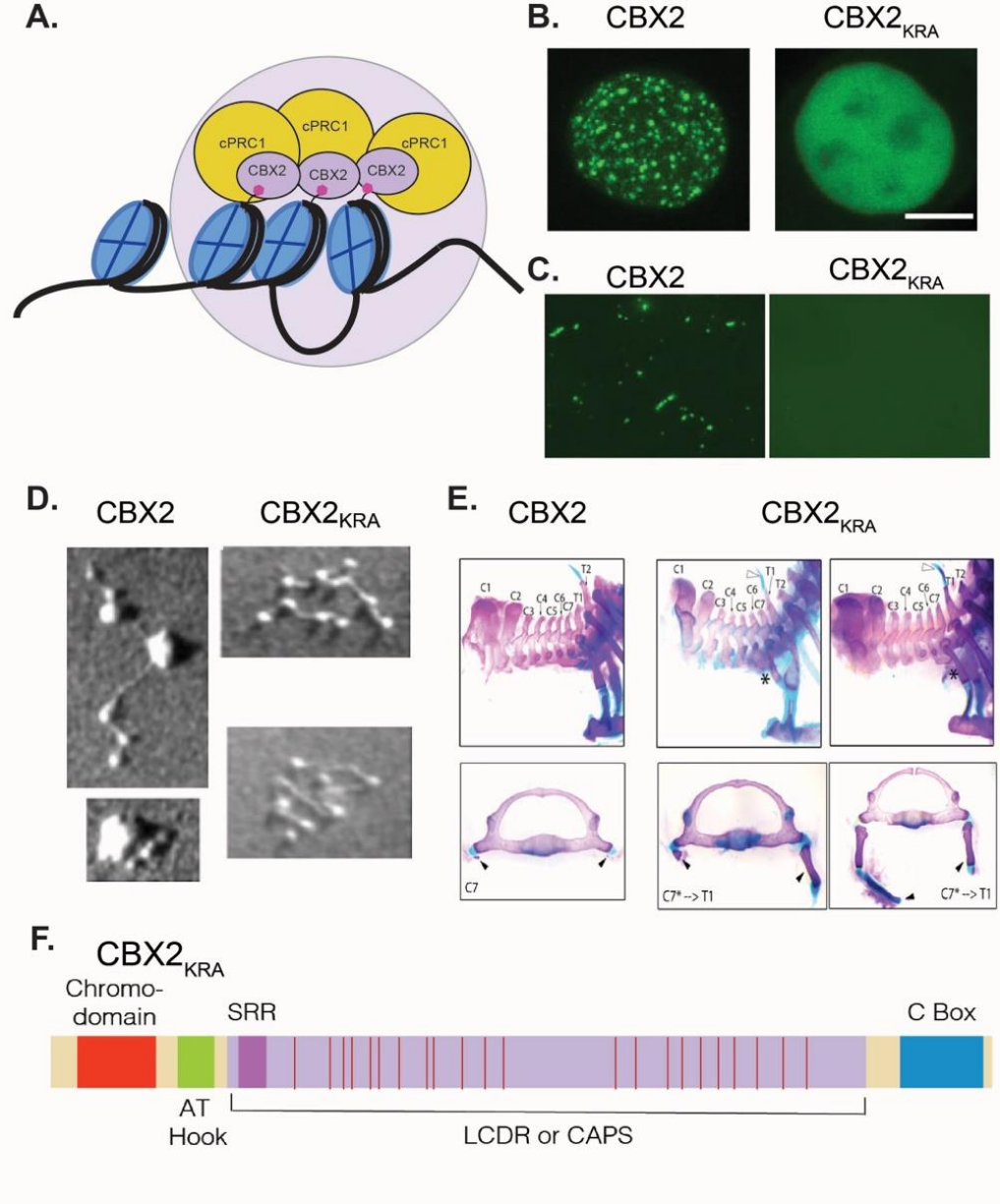


Figure 1.4 CBX2 compaction and phase separation mutant

Positively charged residues in the disordered region of CBX2 are important for phase separation, in vitro nucleosome compaction, and axial patterning in mice. **(A)** Schematic of the proposed model of CBX2-driven phase separation and compaction of nucleosomes. **(B)** Z-stacks of live cell imaging of mEGFP-CBX2 and mEGFP-CBX2$_{KRA}$ expressed in NIH-3T3 cells. Scale bar is 10 μm.[30] **(C)** *In vitro* droplet assay for CBX2 and CBX2$_{KRA}$ purified as a heterodimer with mEGFP-RING1B.[30] **(D)** Scanning electron microscopy of CBX2 and CBX2$_{KRA}$ with nucleosomal arrays.[29] **(E)** Axial patterning defect in mice expressing the mutant CBX2$_{KRA}$.[63] **(F)** Schematic of the domains of CBX2 highlighting the twenty-three lysine and arginine residues mutated in the CBX2$_{KRA}$.

1.3 PcG proteins in cancer

Understanding the mechanisms of phase separation for PcG bodies will also have significant implications for disease. As regulators of development and cell identity, the Polycomb group of proteins have long been associated with a myriad of cancers.[64] PcG proteins can silence tumor suppressor genes and they have been postulated as crucial regulators of cancer stem cell populations.[64] Gain-of-function and loss-of-function mutations in PcG proteins alongside misregulation of expression have all been linked to cell transformation.[64] Importantly, the role of PcG proteins is highly dependent upon the cellular context. In distinct cancers, PcG proteins can have oncogenic or tumor-suppressing effects.[64] Therefore, developing cancer therapeutics targeting PcG proteins will not have a one-size-fits-all strategy. A significant amount of cancer research into PcG proteins has focused on PRC2 and global changes in H3K27me3.[65] EZH2, the catalytic subunit of PRC2 responsible for writing the H3K27me3 mark, has gain-of-function mutations in multiple cancers that lead to hyperactivity.[64,65] Tazemetostat, the first EZH2 inhibitor, was recently approved for epithelioid sarcoma and follicular lymphoma.[66] EZH2 inhibitors are also in clinical trials for several other types of cancers, including in combination with other therapeutics.[67]

PRC1 has also been linked to cancer development. Significantly, PRC1 proteins are less prone to mutation in cancer compared to PRC2, rather changes in expression levels are the bigger culprit.[64] PCGF4, also known as BMI1, is often highly expressed in cancer and can be a biomarker for certain malignancies.[68] PCGF4 regulates the cancer stem cell niche mainly through increased repression of the Ink4a/Arf locus, which encodes two tumor suppressors.[68] PCGF4 also has roles in other hallmarks of cancer such as proliferation, anti-apoptosis, and

metastasis.[64] PCGF2, or MEL18, has been shown to have an antagonistic relationship with PCGF4 in cancer.[64] PCGF2 acts as a tumor suppressor by regulating c-Myc expression.[64] Therefore, downregulation of PCGF2 promotes tumorigenesis and PCGF2 is lowly expressed in several cancers including gastric cancer.[64,69] RING1B can be upregulated in certain cancers, however it has also been shown to have potential activation effects at target genes contrary to canonical mechanisms of PRC1 repression.[70] In additional complexity, RING1B's expression in breast cancer can be differentially correlated with metastasis or survival depending upon the cancer type.[70]

The CBX family of proteins are also misexpressed in cancer.[64] CBX2 is highly expressed in brain cancers including neuroblastoma and gliomas.[71,72] Neuroblastoma is thought to be dependent on CBX2 expression to promote tumorigenesis according to the Cancer Dependency Map.[73] CBX2 regulates the cancer stem cell population and promotes invasion through the Akt/PI3K signaling pathway in brain tumors.[71] CBX2 is also highly expressed in Triple Negative Breast Cancers (TNBCs), where it contributes to tumorigenesis through the same Akt/PI3K pathway.[74] CBX2 promotes the invasiveness of breast cancer cells contributing to an aggressive, metastatic phenotype.[74] High CBX2 expression also strongly correlates with poor prognosis in hepatocellular carcinoma and is a potential therapeutic target in gastric cancer as well.[75] CBX4 is also highly expressed in gastric cancer with a role in maintaining the cancer stem cell niche via expression of CDC20, which promotes stemness of the cancer cells.[76] Additionally, CBX4's sumoylation activity drives HIF-1a's expression in hepatocellular carcinoma to support increased angiogenesis.[77] Pan-cancer meta-analysis for CBX8 demonstrates high correlation with carcinogenesis and poor prognosis.[78] CBX8 is highly expressed in breast cancer and in

leukemia.[79,80] In hepatocellular carcinoma, CBX8 upregulates the expression of Bmp4, a known oncogene, and thereby induces the epithelial-mesenchymal transition contributing to metastasis.[81] CBX6 and CBX7 both have more complicated, context-dependent roles in cancer development. Alongside CBX2, CBX4, and CBX8, CBX6 is upregulated in hepatocellular carcinoma, however it is downregulated in glioblastoma and plays a tumor-suppressive role in that cellular context.[64] Similarly, CBX7 has a tumor-suppressing role in breast cancer or colon cancer where it is lowly expressed.[82] However, in gastric cancer or B-cell lymphomas, CBX7's increased expression promotes cancer cell stemness and proliferation.[82] Overall, the CBX family of proteins may be important cancer biomarkers and potential therapeutic targets.[83] Small molecules are being developed to target the chromodomain of the CBX proteins to prevent binding to H3K27me3 with sufficient specificity.[84] While the CBX proteins clearly play a role in cancer development, the mechanisms of their contributions to the hallmarks of cancer are under-explored. This research and other studies will be significant for furthering our understanding of this family of proteins as therapeutic targets for drug development.

Generally, CBX2, CBX4, and CBX8 tend to be upregulated in cancer suggests that increased phase separation activity caused by increased protein concentration in the nucleus may be linked to a functional role in promoting tumorigenesis. Given the emerging role of the CBX family in cancer development, we chose to use several cancer cell lines throughout this research. We use a neural differentiation assay with the mouse neuroblastoma cell line, Neuro2a, to demonstrate the role of CBX2 phase separation in driving the cancer stem cell promoting phenotype of CBX2 in neuroblastoma. Additionally, we study CBX2 in a TNBC cell line, 4T1, to examine its role in driving an invasive phenotype. Overall, greater understanding of the relationships among PcG

protein misexpression, H3K27me3, and Polycomb body formation in establishing and maintaining silent chromatin will facilitate the development of new therapeutic opportunities in oncology.

References

1. Waddington, C. H. (2012). The Epigenotype. *International Journal of Epidemiology, 41*(1), 10-13. https://doi.org/10.1093/ije/dyr184.

2. Deichmann U. (2016). Epigenetics: The origins and evolution of a fashionable topic. *Developmental biology*, *416*(1), 249–254. https://doi.org/10.1016/j.ydbio.2016.06.005

3. Schuettengruber, B., Bourbon, H. M., Di Croce, L., and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell*, *171*(1), 34–57. https://doi.org/10.1016/j.cell.2017.08.002

4. Blackledge, N. P., and Klose, R. J. (2021). The molecular principles of gene regulation by Polycomb repressive complexes. *Nature reviews. Molecular cell biology*, *22*(12), 815–833. https://doi.org/10.1038/s41580-021-00398-y

5. Loh, C. H., and Veenstra, G. J. C. (2022). The Role of Polycomb Proteins in Cell Lineage Commitment and Embryonic Development. *Epigenomes*, *6*(3), 23. https://doi.org/10.3390/epigenomes6030023

6. Lewis, P. H. (1947). Melanogaster-New mutants: Report of Pamela H. Lewis. *Drosophila Information Service, 21,* 69.

7. Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature, 276,* 565-570.

8. Sowpati, D. T., Ramamoorthy, S., and Mishra, R. K. (2015). Expansion of the polycomb system and evolution of complexity. *Mechanisms of development*, *138 Pt 2*, 97–112. https://doi.org/10.1016/j.mod.2015.07.013

9. Kim, J. J., and Kingston, R. E. (2022). Context-specific Polycomb mechanisms in development. *Nature reviews. Genetics*, *23*(11), 680–695. https://doi.org/10.1038/s41576-022-00499-0

10. Ma, R. G., Zhang, Y., Sun, T. T., and Cheng, B. (2014). Epigenetic regulation by polycomb group complexes: focus on roles of CBX proteins. *Journal of Zhejiang University. Science. B*, *15*(5), 412–428. https://doi.org/10.1631/jzus.B1400077

11. Voncken, J. W., Roelen, B. A., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., and van Lohuizen, M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(5), 2468-2473. https://doi.org/10.1073/pnas.0434312100

12. O'Carroll, D., Erhardt, S., Pagani, M., Barton, S. C., Surani, M. A., and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. *Molecular and cellular biology*, *21*(13), 4330–4336. https://doi.org/10.1128/MCB.21.13.4330-4336.2001

13. Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. *Development*, *121*(2), 273–285. https://doi.org/10.1242/dev.121.2.273

14. Pasini, D., Bracken, A. P., Jensen, M. R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *The EMBO journal*, *23*(20), 4061–4071. https://doi.org/10.1038/sj.emboj.7600402

15. Aranda, S., Mas, G., and Di Croce, L. (2015). Regulation of gene transcription by Polycomb proteins. *Science advances*, *1*(11), e1500737. https://doi.org/10.1126/sciadv.1500737

16. Farcas, A. M., Blackledge, N. P., Sudbery, I., Long, H. K., McGouran, J. F., Rose, N. R., Lee, S., Sims, D., Cerase, A., Sheahan, T. W., Koseki, H., Brockdorff, N., Ponting, C. P., Kessler, B. M., and Klose, R. J. (2012). KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *eLife*, *1*, e00205. https://doi.org/10.7554/eLife.00205

17. Li, H., Liefke, R., Jiang, J., Kurland, J. V., Tian, W., Deng, P., Zhang, W., He, Q., Patel, D. J., Bulyk, M. L., Shi, Y., and Wang, Z. (2017). Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature*, *549*(7671), 287–291. https://doi.org/10.1038/nature23881

18. Du, Z., Zheng, H., Kawamura, Y. K., Zhang, K., Gassler, J., Powell, S., Xu, Q., Lin, Z., Xu, K., Zhou, Q., Ozonov, E. A., Véron, N., Huang, B., Li, L., Yu, G., Liu, L., Au Yeung, W. K., Wang, P., Chang, L., Wang, Q., He, A., Sun, Y., Na, J., Sun, Q., Sasaki, H., Tachibana, K., Peters, A. H. F. M., and Xie, W. (2020). Polycomb Group Proteins Regulate Chromatin Architecture in Mouse Oocytes and Early Embryos. *Molecular cell*, *77*(4), 825–839.e7. https://doi.org/10.1016/j.molcel.2019.11.011

19. Guo, Y., and Wang, G. G. (2022). Modulation of the high-order chromatin structure by Polycomb complexes. *Frontiers in cell and developmental biology*, *10*, 1021658. https://doi.org/10.3389/fcell.2022.1021658

20. Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J. T., Sadreyev, R. I., Dekker, J., and Kingston, R. E. (2017). Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Molecular cell*, *65*(3), 432–446.e5. https://doi.org/10.1016/j.molcel.2017.01.009

21. Kraft, K., Yost, K. E., Murphy, S. E., Magg, A., Long, Y., Corces, M. R., Granja, J. M., Wittler, L., Mundlos, S., Cech, T. R., Boettiger, A. N., and Chang, H. Y. (2022). Polycomb-mediated genome architecture enables long-range spreading of H3K27 methylation. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(22), e2201883119. https://doi.org/10.1073/pnas.2201883119

22. Cao, Q., Wang, X., Zhao, M., Yang, R., Malik, R., Qiao, Y., Poliakov, A., Yocum, A. K., Li, Y., Chen, W., Cao, X., Jiang, X., Dahiya, A., Harris, C., Feng, F. Y., Kalantry, S., Qin, Z. S., Dhanasekaran, S. M., and Chinnaiyan, A. M. (2014). The central role of EED in the

orchestration of polycomb group complexes. *Nature communications*, *5*, 3127. https://doi.org/10.1038/ncomms4127

23. Kim, J., and Kingston, R. E. (2020). The CBX family of proteins in transcriptional repression and memory. *Journal of biosciences*, *45*, 16.

24. Geng, Z., and Gao, Z. (2020). Mammalian PRC1 Complexes: Compositional Complexity and Diverse Molecular Mechanisms. *International journal of molecular sciences*, *21*(22), 8594. https://doi.org/10.3390/ijms21228594

25. Isono, K., Endo, T. A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B. E., and Koseki, H. (2013). SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. *Developmental cell*, *26*(6), 565–577. https://doi.org/10.1016/j.devcel.2013.08.016

26. Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M., and Sixma, T. K. (2006). Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. *The EMBO journal*, *25*(11), 2465–2474. https://doi.org/10.1038/sj.emboj.7601144

27. Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Molecular and cellular biology*, *26*(7), 2560–2569. https://doi.org/10.1128/MCB.26.7.2560-2569.2006

28. Simon, J. A., and Kingston, R. E. (2013). Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Molecular cell*, *49*(5), 808–824. https://doi.org/10.1016/j.molcel.2013.02.013

29. Grau, D. J., Chapman, B. A., Garlick, J. D., Borowsky, M., Francis, N. J., and Kingston, R. E. (2011). Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes & development*, *25*(20), 2210–2221. https://doi.org/10.1101/gad.17288211

30. Plys, A. J., Davis, C. P., Kim, J., Rizki, G., Keenen, M. M., Marr, S. K., and Kingston, R. E. (2019). Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes & development*, *33*(13-14), 799–813. https://doi.org/10.1101/gad.326488.119

31. Jaensch, E. S., Zhu, J., Cochrane, J. C., Marr, S. K., Oei, T. A., Damle, M., McCaslin, E. Z., and Kingston, R. E. (2021). A Polycomb domain found in committed cells impairs differentiation when introduced into PRC1 in pluripotent cells. *Molecular cell*, *81*(22), 4677–4691.e8. https://doi.org/10.1016/j.molcel.2021.09.018

32. Santanach, A., Blanco, E., Jiang, H., Molloy, K. R., Sansó, M., LaCava, J., Morey, L., and Di Croce, L. (2017). The Polycomb group protein CBX6 is an essential regulator of embryonic

stem cell identity. *Nature communications*, *8*(1), 1235. https://doi.org/10.1038/s41467-017-01464-w

33. 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*, *1804*(4), 996–1010. https://doi.org/10.1016/j.bbapap.2010.01.011

34. Peng, K., Radivojac, P., Vucetic, S., Dunker, A. K., and Obradovic, Z. (2006). Length-dependent prediction of protein intrinsic disorder. *BMC bioinformatics*, *7*, 208. https://doi.org/10.1186/1471-2105-7-208

35. Piovesan, D., Del Conte, A., Clementel, D., Monzon, A. M., Bevilacqua, M., Aspromonte, M. C., Iserte, J. A., Orti, F. E., Marino-Buslje, C., and Tosatto, S. C. E. (2023). MobiDB: 10 years of intrinsically disordered proteins. *Nucleic acids research*, *51*(D1), D438–D444. https://doi.org/10.1093/nar/gkac1065

36. Pirrotta, V., and Li, H. B. (2012). A view of nuclear Polycomb bodies. *Current opinion in genetics & development*, *22*(2), 101–109. https://doi.org/10.1016/j.gde.2011.11.004

37. Saurin, A. J., Shiels, C., Williamson, J., Satijn, D. P., Otte, A. P., Sheer, D., and Freemont, P. S. (1998). The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. *The Journal of cell biology*, *142*(4), 887–898. https://doi.org/10.1083/jcb.142.4.887

38. Sabari, B. R., Dall'Agnese, A., and Young, R. A. (2020). Biomolecular Condensates in the Nucleus. *Trends in biochemical sciences*, *45*(11), 961–977. https://doi.org/10.1016/j.tibs.2020.06.007

39. Lafontaine, D. L. J., Riback, J. A., Bascetin, R., and Brangwynne, C. P. (2021). The nucleolus as a multiphase liquid condensate. *Nature reviews. Molecular cell biology*, *22*(3), 165–182. https://doi.org/10.1038/s41580-020-0272-6

40. Liao, S. E., and Regev, O. (2021). Splicing at the phase-separated nuclear speckle interface: a model. *Nucleic acids research*, *49*(2), 636–645. https://doi.org/10.1093/nar/gkaa1209

41. Shin, Y., and Brangwynne, C. P. (2017). Liquid phase condensation in cell physiology and disease. *Science*, *357*(6357), eaaf4382. https://doi.org/10.1126/science.aaf4382

42. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K., and Sharp, P. A. (2017). A Phase Separation Model for Transcriptional Control. *Cell*, *169*(1), 13–23. https://doi.org/10.1016/j.cell.2017.02.007

43. Strom, A. R., Emelyanov, A. V., Mir, M., Fyodorov, D. V., Darzacq, X., and Karpen, G. H. (2017). Phase separation drives heterochromatin domain formation. *Nature*, *547*(7662), 241–245. https://doi.org/10.1038/nature22989

44. Cho, W. K., Spille, J. H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I. I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science, 361*(6400), 412–415. https://doi.org/10.1126/science.aar4199

45. 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., Li, C. H., Guo, Y. E., Day, D. S., Schuijers, J., Vasile, E., Malik, S., Hnisz, D., Lee, T. I., Cisse, I. I., Roeder, R. G., Sharp, P. A., Chakraborty, A.K., and Young, R. A. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science*, *361*(6400), eaar3958. https://doi.org/10.1126/science.aar3958

46. Lyons, H., Veettil, R. T., Pradhan, P., Fornero, C., De La Cruz, N., Ito, K., Eppert, M., Roeder, R. G., and Sabari, B. R. (2023). Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell*, *186*(2), 327–345.e28. https://doi.org/10.1016/j.cell.2022.12.013

47. Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S., and Narlikar, G. J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature*, *547*(7662), 236–240. https://doi.org/10.1038/nature22822

48. Keenen, M. M., Brown, D., Brennan, L. D., Renger, R., Khoo, H., Carlson, C. R., Huang, B., Grill, S. W., Narlikar, G. J., and Redding, S. (2021). HP1 proteins compact DNA into mechanically and positionally stable phase separated domains. *eLife*, *10*, e64563. https://doi.org/10.7554/eLife.64563

49. Sanulli, S., Trnka, M. J., Dharmarajan, V., Tibble, R. W., Pascal, B. D., Burlingame, A. L., Griffin, P. R., Gross, J. D., and Narlikar, G. J. (2019). HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature*, *575*, 390–394. https://doi.org/10.1038/s41586-019-1669-2

50. Tortora, M. M. C., Brennan, L. D., Karpen, G., and Jost, D. (2023). HP1-driven phase separation recapitulates the thermodynamics and kinetics of heterochromatin condensate formation. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(33), e2211855120. https://doi.org/10.1073/pnas.2211855120

51. Gibson, B. A., Doolittle, L. K., Schneider, M. W. G., Jensen, L. E., Gamarra, N., Henry, L., Gerlich, D. W., Redding, S., and Rosen, M. K. (2019). Organization of chromatin by intrinsic and regulated phase separation. *Cell*, *179*(2), 470-484. https://doi.org/10.1016/j.cell.2019.08.037

52. Shakya, A., Park, S., Rana, N., and King, J. T. (2019). Liquid-liquid phase separation of histone proteins in cells: Role in chromatin organization. *Biophysical Journal*, *118*(3), 753-764. https://doi.org/10.1016/j.bpj.2019.12.022

53. Zhang, M., Díaz-Celis, C., Onoa, B., Cañari-Chumpitaz, C., Requejo, K. I., Liu, J., Vien, M., Nogales, E., Ren, G., and Bustamante, C. (2022). Molecular organization of the early stages of nucleosome phase separation visualized by cryo-electron tomography. *Molecular Cell*, *82*(16), 3000–3014.e9. https://doi.org/10.1016/j.molcel.2022.06.032

54. Patil, A., Strom, A. R., Paulo, J. A., Collings, C. K., Ruff, K. M., Shinn, M. K., Sankar, A., Cervantes, K. S., Wauer, T., St Laurent, J. D., Xu, G., Becker, L. A., Gygi, S. P., Pappu, R. V., Brangwynne, C. P., and Kadoch, C. (2023). A disordered region controls cBAF activity via condensation and partner recruitment. *Cell*, *186*(22), 4936–4955.e26. https://doi.org/10.1016/j.cell.2023.08.032

55. Tatavosian, R., Kent, S., Brown, K., Yao, T., Duc, H. N., Hunyh, T. N., Zhen, C. Y., Ma, B., Wang, H., and Ren, X. (2019). Nuclear condensates of the Polycomb protein chromobox2 (CBX2) assemble through phase separation. *Journal of Biological Chemistry, 294*, 1451-1463. https://doi.org/10.1074/jbc.RA118.006620

56. Gallego, L. D., Schneider, M., Mittal, C., Romanauska, A., Gudino Carrillo, R. M., Schubert, T., Pugh, B. F., and Köhler, A. (2020). Phase separation directs ubiquitination of gene-body nucleosomes. *Nature*, *579*(7800), 592–597. https://doi.org/10.1038/s41586-020-2097-z

57. Brown, K., Andrianakos, H., Ingersoll, S., and Ren, X. (2021). Single-molecule imaging of epigenetic complexes in living cells: insights from studies on Polycomb group proteins. *Nucleic acids research*, *49*(12), 6621–6637. https://doi.org/10.1093/nar/gkab304

58. Uckelmann, M., Levina, V., Taveneau, C., Han Ng, X., Pandey, V., Martinez, J., Mendiratta, S., Houx, J., Boudes, M., Venugopal, H., Trépout, S., Zhang, Q., Flanigan, S., Li, M., Sierecki, E., Gambin, Y., Das, P. P., Bell, O., de Marco, A., and Davidovich, C. (2024). Dynamic PRC1-CBX8 stabilizes a porous structure of chromatin condensates. *bioRxiv*, 2023.05.08.539931. https://doi.org/10.1101/2023.05.08.539931

59. Murphy, S. E., and Boettiger, A. N. (2024). Polycomb repression of Hox genes involves spatial feedback but not domain compaction or phase transition. *Nature genetics*, *56*(3), 493–504. https://doi.org/10.1038/s41588-024-01661-6

60. Eeftens, J. M., Kapoor, M., Michieletto, D., and Brangwynne, C. P. (2021). Polycomb condensates can promote epigenetic marks but are not required for sustained chromatin compaction. *Nature communications*, *12*(1), 5888. https://doi.org/10.1038/s41467-021-26147-5

61. Seif, E., Kang, J. J., Sasseville, C., Senkovich, O., Kaltashov, A., Boulier, E. L., Kapur, I., Kim, C. A., and Francis, N. J. (2020). Phase separation by the polyhomeotic sterile alpha motif compartmentalizes Polycomb Group proteins and enhances their activity. *Nature communications*, *11*(1), 5609. https://doi.org/10.1038/s41467-020-19435-z

62. Niekamp, S., Marr, S. K., Oei, T. A., Subramanian, R., and Kingston, R. E. (2024). Modularity of PRC1 composition and chromatin interaction define condensate

properties. *Molecular cell*, S1097-2765(24)00181-3. Advance online publication. https://doi.org/10.1016/j.molcel.2024.03.001

63. Lau, M. S., Schwartz, M. G., Kundu, S., Savol, A. J., Wang, P. I., Marr, S. K., Grau, D. J., Schorderet, P., Sadreyev, R. I., Tabin, C. J., and Kingston, R. E. (2017). Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. *Science, 355*(6329), 1081–1084. https://doi.org/10.1126/science.aah5403

64. Parreno, V., Martinez, A. M., and Cavalli, G. (2022). Mechanisms of Polycomb group protein function in cancer. *Cell research*, *32*(3), 231–253. https://doi.org/10.1038/s41422-021-00606-6

65. Gan, L., Yang, Y., Li, Q., Feng, Y., Liu, T., and Guo, W. (2018). Epigenetic regulation of cancer progression by EZH2: from biological insights to therapeutic potential. *Biomarker research*, *6*, 10. https://doi.org/10.1186/s40364-018-0122-2

66. Straining, R., and Eighmy, W. (2022). Tazemetostat: EZH2 Inhibitor. *Journal of the advanced practitioner in oncology*, *13*(2), 158–163. https://doi.org/10.6004/jadpro.2022.13.2.7

67. Puram, S., Oppelt, P. J., Ley, J. C., Cohen, J., Alberti, J., Shannon, E., Liu, J., and Adkins, D. (2023). Tasemetostat, a selective EZH2 inhibitor, with pembrolizumab as treatment of anti-PD-1 resistant head and neck squamous-cell carcinoma (HNSCC): A phase 1-2 trial. *Journal of clinical oncology*, *41*(16), 6022.

68. Xu, J., Li, L., Shi, P., Cui, H., and Yang, L. (2022). The Crucial Roles of Bmi-1 in Cancer: Implications in Pathogenesis, Metastasis, Drug Resistance, and Targeted Therapies. *International journal of molecular sciences*, *23*(15), 8231. https://doi.org/10.3390/ijms23158231

69. Zhang, X. W., Sheng, Y. P., Li, Q., Qin, W., Lu, Y. W., Cheng, Y. F., Liu, B. Y., Zhang, F. C., Li, J., Dimri, G. P., and Guo, W. J. (2010). BMI1 and Mel-18 oppositely regulate carcinogenesis and progression of gastric cancer. *Molecular cancer*, *9*, 40. https://doi.org/10.1186/1476-4598-9-40

70. Chan, H. L., Beckedorff, F., Zhang, Y., Garcia-Huidobro, J., Jiang, H., Colaprico, A., Bilbao, D., Figueroa, M. E., LaCava, J., Shiekhattar, R., and Morey, L. (2018). Polycomb complexes associate with enhancers and promote oncogenic transcriptional programs in cancer through multiple mechanisms. *Nature communications*, *9*(1), 3377. https://doi.org/10.1038/s41467-018-05728-x

71. Wang, L., Ren, B., Zhuang, H., Zhong, Y., and Nan, Y. (2021). CBX2 Induces Glioma Cell Proliferation and Invasion Through the Akt/PI3K Pathway. *Technology in cancer research & treatment*, *20*, 15330338211045831. https://doi.org/10.1177/15330338211045831

72. Gu, X., Wang, X., Su, D., Su, X., Lin, L., Li, S., Wu, Q., Liu, S., Zhang, P., Zhu, X., and Jiang, X. (2018). CBX2 Inhibits Neurite Development by Regulating Neuron-Specific Genes Expression. *Frontiers in molecular neuroscience*, *11*, 46. https://doi.org/10.3389/fnmol.2018.00046

73. Tsherniak, A., Vazquez, F., Montgomery, P. G., Weir, B.A., Kryukov, G., Cowley, G. S., Gill, S., Harrington, W. F., Pantel, S., Krill-Burger, J. M., Meyers, R. M., Ali, L., Goodale, A., Lee, Y., Jiang, G., Hsiao, J., Gerath, W. F. J., Howell, S., Merkel, E., Ghandi, M., Garraway, L. A., Root, D. E., Golub, T. R., Boehm, J. S., and Hahn, W. C. (2017). Defining a cancer dependency map. *Cell, 170*(3), 564-576. https://doi.org/10.1016/j.cell.2017.06.010

74. Zheng, S., Lv, P., Su, J., Miao, K., Xu, H., and Li, M. (2019). Overexpression of CBX2 in breast cancer promotes tumor progression through the PI3K/AKT signaling pathway. *American journal of translational research*, *11*(3), 1668–1682.

75. Mao, J., Tian, Y., Wang, C., Jiang, K., Li, R., Yao, Y., Zhang, R., Sun, D., Liang, R., Gao, Z., Wang, Q., and Wang, L. (2019). CBX2 Regulates Proliferation and Apoptosis via the Phosphorylation of YAP in Hepatocellular Carcinoma. *Journal of Cancer*, *10*(12), 2706–2719. https://doi.org/10.7150/jca.31845

76. Zhao, W., Ma, B., Tian, Z., Han, H., Tang, J., Dong, B., An, G., Cao, B., and Wang, B. (2021). Inhibiting CBX4 efficiently protects hepatocellular carcinoma cells against sorafenib resistance. *British journal of cancer*, *124*(7), 1237–1248. https://doi.org/10.1038/s41416-020-01240-6

77. Li, J., Xu, Y., Long, X. D., Wang, W., Jiao, H. K., Mei, Z., Yin, Q. Q., Ma, L. N., Zhou, A. W., Wang, L. S., Yao, M., Xia, Q., and Chen, G. Q. (2014). Cbx4 governs HIF-1α to potentiate angiogenesis of hepatocellular carcinoma by its SUMO E3 ligase activity. *Cancer cell*, *25*(1), 118–131. https://doi.org/10.1016/j.ccr.2013.12.008

78. Shi, D., Ao, L., Yu, H., Xia, Y., Li, J., Zhong, W., and Xia, H. (2021). Chromobox Homolog 8 (CBX8) in Human Tumor Carcinogenesis and Prognosis: A Pancancer Analysis Using Multiple Databases. *Frontiers in genetics*, *12*, 745277. https://doi.org/10.3389/fgene.2021.745277

79. Chung, C. Y., Sun, Z., Mullokandov, G., Bosch, A., Qadeer, Z. A., Cihan, E., Rapp, Z., Parsons, R., Aguirre-Ghiso, J. A., Farias, E. F., Brown, B. D., Gaspar-Maia, A., and Bernstein, E. (2016). Cbx8 Acts Non-canonically with Wdr5 to Promote Mammary Tumorigenesis. *Cell reports*, *16*(2), 472–486. https://doi.org/10.1016/j.celrep.2016.06.002

80. Tan, J., Jones, M., Koseki, H., Nakayama, M., Muntean, A. G., Maillard, I., and Hess, J. L. (2011). CBX8, a polycomb group protein, is essential for MLL-AF9-induced leukemogenesis. *Cancer cell*, *20*(5), 563–575. https://doi.org/10.1016/j.ccr.2011.09.008

81. Tang, B., Tian, Y., Liao, Y., Li, Z., Yu, S., Su, H., Zhong, F., Yuan, G., Wang, Y., Yu, H., Tomlinson, S., Qiu, X., and He, S. (2019). CBX8 exhibits oncogenic properties and serves as

a prognostic factor in hepatocellular carcinoma. *Cell death & disease*, *10*(2), 52. https://doi.org/10.1038/s41419-018-1288-0

82. Li, J., Ouyang, T., Li, M., Hong, T., Alriashy, M., Meng, W., and Zhang, N. (2021). CBX7 is Dualistic in Cancer Progression Based on its Function and Molecular Interactions. *Frontiers in genetics*, *12*, 740794. https://doi.org/10.3389/fgene.2021.740794

83. Wang, J., Yang, B., Zhang, X., Liu, S., Pan, X., Ma, C., Ma, S., Yu, D., and Wu, W. (2023). Chromobox proteins in cancer: Multifaceted functions and strategies for modulation. *International journal of oncology*, *62*(3), 36. https://doi.org/10.3892/ijo.2023.5484

84. Stuckey, J. I., Dickson B. M., Cheng, N., Liu, Y., Norris, J. L., Cholensky, S. H., Tempel, W., Qin, S., Huber, K. G., Sagum, C., Black, K., Li, F., Huang, X. P., Roth, B. L., Baughman, B. M., Senisterra, G., Pattenden, S. G., Vedadi, M., Brown, P. J., Bedford, M. T., Min, J., Arrowsmith, C. H., James, L. I., and Frye, S. V. (2016). A cellular chemical probe targeting the chromodomains of Polycomb repressive complex 1. *Nature Chemical Biology*, *12*(3):180-187. https://doi.org/10.1038/nchembio.2007.

Chapter 2 : Characterizing the phase separation behavior of CBX4 and CBX8

Theresa A. Oei[1,2,3] and Robert E. Kingston[2,3]

Affiliations:

[1]Department of Chemical Biology, Harvard University; Cambridge, MA, 02138, USA

[2]Department of Molecular Biology, Massachusetts General Hospital; Boston, MA, 02114, USA

[3]Department of Genetics, Harvard Medical School; Boston, MA, 02115, USA

2.1 Introduction to CBX4 and CBX8

The CBX family of proteins plays a central role in targeting canonical PRC1 to genes to facilitate heterochromatin formation and maintain gene silencing by the PcG proteins.[1-4] They have a chromodomain that binds to the H3K27me3 mark deposited by PRC2.[5] There are likely other means of targeting PRC1 to promoters in mammalian cells, but the exact mechanism is unknown.[6,7] The chromodomains of each CBX protein have different affinities for H3K27me3 and H3K9me3.[8] There has been some contradictory studies of *in vitro* nucleosome binding by the chromodomains of the CBX proteins. One study demonstrated that CBX2 and CBX7 bind preferentially to H3K27me3, whereas CBX4 binds preferentially to H3K9me3.[8] CBX6 and CBX8 did not bind to either mark under the conditions of this assay.[8] In contrast, a different group demonstrated that CBX8 association with chromatin was mainly drive by its chromodomain with preferential binding for H3K27me3 over H3K9me3 alongside DNA binding.[9] Indeed, CBX8 peaks in ChIP-seq overlap with H3K27me3 peaks by more than 70% in K562 and fibroblast cell lines.[9] They conclude that binding to H3K27me3 drives CBX8's association with chromatin *in vivo*.[9] Besides the chromodomain, CBX2 has an AT-Hook region that binds to DNA and the CBX4 and CBX8 also have AT-Hook like regions, which may play a role in DNA targeting.[2] The CBX proteins differ greatly in size and amino acid sequence outside of the conserved structured domains such as the chromodomain or the C box, and this has been shown to facilitate distinct chromatin association patterns in the nucleus.[7] CBX2, CBX4, and CBX8 all have highly disordered regions at the center of the protein that are strongly charged.[10,11] Alphafold2 predictions show the central disordered regions as an unstructured cloud of spaghetti surrounding the structured domains (**Fig. 2.1)**.[12] Proteins that interact with chromatin are enriched with IDRs in comparison to the rest of the proteome, suggesting an evolutionary

conservation of disorder to facilitate chromatin modulation.[13] Their proposed function includes regulating intranuclear localization and facilitating higher order complex formation via multivalent interactions.[13] Nucleic acid targeting may also be specified by the interaction of IDRs with more structured binding domains.[13,14]

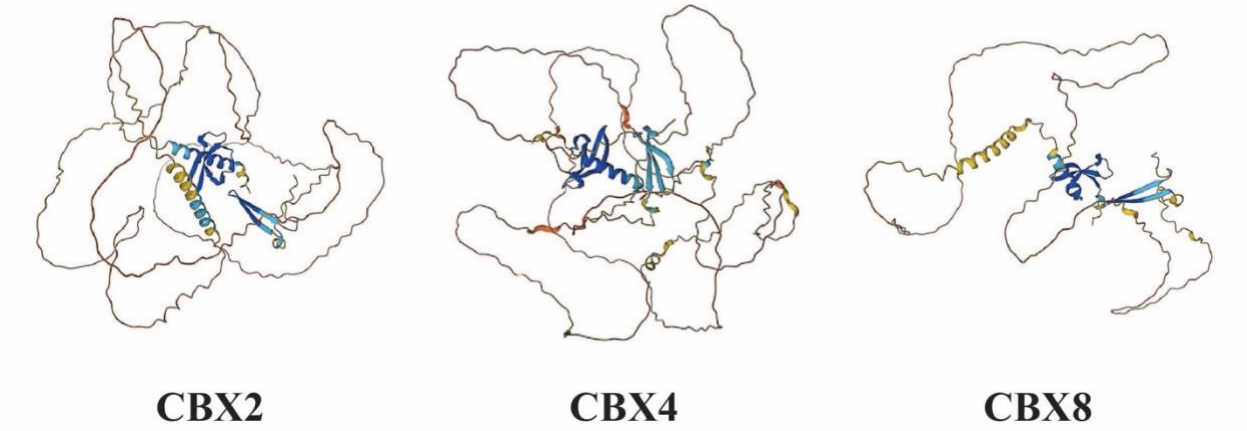


Figure 2.1 Predicted protein structure for CBX2, CBX4, and CBX8

CBX2, CBX4, and CBX8 have long disordered regions. Alphafold 2 structure predictions for *mus muscu*lus CBX2, CBX4, and CBX8.[12]

The behavior of IDRs may be modulated by post-transcriptional modifications.[15] CBX2 has an enriched serine patch that is known to be phosphorylated.[16] Phosphorylation of the CBX2 serine rich region contributes to its ability to phase separate.[11] CBX4 has a long histidine stretch. Histidine rich regions have been shown to play a role in the phase separation of other proteins and are often phosphorylated.[17-19] CBX8 is known to be phosphorylated by the PIM1 kinase.[20] This phosphorylation promotes the degradation of CBX8 controlling the concentration of CBX8 in the nucleus to induce senescence in fibroblasts.[20] Modification of the CBX4 and CBX8 IDRs may play a role in tuning the proposed phase separation activity of these proteins.

The CBX family of proteins are differentially expressed in distinct cell types.[21,22] Differences in expression levels may help to modulate the silencing of target genes to maintain chromatin states specific to a particular tissue.[21,22] For example, CBX4 plays an important role in thymus development, whereas CBX8 is thought to regulate erythroid differentiation.[23,24] There is evidence that they often target the same genes, however some studies have shown that they have distinct target genes subsets.[21,22,25] Additionally, the differential roles of CBX proteins in different cancer types suggests that they have unique roles in maintaining distinct cell populations.[26]

Another perspective on the role of CBX proteins includes their potential ability to compensate for one another through some overlapping gene targets.[2] RING1B knockout leads to defects in gastrulation.[27] However, knockout of the individual CBX proteins has less severe effects, and knockout of CBX2 in spermatogonia upregulates CBX4 and CBX8 expression.[2,28] Although it is challenging to compare since RING1B is also part of the ncPRC1 complex, this could been

interpreted as potential evidence for the redundancy of the CBX proteins. CBX2, CBX4, and CBX8 are all found within a 65 kb region of the genome.[29] Similarly, CBX6 and CBX7 are located together in the genome.[29] This suggests that the diversity of the CBX family may have come about through evolutionary gene duplication events during mammalian evolution.[30] There is a single homolog in *Drosophila melanogaster* known as the *Pc* gene, whereas the CBX paralogs diversified in vertebrates.[30] Recent work suggests that ncPRC1 preceded cPRC1 in evolution and the CBX family might have evolved as an ortholog of RYBP from ncPRC1 at the start of metazoan evolution.[31] The expansion of the CBX locus may be related to the increasing number of cell types requiring differential silent chromatin states in mammalian development.[30]

Previous work from the lab characterized the ability of CBX2 to form condensates *in vitro* and in cells.[11] Positive charges from its IDR contribute to its ability to phase separate *in vitro* and form condensates in cells.[11] These residues are also responsible for the compaction of nucleosomal arrays *in vitro*.[10] This suggests that CBX2's phase separation and compaction activity are inextricably linked in its IDR. Thus, this region was named the Compaction and Phase Separation (CaPS) region of CBX2.[11] Mutations of the same lysine and arginine residues of CBX2 in a mouse led to the development of axial transformations effected by *Hox* gene misexpression.[32] Thus, there was strong evidence that the CaPS region was important for cPRC1 function. CaPS-like disordered regions can be found in CBX4 and CBX8 as well. Previous *in vitro* assays demonstrated compaction of nucleosomal arrays by CBX8 as well.[10] Therefore, the goal of this work was to determine if CBX4 and CBX8 also contribute to phase separation of cPRC1 and Polycomb body formation. Additionally, it was hypothesized that the positive charges in the IDRs of CBX4 and CBX8 would be important for phase separation activity.

2.2 CBX4 and CBX8 form droplets in vitro and condensates in cells

CBX4 and CBX8 were individually co-expressed with GFP tagged Ring1b as a heterodimer and purified from insect cells. The purified proteins form round fluorescent droplets *in vitro* even at low concentrations in physiological salt conditions as visualized by fluorescence microscopy (**Fig. 2.2**). These spherical condensates increase in size with an increase in protein concentration. Additionally, in a spin down assay that involves centrifugation of the purified heterodimer, CBX4-RING1B and CBX8-RING1B both form condensates at the bottom of an Eppendorf tube (**Fig. 2.2**). Both experiments suggest that CBX4 and CBX8 self-interact to facilitate liquid liquid phase separation.

To test if positive charges in the IDRs of CBX4 and CBX8 are responsible for phase separation, lysine and arginine residues in the disordered regions were mutated to alanine. The CBX4 and CBX8 mutants do not form droplets or condense in the spin-down assay (**Fig. 2.2**). This data suggests that the positive charges in the CBX4 and CBX8 IDRs are important for *in vitro* phase separation.

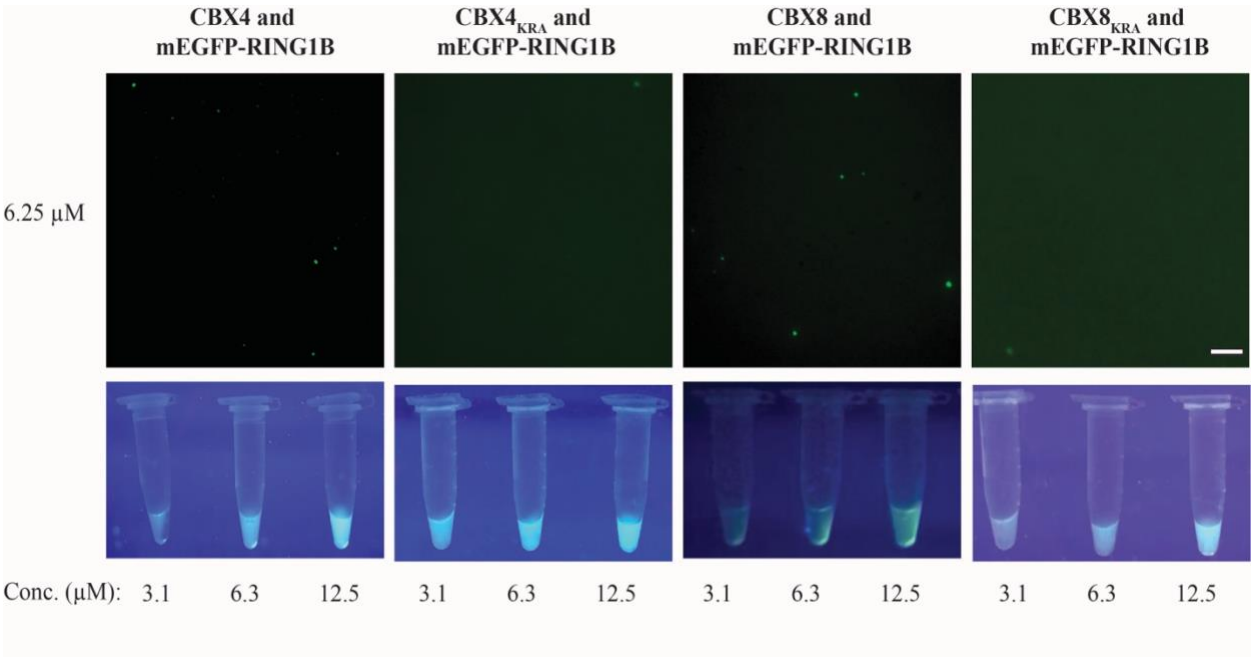


Figure 2.2 CBX4 and CBX8 in vitro assays

Lysine and arginine residues in the IDRs of CBX4 and CBX8 drive condensation in vitro. Purified CBX4 and CBX8 heterodimerized with mEGFP-RING1B forms droplets *in vitro* and condenses in a spin-down assay. Lysine and arginine mutations in the CBX4 or CBX8 IDRs prevent phase separation. Scale bar is 10 μm.

2.3 CBX4 and CBX8 drive condensate formation in cells

I examined the behavior of CBX4 and CBX8 in cells by overexpressing mEGFP-CBX4 and mEGFP-CBX8 in NIH-3T3 fibroblasts through lentiviral integration and doxycycline induction. Both cell lines had bright, fluorescent puncta in the nucleus in live cell imaging on a confocal microscope (**Fig. 2.3**). In contrast, the IDR mutants, mEGFP-CBX4$_{KRA}$ or mEGFP-CBX8$_{KRA}$, do not form puncta in the nuclei of NIH-3T3 cells (**Fig. 2.3**). Therefore, we conclude that the positive charges in the IDRs of CBX4 and CBX8 contribute to condensate formation in cells.

Although the IDRs of CBX2, CBX4, and CBX8 seem to have similar functions in driving *in vitro* phase separation and condensate formation in cells, they are different in amino acid composition and length. Perhaps these differences alter the dynamics of the condensates formed in cells. In order to study these dynamics, Fluorescence Recovery After Photobleaching (FRAP) was performed on mEGFP-CBX4 and mEGFP-CBX8 puncta in NIH-3T3 cells (**Fig. 2.3**). Both show full recovery within sixty seconds similarly to CBX2.[11] This recovery time is also similar to FRAP recovery for the nucleolus, a well-established liquid liquid phase separated nuclear body.[33] The unrecovered fraction, which is ~30% for CBX4 and ~10% for CBX8, is likely an immobile chromatin bound fraction of respective CBX proteins. Previous data on CBX2 shows ~30% immobile fraction.[11] This may indicate that less CBX8 is bound in this cell line, which is in line with *in vitro* data suggesting that CBX4 and CBX2 both have stronger affinities for H3K27me3 than CBX8.[8] CBX4 puncta are also more stable than CBX8 puncta with CBX8 being more highly dynamic. The half-time for CBX4 recovery is 27 seconds whereas the half-time for CBX8 recovery is 15 seconds. This is consistent with previous findings in U2OS cells.[34] Previous FRAP data for CBX2 shows a similar half-life.[11] Concerns have been raised over the

use of FRAP to quantify phase separation activity.[35] With this in mind, FRAP is most likely a qualitative perspective for CBX dynamics. Regardless, this data demonstrates that CBX4 and CBX8 diffuse dynamically into nuclear Polycomb bodies in a manner that is consistent with LLPS.

Additionally, the dynamics of CBX2, CBX4 and CBX8 phase separation were explored with the use of OptoDroplet technology.[36] Fusion of Cry2, a light inducible homo-oligomerizing protein from *Arabidopsis thaliana*, and mCherry to CBX2, CBX4, and CBX8 allows for light induction of Polycomb bodies in NIH-3T3 cells.[36] The N-terminus of FUS, a classic phase separating IDR, was used as a control. Droplets were induced with blue light activation for one minute and then the rate of droplet dissolution was measured with the hypothesis that weaker multivalent interactions would lead to higher dissolution rates. Consistent with this theory, CBX8 droplets seem to dissolve the fastest when the blue light is turned off (**Fig. 2.3**). However, the differences between the dissolution rates for CBX2, CBX4, and CBX8 were not statistically significant. Encouragingly, droplets formed by CBX2, CBX4, or CBX8 dissolved at similar rates to droplets formed by FUS. This assay likely does not have the resolution to truly differentiate the phase separation dynamics of distinct CBX proteins, however it does give an interesting perspective on the potential relative strengths of homotypic and heterotypic interactions within Polycomb bodies. Although these dynamics may be dependent on protein expression in the nucleus, and the particular cellular context of distinct cell lineages. Overall, examination of live cell dynamics for CBX2, CBX4, and CBX8 demonstrate that they are relatively similar in their abilities to phase separate, perhaps unsurprising given that they often colocalize at target genes in the same Polycomb bodies.

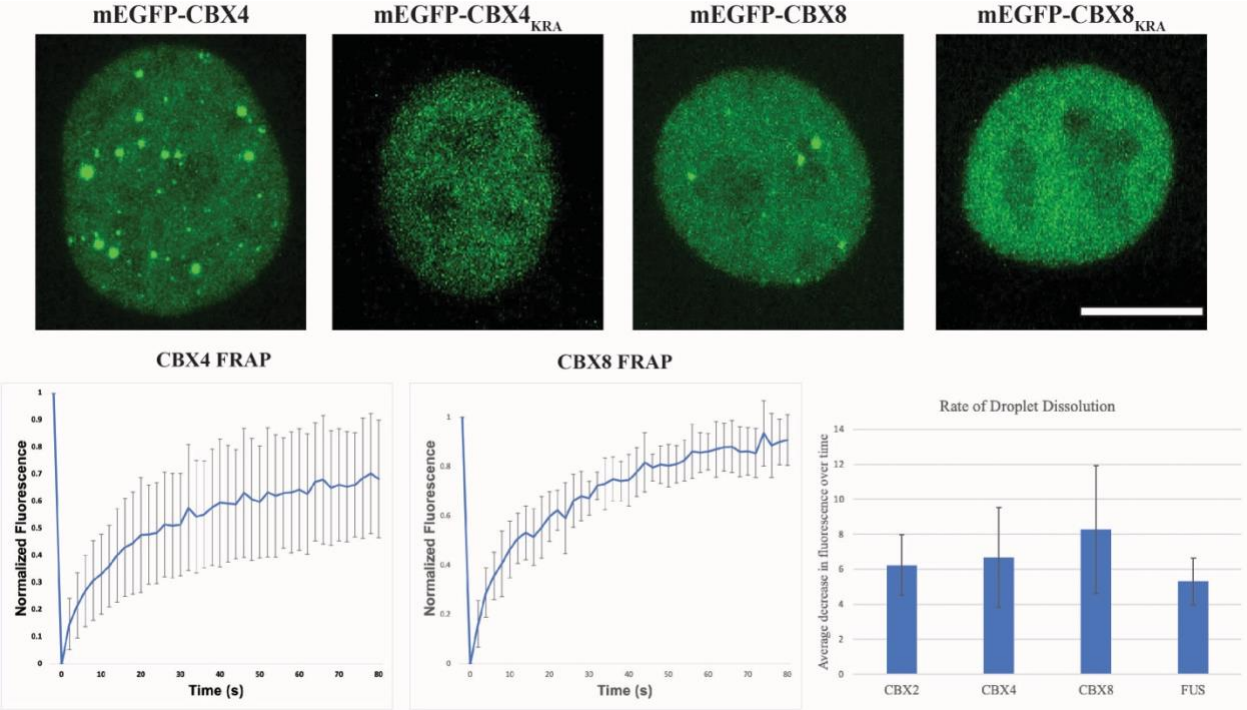


Figure 2.3 CBX4 and CBX8 form condensates in cells

Lysine and arginine residues in the IDRs of CBX4 and CBX8 drive dynamic condensates in live cells. mEGFP-tagged CBX4, CBX8, $CBX4_{KRA}$, and $CBX8_{KRA}$, were expressed in NIH-3T3 cells by lentivirus integration and imaged by live cell confocal microscopy. Scale bar is 10 μm. Fluorescence Recovery After Photobleaching (FRAP) was performed on mEGFP-CBX4 and mEGFP-CBX8 puncta to measure dynamics. OptoDroplet constructs were used to compare the rates of droplet dissolution for CBX2, CBX4, CBX8 and the FUS^{IDR}.[36]

2.4 Conclusions and future directions

CBX4 and CBX8 phase separate *in vitro* and form condensates in cells. This confirms that CBX4 and CBX8 are more similar to CBX2 and distinct from CBX7, which does not phase separate *in vitro*.[11,37] Further characterization of CBX4 and CBX8 phase separation activity will include *in vitro* phase diagrams and assays to measure nucleosomal array incorporation for purified CBX4 and CBX8 containing cPRC1. Phase diagrams for CBX4 and CBX8 would indicate the threshold concentration required for *in vitro* phase separation, which should be in the range of intracellular concentrations of the protein to confirm LLPS in cells.[38] CBX2 has been shown to coat nucleosomal arrays before initiating condensate formation in TIRF microscopy experiments.[39] It would be interesting to compare CBX4 and CBX8 to CBX2 in these types of assays to further understand the differences between the paralogs and the impact of their distinct IDR composition on nucleosome compaction and thresholds of condensate formation.

CBX6 and CBX7 are typically expressed in the embryonic stem cell stage.[40,41] CBX2, CBX4 and CBX8 turn on as cells differentiate.[21] Therefore, it would be helpful to understand the cooperativity or competition between CBX2, CBX4, and CBX8 driven Polycomb bodies. CBX2 has previously been shown to concentrate CBX4 in condensates, but to disperse CBX8 condensates.[42] The relationship between these paralogs may be distinct in different cell types. Measuring all three endogenous paralogs in the same cell by single molecule tracking would help to illuminate the relative dynamics of these proteins in condensates. Previous single molecule tracking studies suggest low target occupancy of cPRC1.[43] This might seem counterintuitive to stable repression, but it likely highlights the importance of the liquid-like dynamics contributed by each of the CBX proteins to Polycomb body function in repression. Co-IP of RING1B

coupled with mass spec methods could be used to determine the stoichiometry of CBX paralogs in cPRC1 complexes throughout multiple cell lines. The ratio of CBX2, CBX4, and CBX8 found at different genomic loci and the corresponding chromatin state or transcriptional profile of the gene would give insight into the relationship between Polycomb body composition and gene repression.

The relationship between Polycomb body dynamics and cPRC1 DNA binding is also unclear. DNA Fluorescence In Situ Hybridization (FISH) combined with Immunofluorescence could be used to visualize the colocalization of CBX4 and CBX8 containing Polycomb bodies with canonical target genes such as the *Hox* clusters. It is known that often the CBX proteins co-occupy target genes, however they can also have different targets depending on the cell type.[21,22,25] Profiling of CBX target genes in the NIH-3T3 cell lines by CUT&RUN or ChIP-seq with CBX4 or CBX8 would give insight into targeting of CBX4 and CBX8, which are endogenously expressed in this cell line. CBX2 is also expressed in this cell line at very low levels. Comparison of binding profiles for each CBX and the effect of overexpression of each could distinguish the effect of increasing protein concentration in the nucleus on genomic targeting.

Mutation of all lysine and arginine residues in the CBX4 and CBX8 IDRs strongly impacted the ability of CBX4 and CBX8 to phase separate. However, these mutants represent drastic changes to the amino acid composition of these proteins. Narrowing down the minimum number of positive charges that drive phase separation would be a necessary next step. For example, the positive charges in the first half of CBX2's IDR seem to be most important for driving its phase

separation activity.[11,37] CBX4 has a pattern of alternating charged and uncharged stretches in its disordered region. This type of distribution pattern has been shown to be important in phase separation of other transcriptional regulators such as MED1.[44] The blocks of charge are thought to give a sticker and spacer arrangement that facilitates multivalent interactions.[44] Shuffling of these blocks of charge within the CaPS-like region of CBX4 could be studied. Additionally, CBX4 has a stretch of histidine repeats, which have been shown to regulate transcriptional control via phase separation.[45] The histidine-rich region of the Positive transcription elongation factor b (P-TEFb) allows for phase separation partitioning with RNA Polymerase (Pol) II and hyperphosphorylation of Pol II's C terminal domain.[45] Further work to examine the impact of this histidine stretch, by mutation or deletion, on CBX4 phase separation *in vitro* and in cells would further our understanding of key amino acid drivers of CBX family condensation activity.

The regulation of Polycomb condensate formation is another underexplored topic. Post-translational modifications are known to regulate other nuclear bodies.[15] Previous work has shown that serine phosphorylation of CBX2 is important for its ability to phase separate in cells.[11] CBX8 is also phosphorylated in its disordered region.[20] Mutations could be made to the phosphorylated amino acids to understand the significance of phosphorylation in CBX8 phase separation. The PIM1 kinase has been shown to phosphorylate CBX8.[20] A PIM1 inhibitor could also be used to visualize the effects of phosphorylation on CBX8 phase separation, although this approach may have more non-specific effects given PIM1's other target proteins.

In contrast to the other CBX paralogs, CBX4 is an E3 SUMO ligase that also contains two SUMO Interaction Motifs (SIMs) that affect its sumoylation activity. Mutations of these SIMs

has been shown to alter recruitment of CBX4 to Polycomb bodies and decrease the downstream sumoylation of proteins such as CtBP.[46] Studies on other condensate forming proteins have shown that sumoylation can play a role in phase separation.[47] Purification of a SUMO-modified CBX4 or SIMs deletion mutant will be useful for understanding the effect of this PTM on *in vitro* phase separation. Deletion of the SIMs in endogenous CBX4 by CRISPR could also measure the impact of the SIMs on CBX4 localization to the genome and the effect on the recruitment of other CBX proteins to Polycomb bodies. Overall, the impact of CBX4's sumoylation activity on its canonical role in PRC1 repression should be more thoroughly examined.

This work demonstrates that CBX4 and CBX8 can also drive the phase separation of cPRC1 in a similar manner to CBX2. Lysine and arginine residues in the disordered regions of these proteins are important for generating phase separation activity. There are minor differences in the dynamics of CBX2, CBX4, and CBX8 that deserve further elucidation. Already this points to a system where the evolution of distinct CBX paralogs may tune condensate formation to adjust Polycomb-mediated repression. This may come about both through the relative levels of CBX2, CBX4, and CBX8 protein expression in each cell type or differential relative stoichiometry of CBX-containing cPRC1 complexes within Polycomb bodies or bound at Polycomb target genes. Further research will correlate the comparative dynamics of CBX2, CBX4, and CBX8 phase separation and the ability to suppress target genes. A recent preprint on CBX8-PRC1 suggests a porous chromatin structure that allows for both static chromatin and dynamic diffusion of CBX8-cPRC1 through multivalent interactions with histone tails, DNA, and nucleosomes.[48] This may contradict more simplistic hypotheses that Polycomb condensates form as impenetrable globules.

However, although the CBX8-cPRC1 structures are porous to cPRC1 complexes there is still proposed to be size exclusion of large transcription factors at target sites.[48] This contributes to a view of Polycomb bodies as dynamic nuclear structures in line with single molecule studies and FRAP dynamics while maintaining static, closed chromatin that is inaccessible to transcription machinery.[43] Perhaps the ratio of different CBX proteins present at a particular genomic locus tunes accessibility to transcriptional activators or the maintenance of gene repression. Further experiments to test the correlation of accessibility with cPRC1 composition will be required. Overall, the CBX4 and CBX8 data herein presented supports the hypothesis that phase separation and compaction can be attributed to an intrinsically disordered region that is functionally conserved even while amino acid composition has evolved variation between paralogs for tunable function of cPRC1. More generally, this work supports a growing body of evidence that cells use nuclear phase separation to regulate transcriptional control within the chromatin environment.

As discussed in Chapter 1, CBX4 and CBX8 have distinct roles in cancer.[49] Changes in CBX protein expression by gene amplification or misregulation of expression are often culprits in disease pathology.[22,26,49] This suggests that phase separation, which is sensitive to protein concentration, may play a role in CBX4 and CBX8 function in disease. Further understanding the mechanisms of CBX4 and CBX8 phase separation and their role in driving Polycomb-mediated repression will give further insight into CBX4 or CBX8-dependent cancers. Small molecules have been designed to target CBX4 and CBX8. UNC3866 has been developed to inhibit the chromodomain of CBX4, although selectivity amongst CBX proteins is still a challenge.[50] Treatment of hepatocellular carcinoma cells with UNC3866 suppressed tumor

growth in xenograft mouse models.[50] In contrast, UNC7040 has been developed as a positive allosteric modulator of CBX8 that displaces CBX8 from H3K27me3 binding.[51] Treatment of colorectal cancer cells with this compound led to decreased proliferation, suggesting a CBX8 dependency in this type of cancer.[51] In the future, developing small molecules to modulate CBX4 and CBX8 condensate formation may be a potential strategy to target disease pathology in many different cancer types.

2.5 Materials and methods

Protein Purification

Fastbac plasmids containing cDNA encoding Flag-tagged CBX4 or CBX8 and mEGFP-tagged RING1B were used to make baculovirus utilizing the Bac-to-Bac system from Thermo Fisher. Sf9 cells were infected with baculovirus and incubated with gentle shaking for 72hr at 27°C to express proteins. Sf9 cells were harvested by centrifugation. Nuclear extracts were prepared as follows on ice. Harvested Sf9 cells were washed once with PBS and then spun down and resuspended in Hypotonic buffer (10 mM Hepes pH 7.9, 1.5 mM MgCl2, 10 mM KCl, 0.2 mM PMSF, 0.5 mM DTT) at 2x the pelleted cell volume. Spun down for 5 min at 3000 rpm, then resuspended in Hypotonic buffer at 3x the pelleted cell volume. The cells were incubated on ice for 10 minutes. Then, the cells were transferred to a dounce homogenizer and dounced for ~25 strokes. The cells were transferred to 50mL Falcon tubes and centrifuged for 15 min at 4000 rpm. The cytoplasmic supernatant was removed. The sample was resuspended in low salt buffer (20 mM HEPES pH 7.9, 25% glycerol, 1.5 mM $MgCl_2$, 0.02 M KCl, 0.2 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT) at half the pelleted nuclei volume. The same volume of high salt buffer (20 mM HEPES pH 7.9, 25% glycerol, 1.5 mM $MgCl_2$, 1.2 M KCl, 0.2 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT) was added drop-wise to the tube while vortexing at low speed. The tube was placed on a rotator at 4^{o}C for 30 minutes. Then the sample was spun at 14,500 rpm for 30 min, The nuclear extract was removed to new tubes and the pellet was discarded. The nuclear extract was incubated with anti-Flag M2 affinity resin (Sigma) for 2 hr and then washed with BC buffer (20 mM HEPES pH 7.9, 0.2 mM EDTA, 20% glycerol, 0.05% NP-40, 0.5 mM DTT, 0.1 mM PMSF, cOmplete EDTA-free protease inhibitor [Roche]) containing 300 mM KCl. The resin was washed with BC buffer containing increasing concentrations (300, 600, 1200, and 2000 mM) of

KCl and then washed with BC buffer in decreasing order of KCl concentration to 300 mM KCl. Proteins were eluted from the resin using BC buffer containing 300 mM KCl and 0.8 mg/mL of Flag peptide. Fractions were taken every 30 min for several hours. The purity of the protein was assessed by Coomassie staining. The fractions were combined and concentrated using Amicon Ultra-4 centrifugal filter units and the protein samples quantified by Bradford assay.

Spin Down Assay

Purified protein was diluted into Eppendorf tubes on ice with buffer containing a final concentration of 20 mM HEPES (pH 7.9), 100 mM KCl, and 1 mM $MgSO_4$.
The samples were incubated at room temperature for 5 minutes and then centrifuged at 10,000 g for 5 minutes. The tubes were then visualized with a UV light box.

Droplet Assay

Purified mEGFP-tagged RING1B heterodimerized with CBX4 or CBX8 was diluted proteins in buffer containing a final concentration of 20 mM HEPES (pH 7.9), 100 mM KCl, and 1 mM $MgSO_4$ on ice. The solution was pipetted onto glass slides with coverslips. Samples were imaged at room temperature with a Nikon 90i Eclipse epifluorescence microscope equipped with an Orca ER camera (Hamamatsu). A 100× oil objective and images were collected using Volocity software (Perkin Elmer).

Cell Culture

NIH-3T3 cells were obtained from the American Type Culture Collection (ATCC) and cultured in sterile filtered DMEM with 10% Fetal Bovine Serum (FBS) and penicillin/streptomycin (100U/mL).

HEK293T cells were a gift from the lab of Prof. Paul Lehner (CITIID, Cambridge, UK). They were cultured in sterile filtered IMDM with 10% FBS and penicillin/streptomycin (100 U/mL). All cell lines were grown at 37^{o} C with 5% CO_2.

Lentiviral Cell Line Generation

Second generation lentiviral plasmids, pCMVΔR8.91 and pMD.G were transfected into HEK293T cells alongside the lentiviral vector containing the transgene of interest using the TransIT-293 Reagent from Mirus bio. The viral supernatant was harvested and filtered after 48 hours. The viral supernatant was used fresh or stored at -80^{o}C. NIH-3T3 cells were seeded at 0.5*10^5 cells/mL in a 24 well plate. After 24 hours, the media was aspirated from the 24 well plate and 1mL of the desired viral supernatant was used to transduce the cells. The cells were incubated with virus for 4-6 hours at 37^{o} C. Then, the virus was removed and replaced with regular growth media. After 48 hours, the cells were selected in puromycin at 1ug/mL for several days. For the cell lines transduced with a GFP or mCherry containing transgene, the selected lines were sorted on a BD FACS Melody cell sorter for high GFP or high mCherry expression.

Live Cell Imaging

Cell lines were seeded on round glass bottom dishes from Mattek at $0.5*10^5$ cells/mL and incubated overnight in imaging media, DMEM without phenol red supplemented with 10% FBS.

Expression of GFP-CBX4 and GFP-CBX8 was induced in the engineered NIH-3T3 cell lines with doxycycline at 1ug/mL concentration. Twenty-four hours later, the samples were imaged on a confocal Nikon A1R with a temperature-controlled live cell chamber at 63x magnification. Z-stack images were generated with 0.25uM intervals through the whole cell. Z-stacks were collapsed using Max intensity settings on FIJI.

Fluorescence Recovery After Photobleaching (FRAP)

FRAP was performed on a Nikon A1R confocal microscope. Images were acquired every 2 sec for 90 sec (45 frames). The first few frames were collected before the bleach pulse to determine the baseline fluorescence. A circular region of interest (ROI) was selected for bleaching puncta with 100% laser power (488 nm). FRAP curves were generated used the FIJI plugin FRAP Profiler V2. Fluorescence was normalized to the whole cell to account for bleaching over the timecourse of the experiment. The difference between the normalized FRAP intensity before and at the first frame after bleach pulse was calculated and normalized to 100%. FRAP recovery measurements were averaged over 10-15 replicates in different cells. The immobile fraction was estimated as the percentage of fluorescence intensity unrecovered when the recovery curve levels out. The half-life was calculated as the amount of time for recovery to reach half of the final recovered fluorescence.

OptoDroplet Residency Assay

OptoDroplet plasmids from Cliff Brangwynne's lab were ordered from Addgene. CBX2, CBX4, and CBX8 cDNA was cloned into the plasmid backbone to generate the fusion protein with Cry2, and mCherry. NIH-3T3 cells expressing these constructs were engineered as described

above. Cell lines were seeded on round glass bottom dishes from Mattek at $0.5*10^5$ cells/mL and incubated overnight in imaging media, DMEM without phenol red supplemented with 10% FBS. Twenty-four hours later, the samples were imaged on a confocal Nikon A1R with a temperature-controlled live cell chamber at 63x magnification. The Cry2 protein was activated by a minute of blue light exposure (488 nm). Then, the blue light was turned off and the mcherry signal was imaged for 2 min. Round ROIs of the induced puncta were selected using FIJI and the fluorescence of these ROIs was quantified across the timecourse. The average slope for fluorescence change over the timecourse was measured as the rate of droplet dissolution.

References

1. Schuettengruber, B., Bourbon, H. M., Di Croce, L., and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell*, *171*(1), 34–57. https://doi.org/10.1016/j.cell.2017.08.002

2. Ma, R. G., Zhang, Y., Sun, T. T., and Cheng, B. (2014). Epigenetic regulation by polycomb group complexes: focus on roles of CBX proteins. *Journal of Zhejiang University. Science. B*, *15*(5), 412–428. https://doi.org/10.1631/jzus.B1400077

3. Aloia, L., Di Stefano, B., and Di Croce, L. (2013). Polycomb complexes in stem cells and embryonic development. *Development*, *140*(12), 2525–2534. https://doi.org/10.1242/dev.091553

4. Kim, J. J., and Kingston, R. E. (2022). Context-specific Polycomb mechanisms in development. *Nature reviews. Genetics*, *23*(11), 680–695. https://doi.org/10.1038/s41576-022-00499-0

5. Kaustov, L., Ouyang, H., Amaya, M., Lemak, A., Nady, N., Duan, S., Wasney, G. A., Li, Z., Vedadi, M., Schapira, M., Min, J., and Arrowsmith, C. H. (2011). Recognition and specificity determinants of the human cbx chromodomains. *The Journal of biological chemistry*, *286*(1), 521–529. https://doi.org/10.1074/jbc.M110.191411

6. Simon, J. A. and Kingston, R. E. (2013). Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Molecular cell*, *49*(5), 808–824. https://doi.org/10.1016/j.molcel.2013.02.013

7. Vincenz, C. and Kerppola, T. K. (2008). Different polycomb group CBX family proteins associate with distinct regions of chromatin using nonhomologous protein sequences. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(43), 16572–16577. https://doi.org/10.1073/pnas.0805317105

8. Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Molecular and cellular biology*, *26*(7), 2560–2569. https://doi.org/10.1128/MCB.26.7.2560-2569.2006

9. Connelly, K. E., Weaver, T. M., Alpsoy, A., Gu, B. X., Musselman, C. A., and Dykhuizen, E. C. (2019). Engagement of DNA and H3K27me3 by the CBX8 chromodomain drives chromatin association. *Nucleic acids research*, *47*(5), 2289–2305. https://doi.org/10.1093/nar/gky1290

10. Grau, D. J., Chapman, B. A., Garlick, J. D., Borowsky, M., Francis, N. J., and Kingston, R. E. (2011). Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes & development*, *25*(20), 2210–2221. https://doi.org/10.1101/gad.17288211

11. Plys, A. J., Davis, C. P., Kim, J., Rizki, G., Keenen, M. M., Marr, S. K., and Kingston, R. E. (2019). Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes & development*, *33*(13-14), 799–813. https://doi.org/10.1101/gad.326488.119

12. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

13. Musselman, C. A. and Kutateladze, T. G. (2021). Characterization of functional disordered regions within chromatin-associated proteins. *iScience*, *24*(2), 102070. https://doi.org/10.1016/j.isci.2021.102070

14. Cermakova, K. and Hodges, H. C. (2023). Interaction modules that impart specificity to disordered protein. *Trends in biochemical sciences*, *48*(5), 477–490. https://doi.org/10.1016/j.tibs.2023.01.004

15. Bah, A. and Forman-Kay, J. D. (2016). Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. *The Journal of biological chemistry*, *291*(13), 6696–6705. https://doi.org/10.1074/jbc.R115.695056

16. Kawaguchi, T., Machida, S., Kurumizaka, H., Tagami, H., and Nakayama, J. I. (2017). Phosphorylation of CBX2 controls its nucleosome-binding specificity. *Journal of biochemistry*, *162*(5), 343–355. https://doi.org/10.1093/jb/mvx040

17. Wang, W., Qiao, S., Li, G., Cheng, J., Yang, C., Zhong, C., Stovall, D. B., Shi, J., Teng, C., Li, D., and Sui, G. (2022). A histidine cluster determines YY1-compartmentalized coactivators and chromatin elements in phase-separated enhancer clusters. *Nucleic acids research*, *50*(9), 4917–4937. https://doi.org/10.1093/nar/gkac233

18. Salichs, E., Ledda, A., Mularoni, L., Albà, M. M., and de la Luna, S. (2009). Genome-wide analysis of histidine repeats reveals their role in the localization of human proteins to the nuclear speckles compartment. *PLoS genetics*, *5*(3), e1000397. https://doi.org/10.1371/journal.pgen.1000397

19. Fuhs, S. R. and Hunter, T. (2017). pHisphorylation: the emergence of histidine phosphorylation as a reversible regulatory modification. *Current opinion in cell biology*, *45*, 8–16. https://doi.org/10.1016/j.ceb.2016.12.010

20. Zhan, X., Yang, J., Mao, Z., and Yu, W. (2018). PIM1-catalyzed CBX8 phosphorylation promotes the oncogene-induced senescence of human diploid fibroblast. *Biochemical and*

biophysical research communications, *501*(3), 779–785. https://doi.org/10.1016/j.bbrc.2018.05.070

21. Klauke, K., Radulović, V., Broekhuis, M., Weersing, E., Zwart, E., Olthof, S., Ritsema, M., Bruggeman, S., Wu, X., Helin, K., Bystrykh, L., and de Haan, G. (2013). Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. *Nature cell biology*, *15*(4), 353–362. https://doi.org/10.1038/ncb2701

22. van Wijnen, A. J., Bagheri, L., Badreldin, A. A., Larson, A. N., Dudakovic, A., Thaler, R., Paradise, C. R., and Wu, Z. (2021). Biological functions of chromobox (CBX) proteins in stem cell self-renewal, lineage-commitment, cancer and development. *Bone*, *143*, 115659. https://doi.org/10.1016/j.bone.2020.115659

23. Liu, B., Liu, Y. F., Du, Y. R., Mardaryev, A. N., Yang, W., Chen, H., Xu, Z. M., Xu, C. Q., Zhang, X. R., Botchkarev, V. A., Zhang, Y., and Xu, G. L. (2013). Cbx4 regulates the proliferation of thymic epithelial cells and thymus function. *Development*, *140*(4), 780–788. https://doi.org/10.1242/dev.085035

24. Kim, H. J., Park, J. W., Kang, J. Y., and Seo, S. B. (2021). Negative Regulation of Erythroid Differentiation via the CBX8-TRIM28 Axis. *Molecules and cells*, *44*(7), 444–457. https://doi.org/10.14348/molcells.2021.0012

25. Morey, L., Pascual, G., Cozzuto, L., Roma, G., Wutz, A., Benitah, S. A., and Di Croce, L. (2012). Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. *Cell stem cell*, *10*(1), 47–62. https://doi.org/10.1016/j.stem.2011.12.006

26. Wang, J., Yang, B., Zhang, X., Liu, S., Pan, X., Ma, C., Ma, S., Yu, D., and Wu, W. (2023). Chromobox proteins in cancer: Multifaceted functions and strategies for modulation. *International journal of oncology*, *62*(3), 36. https://doi.org/10.3892/ijo.2023.5484

27. Zhang, J., Sun, Z., Han, Y., Yao, R., Yue, L., Xu, Y., and Zhang, J. (2017). Rnf2 knockdown reduces cell viability and promotes cell cycle arrest in gastric cancer cells. *Oncology letters*, *13*(5), 3817–3822. https://doi.org/10.3892/ol.2017.5868

28. Kim, J. J., Steinson, E. R., Lau, M. S., de Rooij, D. G., Page, D. C., and Kingston, R. E. (2023). Cell type-specific role of CBX2 and its disordered region in spermatogenesis. *Genes & development*, *37*(13-14), 640–660. https://doi.org/10.1101/gad.350393.122

29. Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., Tse, T., Wang, J., Williams, R., Trawick, B. W., Pruitt, K. D., and Sherry, S. T. (2022). Database resources of the national center for biotechnology information. *Nucleic acids research*, *50*(D1), D20–D26. https://doi.org/10.1093/nar/gkab1112

30. Sowpati, D. T., Ramamoorthy, S., and Mishra, R. K. (2015). Expansion of the polycomb system and evolution of complexity. *Mechanisms of development*, *138 Pt 2*, 97–112. https://doi.org/10.1016/j.mod.2015.07.013

31. de Potter, B., Raas, M.W.D., Seidl, M.F., Verrijzer, C. P., and Snel, B. (2023). Uncoupled evolution of the Polycomb system and deep origin of non-canonical PRC1. *Communications Biology*, *6*(1144). https://doi.org/10.1038/s42003-023-05501-x

32. Lau, M. S., Schwartz, M. G., Kundu, S., Savol, A. J., Wang, P. I., Marr, S. K., Grau, D. J., Schorderet, P., Sadreyev, R. I., Tabin, C. J., and Kingston, R. E. (2017). Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. *Science*, *355*(6329), 1081–1084. https://doi.org/10.1126/science.aah5403

33. Frottin, F., Schueder, F., Tiwary, S., Gupta, R., Körner, R., Schlichthaerle, T., Cox, J., Jungmann, R., Hartl, F. U., and Hipp, M. S. (2019). The nucleolus functions as a phase-separated protein quality control compartment. *Science*, *365*(6451), 342–347. https://doi.org/10.1126/science.aaw9157

34. Vandenbunder, B., Fourré, N., Leray, A., Mueller, F., Völkel, P., Angrand, P. O., and Héliot, L. (2014). PRC1 components exhibit different binding kinetics in Polycomb bodies. *Biology of the cell*, *106*(4), 111–125. https://doi.org/10.1111/boc.201300077

35. McSwiggen, D. T., Mir, M., Darzacq, X., and Tjian, R. (2019). Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes & development*, *33*(23-24), 1619–1634. https://doi.org/10.1101/gad.331520.119

36. Shin, Y., Berry, J., Pannucci, N., Haataja, M. P., Toettcher, J. E., and Brangwynne, C. P. (2017). Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell*, *168*(1-2), 159–171.e14. https://doi.org/10.1016/j.cell.2016.11.054

37. Jaensch, E. S., Zhu, J., Cochrane, J. C., Marr, S. K., Oei, T. A., Damle, M., McCaslin, E. Z., and Kingston, R. E. (2021). A Polycomb domain found in committed cells impairs differentiation when introduced into PRC1 in pluripotent cells. *Molecular cell*, *81*(22), 4677–4691.e8. https://doi.org/10.1016/j.molcel.2021.09.018

38. Ceballos, A. V., McDonald, C. J., and Elbaum-Garfinkle, S. (2018). Methods and Strategies to Quantify Phase Separation of Disordered Proteins. *Methods in enzymology*, *611*, 31–50. https://doi.org/10.1016/bs.mie.2018.09.037

39. Niekamp, S., Marr, S. K., Oei, T. A., Subramanian, R., and Kingston, R. E. (2024). Modularity of PRC1 composition and chromatin interaction define condensate properties. *Molecular cell*, S1097-2765(24)00181-3. Advance online publication. https://doi.org/10.1016/j.molcel.2024.03.001

40. Santanach, A., Blanco, E., Jiang, H., Molloy, K. R., Sansó, M., LaCava, J., Morey, L., and Di Croce, L. (2017). The Polycomb group protein CBX6 is an essential regulator of embryonic

stem cell identity. *Nature communications*, *8*(1), 1235. https://doi.org/10.1038/s41467-017-01464-w

41. Morey, L., Aloia, L., Cozzuto, L., Benitah, S. A., and Di Croce, L. (2013). RYBP and Cbx7 define specific biological functions of polycomb complexes in mouse embryonic stem cells. *Cell reports*, *3*(1), 60–69. https://doi.org/10.1016/j.celrep.2012.11.026

42. Brown, K., Chew, P. Y., Ingersoll, S., Espinosa, J. R., Aguirre, A., Espinoza, A., Wen, J., Astatike, K., Kutateladze, T. G., Collepardo-Guevara, R., and Ren, X. (2023). Principles of assembly and regulation of condensates of Polycomb repressive complex 1 through phase separation. *Cell reports*, *42*(10), 113136. https://doi.org/10.1016/j.celrep.2023.113136

43. Brown, K., Andrianakos, H., Ingersoll, S., and Ren, X. (2021). Single-molecule imaging of epigenetic complexes in living cells: insights from studies on Polycomb group proteins. *Nucleic acids research*, *49*(12), 6621–6637. https://doi.org/10.1093/nar/gkab304

44. Lyons, H., Veettil, R. T., Pradhan, P., Fornero, C., De La Cruz, N., Ito, K., Eppert, M., Roeder, R. G., and Sabari, B. R. (2023). Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell*, *186*(2), 327–345.e28. https://doi.org/10.1016/j.cell.2022.12.013

45. Lu, H., Yu, D., Hansen, A. S., Ganguly, S., Liu, R., Heckert, A., Darzacq, X., and Zhou, Q. (2018). Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature*, *558*(7709), 318–323. https://doi.org/10.1038/s41586-018-0174-3

46. Merrill, J. C., Melhuish, T. A., Kagey, M. H., Yang, S. H., Sharrocks, A. D., & Wotton, D. (2010). A role for non-covalent SUMO interaction motifs in Pc2/CBX4 E3 activity. *PloS one*, *5*(1), e8794. https://doi.org/10.1371/journal.pone.0008794

47. Zhao X. (2018). SUMO-Mediated Regulation of Nuclear Functions and Signaling Processes. *Molecular cell*, *71*(3), 409–418. https://doi.org/10.1016/j.molcel.2018.07.027

48. Uckelmann, M., Levina, V., Taveneau, C., Han Ng, X., Pandey, V., Martinez, J., Mendiratta, S., Houx, J., Boudes, M., Venugopal, H., Trépout, S., Zhang, Q., Flanigan, S., Li, M., Sierecki, E., Gambin, Y., Das, P. P., Bell, O., de Marco, A., and Davidovich, C. (2024). Dynamic PRC1-CBX8 stabilizes a porous structure of chromatin condensates. *bioRxiv*, 2023.05.08.539931. https://doi.org/10.1101/2023.05.08.539931

49. Parreno, V., Martinez, A. M., and Cavalli, G. (2022) Mechanisms of Polycomb group protein function in cancer. *Cell Research*, *32*, 231–253. https://doi.org/10.1038/s41422-021-00606-6

50. Stuckey, J. I., Dickson, B. M., Cheng, N., Liu, Y., Norris, J. L., Cholensky, S. H., Tempel, W., Qin, S., Huber, K. G., Sagum, C., Black, K., Li, F., Huang, X. P., Roth, B. L., Baughman, B. M., Senisterra, G., Pattenden, S. G., Vedadi, M., Brown, P. J., Bedford, M. T., Min, J., Arrowsmith, C. H., James, L. I., and Frye, S. V. (2016). A cellular chemical probe

targeting the chromodomains of Polycomb repressive complex 1. *Nature chemical biology*, *12*(3), 180–187. https://doi.org/10.1038/nchembio.2007

51. Suh, J. L., Bsteh, D., Hart, B., Si, Y., Weaver, T. M., Pribitzer, C., Lau, R., Soni, S., Ogana, H., Rectenwald, J. M., Norris, J. L., Cholensky, S. H., Sagum, C., Umana, J. D., Li, D., Hardy, B., Bedford, M. T., Mumenthaler, S. M., Lenz, H. J., Kim, Y. M., Wang, G. G., Pearce, K. H., James, L. I., Kireev, D. B., Musselman, C. A., and Bell, O. (2022). Reprogramming CBX8-PRC1 function with a positive allosteric modulator. *Cell chemical biology*, *29*(4), 555–571.e11. https://doi.org/10.1016/j.chembiol.2021.10.003

Chapter 3 : Phase separation drives Polycomb-mediated gene repression

Adapted from:

Oei, T. A.,[1,2,3] Schneider, P. C.,[2] and Kingston, R. E. (2024).[2,3] Polycomb-mediated gene repression by heterologous phase separation. *Under review.*

Affiliations:

[1]Department of Chemical Biology, Harvard University; Cambridge, MA, 02138, USA

[2]Department of Molecular Biology, Massachusetts General Hospital; Boston, MA, 02114, USA

[3]Department of Genetics, Harvard Medical School; Boston, MA, 02115, USA

3.1 Summary and acknowledgements

Polycomb Group (PcG) proteins play an essential role in development through the maintenance of repressed chromatin states. Proper axial patterning in mice, condensate formation in cells, and nucleosome compaction in vitro, are driven by the Intrinsically Disordered Region (IDR) of CBX2, a subunit of canonical Polycomb Repressive Complex 1 (cPRC1). To investigate the significance of phase separation activity, we fused a condensate deficient CBX2 IDR mutant to heterologous IDRs. The FUS IDR functionally rescued the CBX2 mutant in a neuronal lineage by inducing Polycomb body formation and suppressing neurite elongation. The fusion construct binds to similar genomic targets as wildtype cPRC1 and represses transcription. Orthogonal phase separation activity is sufficient to rescue CBX2 function, demonstrating the central role of phase separation in Polycomb-mediated epigenetic regulation.

Bioinformatics analysis for the RNA-sequencing and CUT&RUN data was performed by Philipp Schneider. The authors would like to thank Kingston lab members: Jongmin Kim, Stefan Niekamp, Christos Tsokos, MacKenzie Mauger, Philipp Schneider and Robert Kingston for critical reading of the manuscript. The protocol for the breast cancer transwell invasion assay was shared by the laboratory of Karen Cichowski at Harvard Medical School, courtesy of Patrick Loi and Naomi Olsen. The MGH Microscopy Core of the Program in Membrane Biology run by Dennis Brown and Richard Bouley provided instrumentation access and training for live cell microscopy. The MGH Sequencing Core run by Ruslan Sadreyev and staffed by Ulandt Kim provided instrumentation access and training for collecting next-generation sequencing data. This work was supported by funding from the National Institute of General Medical Sciences grant: 5R35GM131743 (REK).

3.2 Introduction to CBX2 in Polycomb epigenetic regulation

Polycomb Group (PcG) proteins play a conserved and crucial role in development by regulating chromatin accessibility, silencing genes, and maintaining cell identity.[1,2,3,4,5] Noncanonical Polycomb Repressive Complex 1 (ncPRC1) catalyzes H2AK119ub1 and this histone mark recruits Polycomb Repressive Complex 2 (PRC2) stimulating deposition of H3K27me3.[6,7,8,9] Canonical PRC1 (cPRC1) recognizes the H3K27me3 mark via the chromodomain of the CBX subunit and plays a structural role in compacting the surrounding chromatin while also generating long-distance chromatin contacts.[9,10] Consistent with their role as developmental regulators, PcG proteins are often misexpressed in cancer, leading to repression of tumor suppressor genes or maintenance of the cancer stem cell niche.[11,12] CBX2 overexpression correlates with poor prognosis in several cancers.[12,13,14]

The CBX2 protein has the strongest chromatin compaction activity of the CBX paralogs and contains a long, charged IDR referred to as the Compaction and Phase Separation (CaPS) region.[15,16,17] Mutation of twenty-three lysines and arginines in this region of CBX2 ($CBX2_{KRA}$) prevents chromatin compaction and phase separation *in vitro*, obviates the formation of nuclear puncta known as Polycomb bodies in live cells, and causes homeotic transformation in mice.[18,19,20] Polycomb bodies were first identified as nuclear foci in 1998,[21] and their role in gene regulation has been studied in several systems.[22] Phase separation has been proposed as a mechanism for genome regulation by concentrating and compartmentalizing select biomolecules in the context of a crowded nuclear milieu.[23,24,25] Numerous studies highlight the potential role for phase separation in the regulation of active and repressive chromatin,[26-40] presenting it as a compelling model for transcriptional control. Recently, different amino acid sequences that

promote phase separation have been shown to aid in concentration of specific proteins in transcriptional condensates, supporting a direct role for condensate formation in regulatory mechanisms.[28,37] However, establishing a clear link between condensate formation and regulatory function has been challenging.[38] Here, we investigate the role of phase separation in CBX2 biology, specifically in two cancer cell models with known CBX2 involvement. We show that a heterologous IDR with no biological or sequence similarity to the CBX2 IDR can rescue PRC1 epigenetic function, both at a phenotypic and at a molecular level. These experiments support a key functional role for CBX2's phase separation activity in driving Polycomb-mediated gene repression.

3.3 Heterologous IDRs induce orthogonal phase separation of Polycomb condensates

Inspired by optogenetic phase separation tools,[41] we developed a chemogenetic system to explore cellular phenotypes of Polycomb function. To test whether we could induce condensate formation of $CBX2_{KRA}$, the CBX2 phase separation deficient mutant, with heterologous IDRs, we engineered constructs with an FKBP (F36V variant) homodimerizing protein. We used well-characterized, phase separating IDRs of varied amino acid composition derived from proteins with distinct cellular functions. For example, FUS is involved in transcription, splicing, DNA damage repair and cytoplasmic stress granule formation.[42,43,44,45] HNRNPA1 plays a role in RNA metabolism, splicing, and transport.[46] MED1 regulates transcriptional activation as part of the mediator complex.[47] We generated a fusion construct including the IDR, mEGFP, FKBP(F36V), and $CBX2_{KRA}$ (**Fig. 3.1A**).[48] We also designed three control constructs: one of the wildtype CBX2 protein without an additional IDR, one of $CBX2_{KRA}$ without an IDR, and one of $CBX2_{KRA}$ with the first half of the CBX2 IDR, shown to be sufficient for phase separation

activity in previous work (**Fig. 3.1A**).[49] These constructs were stably expressed in Neuro2a (N2a) cells via lentivirus transduction. The cell lines expressing the following constructs are hereafter referred to by their shortened, bolded names shown in **Fig. 1A** and listed here: mEGFP-FKBP(F36V)-CBX2 (**CBX2$_{WT}$**), mEGFP-FKBP(F36V)-CBX2$_{KRA}$ (**CBX2$_{KRA}$**), FUS^{IDR}-mEGFP-FKBP(F36V)-CBX2$_{KRA}$ (**FUS^{IDR}**), HNRNPA1^{IDR}-mEGFP-FKBP(F36V)-CBX2$_{KRA}$ (**HNRNP^{IDR}**), MED1^{IDR}- mEGFP-FKBP(F36V)-CBX2$_{KRA}$ (**MED1^{IDR}**), CBX2^{IDR}-mEGFP-FKBP(F36V)-CBX2$_{KRA}$ (**CBX2^{IDR}**). The rapalog, AP20187, induces homodimerization of the FKBP(F36V) protein and generated bright nuclear GFP puncta for the FUS^{IDR}, HNRNP^{IDR}, MED1^{IDR}, and CBX2^{IDR} lines (**Fig. 3.1B)**. Homodimerization of FKBP(F36V) brings the fused IDRs into proximity and promotes the weak multivalent homotypic interactions that promote condensate formation. The nuclear puncta formed by different IDRs are qualitatively different in size and number as seen in live cell imaging (**Fig. 3.2A**). The CBX2$_{KRA}$ line did not generate condensates (**Fig. 3.1B)**. We conclude that heterologous IDRs can rescue CBX2 puncta formation in cells.

We performed immunofluorescence experiments with the FUS^{IDR} line to study condensate composition. The GFP signal of the FUS^{IDR} construct colocalized with that of endogenous cPRC1 subunits including CBX2, RING1B, BMI-1, and PHC2 suggesting that these condensates can recruit all components of the full cPRC1 complex (**Fig. 3.1C**). Interestingly, the FUS^{IDR} induced condensates also weakly concentrate CBX4 and CBX7, which are highly expressed in N2A cells (**Fig. 3.2B**). CBX8 is lowly expressed in N2a cells and not specifically concentrated in the condensates (**Fig. 3.2B**). This echoes recent work which demonstrates that CBX2 is mechanistically responsible for concentrating other CBX proteins with weaker phase separation

activity.[38] However, the Polycomb bodies generated in the FUS^{IDR} line do not concentrate members of the PRC2 complex such as EZH2, EED or SUZ12 (**Fig. 3.1C**). This is consistent with previous studies demonstrating that cPRC1 is more stably bound to chromatin than PRC2 and the nucleation of CBX2 condensates is independent of H3K27me3.[50,51] Additionally, non-canonical PRC1 (ncPRC1) is not enriched in the FUS^{IDR} condensates, as shown by staining of RYBP (**Fig. 3.2B**). These results agree with single molecule tracking studies that suggest cPRC1 is more stably bound to chromatin and is more enriched in Polycomb bodies than ncPRC1.[52,53] H3K27me3 and H2AK119ub1, demonstrate diffuse staining throughout the nucleus that is not specifically enriched in the induced puncta consistent with a model of histone modifications widely deposited by dynamic Polycomb complexes as a blanket across the genome.[54] Moreover, the FUS condensates did not colocalize with proteins or histone marks involved in active transcriptional processes such as the RPB1 subunit of RNA Polymerase II (**Fig. 3.1C**), H3K27ac (**Fig. 3.1C**), SMARCC1/BAF155 (**Fig. 3.1C**), or H3K4me3 (**Fig. 3.2B**). We conclude that condensates induced by the FUS^{IDR} fusion are consistent with the composition of typical Polycomb bodies in cells.

Heterologous IDRs induce Polycomb body formation of the CBX2 phase separation mutant. **(A)** Schematic of cPRC1 condensate formation induced by the FUS^{IDR} fusion to mEGFP, FKBP(F36V), and the CBX2 phase separation mutant, $CBX2_{KRA}$. Protein constructs designed for the $CBX2_{WT}$, $CBX2_{KRA}$, FUS^{IDR}, $HNRNP^{IDR}$, $MED1^{IDR}$, and $CBX2^{IDR}$ cell lines. **(B)** Z-projections of live cell images at 63X magnification of Neuro2a (N2a) cells expressing the $CBX2_{WT}$, $CBX2_{KRA}$, FUS^{IDR}, $HNRNP^{IDR}$, $MED1^{IDR}$, and $CBX2^{IDR}$ fusion constructs from (A). The scale bar is 10 μm. **(C)** Immunofluorescence staining for GFP in the FUS^{IDR} line shown pairwise with staining for cPRC1 subunits: CBX2, RING1B, BMI-1, and PHC2; histone modifications: H3K27me3, H2AK119ub1, H3K27ac; and active transcription markers: RPB1 subunit of RNA Pol II, and the SMARCC1/BAF155 subunit of SWI/SNF. Images show a single z-plane at 100x magnification. The scale bar is 10 μm.

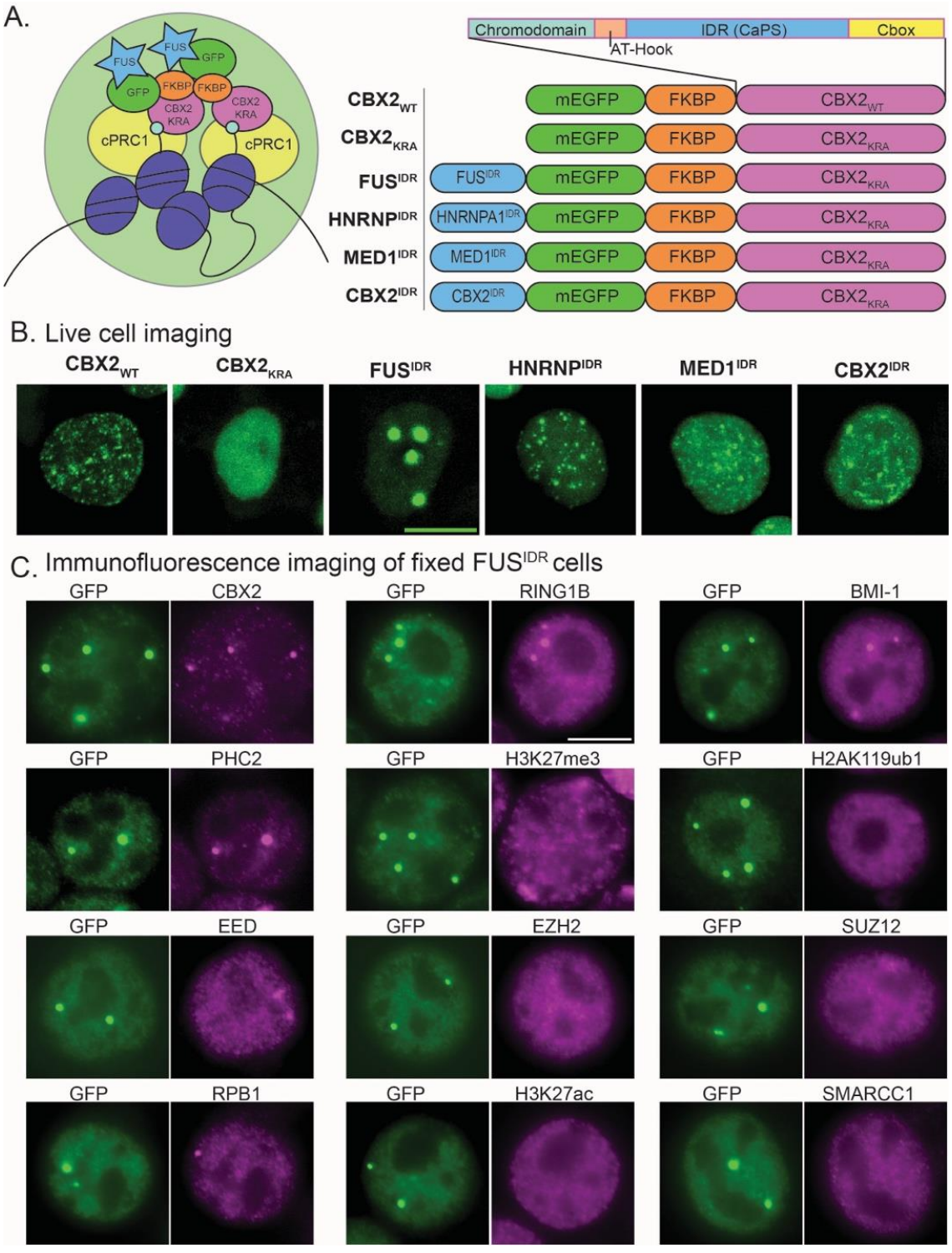


Figure 3.1 Heterologous IDRs can induce Polycomb bodies

Puncta quantification of IDR condensates, additional immunofluorescence experiments for FUS^{IDR} condensates, and images for CBX2 KD and overexpression in the neural differentiation assay*. (A)** Quantification of puncta size for the $CBX2_{KRA}$, FUS^{IDR}, $HNRNP^{IDR}$, $MED1^{IDR}$, and $CBX2^{IDR}$ lines in square pixels. *P*-values calculated by two-tailed t-test are as follows: $CBX2^{IDR}$ vs. FUS^{IDR} (*, <0.0001), $CBX2^{IDR}$ vs. $HNRNP^{IDR}$ (**, 0.0013), $CBX2^{IDR}$ vs. $MED1^{IDR}$ (N.S., 0.1156). Quantification of the average number of puncta per cell for the $CBX2_{KRA}$, FUS^{IDR}, $HNRNP^{IDR}$, $MED1^{IDR}$, and $CBX2^{IDR}$ lines. *P*-values calculated by two-tailed t-test for 15 cells per line are as follows: $CBX2^{IDR}$ vs. $CBX2_{KRA}$ (****, <0.0001), $CBX2^{IDR}$ vs. FUS^{IDR} (***, 0.0007), $CBX2^{IDR}$ vs. $HNRNP^{IDR}$ (**, 0.0014), $CBX2^{IDR}$ vs. $MED1^{IDR}$ (N.S., 0.6186). **(B)** Immunofluorescence staining for GFP in the FUS^{IDR} line shown pairwise with staining for cPRC1 subunits: CBX7, CBX4, CBX8, and PHC1; ncPRC1 subunit: RYBP; and histone mark, H3K4me3. The scale bar is 10 μm. **(C)** Additional fluorescent images of membrane staining were taken at 20x magnification after 24hr differentiation for the control shRNA and CBX2 shRNA lines. **(D)** Z-projections of live cell imaging for mEGFP-$CBX2_{WT}$ and mEGFP- $CBX2_{KRA}$ overexpression lines. Scale bar is 10 μm. **(E)** Western Blot with CBX2 and Histone 3 antibodies for the N2a, mEGFP-$CBX2_{WT}$ overexpression, and mEGFP-$CBX2_{KRA}$ overexpression lines. **(F)** Representative fluorescent images of membrane staining at 20x magnification after 24hr differentiation for the wildtype N2a, mEGFP-$CBX2_{WT}$, and mEGFP-$CBX2_{KRA}$ overexpression lines. Scale bar is 100 μm.

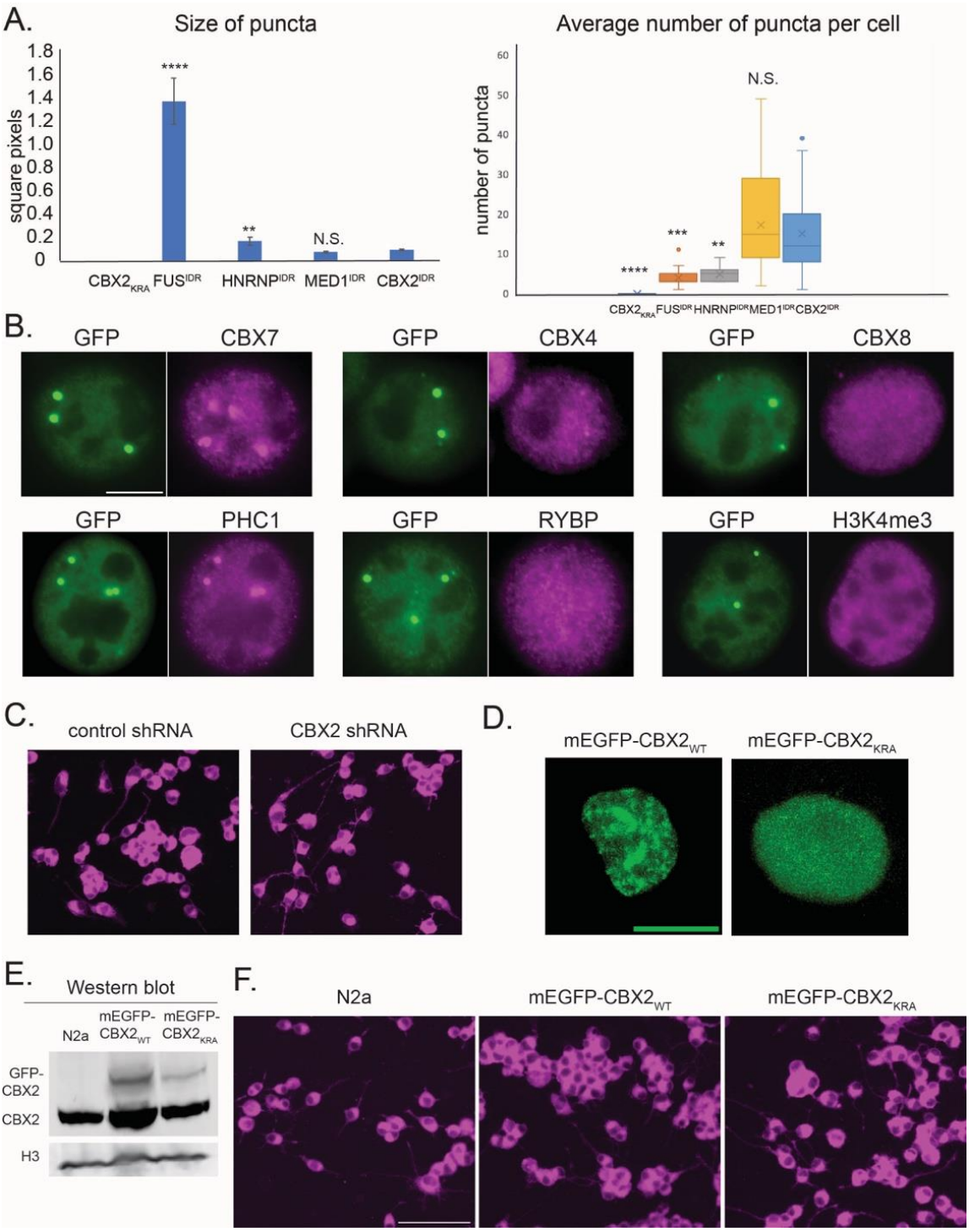


Figure 3.2 Puncta quantification, immunofluorescence, and neurite assay development

3.4 Overexpression of CBX2$_{KRA}$ fails to repress neurite elongation

To test if condensates formed by the heterologous IDRs can also rescue Polycomb biological function, we established a protocol to measure a previously reported neural differentiation phenotype in cells.[55] Data from the Cancer Dependency Map show that neuroblastomas depend significantly on CBX2 expression for tumor development.[56] Results from human neuroblastoma cell lines demonstrated that CBX2 prevents neurite outgrowth during differentiation to maintain a more cancer stem cell-like phenotype.[55] In agreement with previous studies,[55] we found that knockdown of CBX2 in the N2a mouse neuroblastoma cell line leads to a significant increase in the proportion of cells that differentiate and develop neurites, whereas overexpression of CBX2 leads to a decrease in neurite extension across three biological replicates of the differentiation (**Fig. 3.3A**). We found that 49% of cells differentiated in the control condition, while 74% of cells differentiated in the CBX2 knockdown condition (**Fig. 3.3A, 3.2C**). In contrast, overexpression of an mEGFP-tagged CBX2$_{WT}$ results in significant repression of neurite growth during differentiation, with 38% differentiation compared to the 75% differentiation of the wildtype N2a cell line (**Fig. 3.3B, 3.2D-F**). However, overexpression of mEGFP-tagged CBX2$_{KRA}$ fails to repress neurite growth. The percent of neurite-forming cells in the mutant overexpression is not significantly different from the wildtype N2a cells at 74%, but nearly double the 38% percent differentiation of the mEGFP-CBX2$_{WT}$ overexpressing cells (**Fig. 3.3B, 3.2D-F**). These results indicate that the region of CBX2 necessary for phase separation is also necessary for its pathogenic role as a negative regulator of neural differentiation to maintain a cancer stem cell-like state.

Lysine and arginine residues required for CBX2 phase separation are necessary for repression of neurite formation in N2a cells. Heterologous IDRs can phenotypically rescue the CBX2 phase separation mutant. **(A)** Knockdown of CBX2 by shRNA in N2a cells was compared to a scrambled control shRNA treatment as shown by Western Blot with CBX2 and Histone 3 antibodies. Representative fluorescent images of membrane staining were taken at 20x magnification after 24hr differentiation. Scale bar is 100 μm. Quantification of the percent of neurite-forming cells in the CBX2 knockdown experiment. The *P*-value was calculated by two-tailed t-test for three independent differentiations (~200 cells per replicate): control shRNA vs. CBX2 shRNA (**, *P*-value = 0.0024). **(B)** Quantification of the percent of neurite-forming cells in the wildtype N2a, mEGFP-$CBX2_{WT}$ overexpression, or mEGFP- $CBX2_{KRA}$ overexpression. *P*-values were calculated by two-tailed t-test for three independent differentiations (~150 cells per replicate): N2a vs. mEGFP-$CBX2_{WT}$ (**, *P*-value = 0.0012), N2a vs. mEGFP-$CBX2_{KRA}$ (N.S., *P*-value = 0.8490), mEGFP-$CBX2_{WT}$ vs. mEGFP-$CBX2_{KRA}$ (***, *P*-value = 0.0007). **(C)** Quantification of the percentage of neurite-forming cells in the $CBX2_{KRA}$, FUS^{IDR}, $HNRNP^{IDR}$, $MED1^{IDR}$, and $CBX2^{IDR}$ lines after 24hr differentiation. *P*-values were calculated by two-tailed t-test for three independent differentiations (~150 cells per replicate): $CBX2_{KRA}$ vs. FUS^{IDR} (**, *P*-value = 0.002), $CBX2_{KRA}$ vs. $HNRNP^{IDR}$ (*, *P*-value = 0.02), $CBX2_{KRA}$ vs. $MED1^{IDR}$ (N.S., *P*-value = 0.12), and $CBX2_{KRA}$ vs. $CBX2^{IDR}$ (***, *P*-value = 0.0006). Representative fluorescent images of membrane staining were taken at 20x magnification after 24hr differentiation. Scale bar is 100 μm.

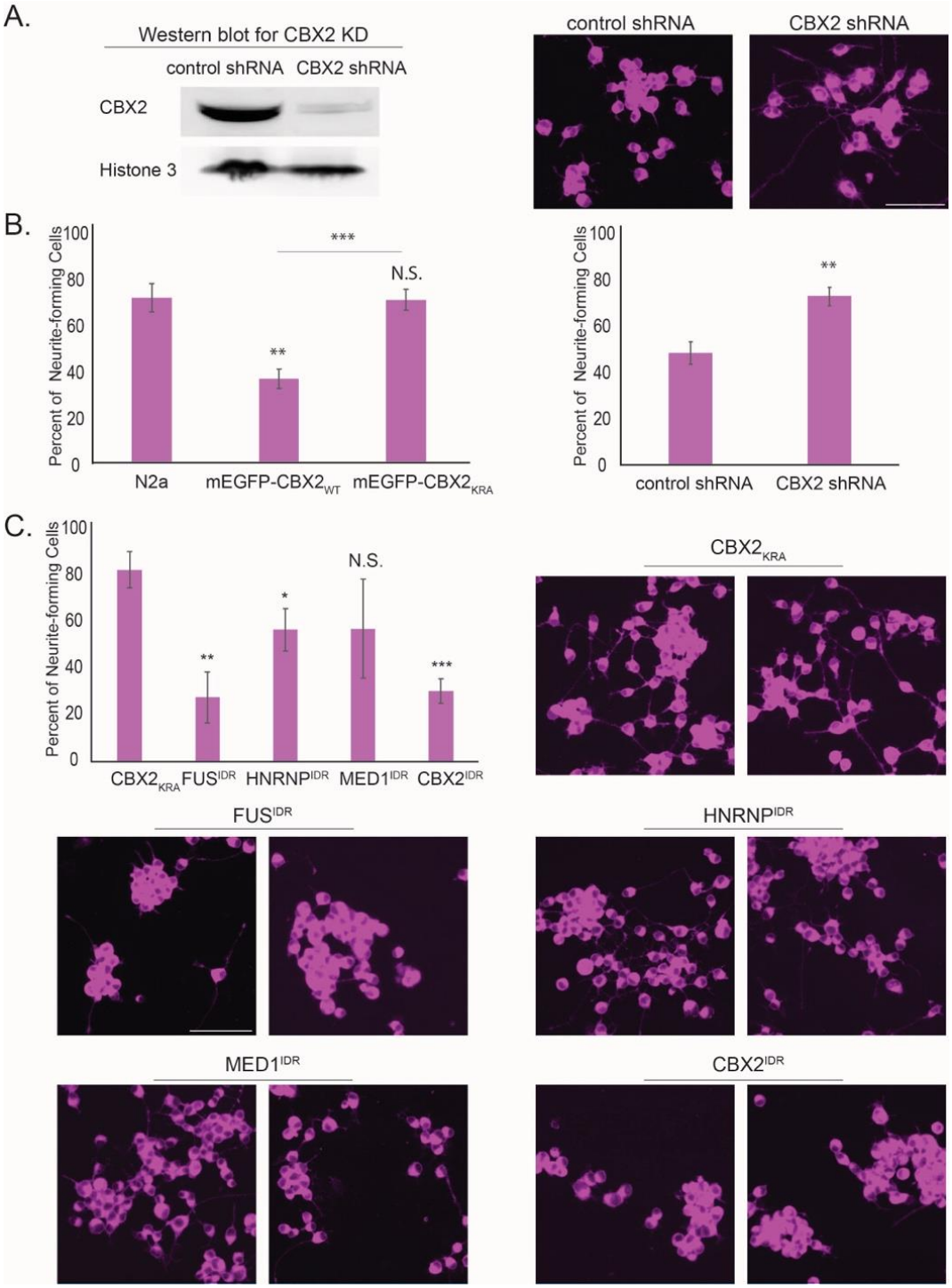


Figure 3.3 Heterologous IDRs can rescue a CBX2 mutant phenotype

3.5 The FUS^{IDR} rescues CBX2$_{KRA}$ in a neural differentiation phenotype

We used this neural differentiation assay to compare biological function of induced phase separation using the cell lines expressing the IDR fusion constructs (**Fig. 3.1A**.) Strikingly, only 27% of cells differentiate in the FUS^{IDR} cell line as compared to 80% of cells differentiating in the CBX2$_{KRA}$ cell line (**Fig. 3.3C**). The FUS^{IDR} line has similar levels of neurite formation as the CBX2^{IDR} line in that ~30% of cells differentiate in both cases. In the HNRNP^{IDR} line, neurite formation is less significantly suppressed – 55% of cells develop neurites, and the MED1^{IDR} line is not significantly different from the CBX2$_{KRA}$ line (**Fig. 3.3C**). These conclusions hold true when we change quantification strategy to average neurite length per cell (**Fig. 3.4A**). As a control, we tested overexpression of the FUS^{IDR}-mEGFP-FKBP portion of the construct alone without fusion to the CBX2 mutant. This construct does not form nuclear condensates, nor does it impact differentiation (**Fig. 3.4C**). Additionally, we tested whether the phenotypic rescue by the FUS^{IDR} requires CBX2 binding to H3K27me3 (chromodomain) or interaction of CBX2 with the rest of cPRC1 (C box) (**Fig. 3.1A**). We found that the deletion of either domain in the FUS^{IDR}-mEGFP-FKBP-CBX2$_{KRA}$ construct abolishes the neurite repression effect (**Fig. 3.4C**). These two deletions are known to strongly impair CBX2 function by preventing binding to H3K27me3 (CDdel) or by preventing the interaction of CBX2 with the rest of cPRC1 (Cboxdel). Thus, rescue by the FUS^{IDR} fusion requires the same domains for function as normal CBX2.

Neurite assay control experiments and quantification of RNA and protein in cell lines used in RNA-seq and CUT&RUN. **(A)** Alternative quantification of neurite assay for the $CBX2_{KRA}$, FUS^{IDR}, $HNRNP^{IDR}$, $MED1^{IDR}$, and $CBX2^{IDR}$ lines achieved by dividing the total measured neurite length by the total number of cells. *P*-values calculated by two-tailed t-test (~130 cells scored per replicate) are as follows: $CBX2_{KRA}$ vs. FUS^{IDR} (**, 0.0032), $CBX2_{KRA}$ vs. $HNRNP^{IDR}$ (*, 0.0219), $CBX2_{KRA}$ vs. $MED1^{IDR}$ (*, 0.0338), $CBX2_{KRA}$ vs. $CBX2^{IDR}$ (**, 0.0037). **(B)** IDR sequences for HNRNPA1 and MED1 highlighting serines, hydrophobic residues, and charged residues. **(C)** Schematic of control constructs designed for the FUS^{IDR} alone, chromodomain deletion (CDdel), and Cbox deletion (Cboxdel). Z-projections of live cell images for FUS^{IDR} alone, CDdel, and Cboxdel cell lines. The scale bar is 10 μm. Quantification of the percentage of differentiation in the neurite assay for control cell lines alongside the N2a, $CBX2_{KRA}$, and FUS^{IDR} lines. *P*-values calculated by two-tailed t-test for three replicates (~150 cells scored per replicate) are as follows: N2a vs. FUS^{IDR} alone (N.S., 0.8965), N2a vs. $CBX2_{KRA}$ (**, 0.0049), N2a vs. FUS^{IDR} (**, 0.0031), N2a vs. CDdel (N.S., 0.5006), N2a vs. Cboxdel (N.S., 0.3646), FUS^{IDR} vs. CDdel (**, 0.006), FUS^{IDR} vs. Cboxdel (**, 0.0081). **(D)** Western Blot for cell lines used in RNA-seq and CUT&RUN analysis including wildtype N2a, $CBX2_{WT}$, $CBX2_{KRA}$, FUS^{IDR}, and $CBX2^{IDR}$. Plot of mRNA expression in CPM of the GFP, $CBX2^{IDR}$ and $CBX2_{KRA}$ sequences for each of the cell lines at 0 and 24hr time points.

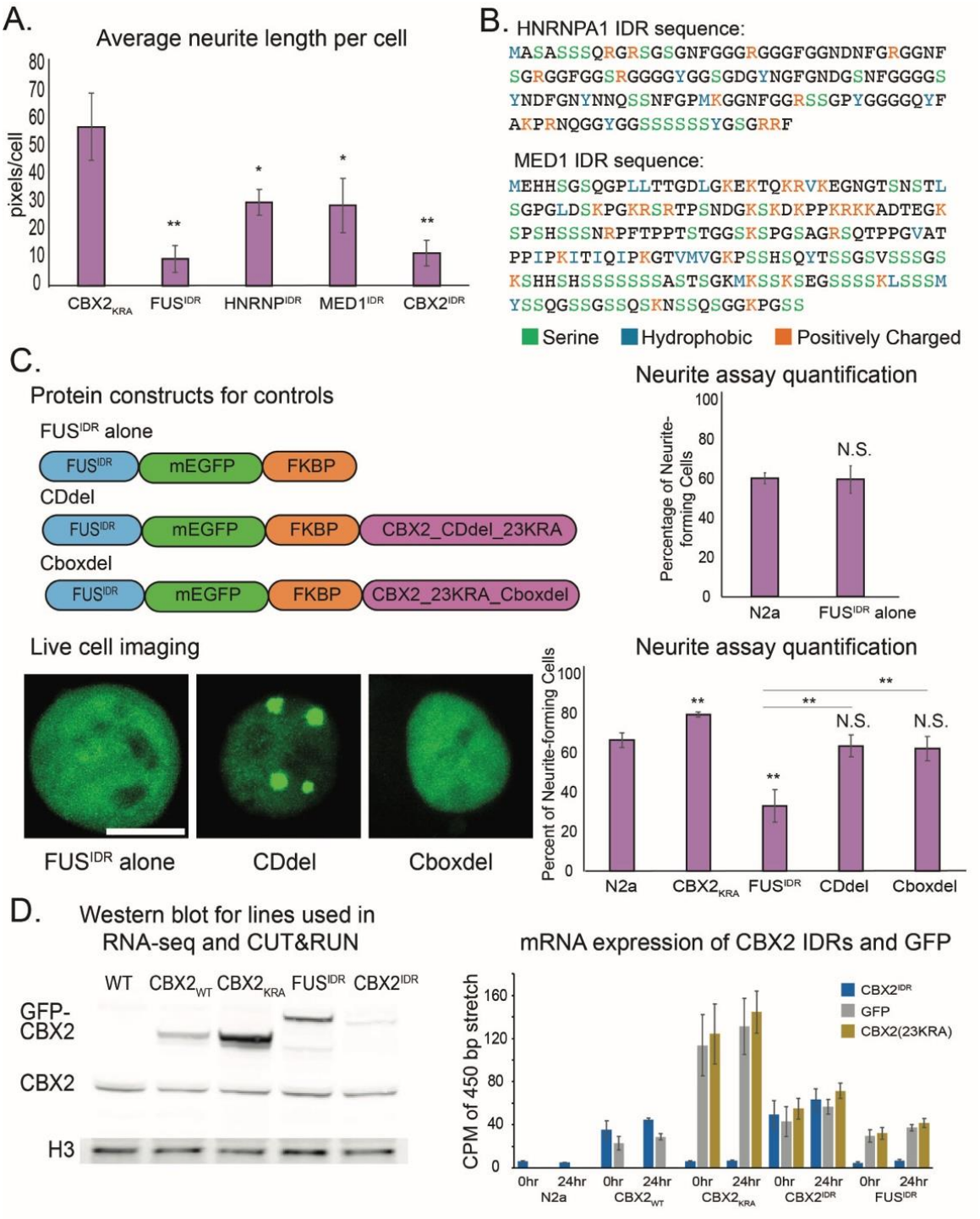


Figure 3.4 Additional neurite assay experiments

We further contextualized the function of the FUS^{IDR} line by comparison to wildtype N2a cells or the $CBX2_{WT}$ line. On average, 66% of wildtype cells differentiate (**Fig. 3.5A**). Differentiation in the $CBX2_{WT}$ line is reduced to 37%, in contrast to the $CBX2_{KRA}$ line in which 79% of cells form neurites (**Fig. 3.5A**). The fusion of the FUS^{IDR} or the $CBX2^{IDR}$ to the CBX2 phase separation mutant suppressed differentiation to 33% and 41%, respectively (**Fig. 3.5A**). These data confirm the functional rescue; the $CBX2_{WT}$ line, the FUS^{IDR} line, and the $CBX2^{IDR}$ line all have a similar phenotype in the neural differentiation assay. This result is surprising given the differences in amino acid composition between the FUS IDR and the CBX2 IDR. Interestingly, both IDRs are enriched for serine residues, but the CBX2 IDR is highly charged, whereas the FUS IDR is hydrophobic in nature (**Fig. 3.5B**). For comparison, the HNRNPA1 IDR is enriched with glycine residues and the MED1 IDR is both charged, and serine enriched (**Fig. 3.4B**). Together with the imaging data, the neurite assay results suggest that the FUS^{IDR} induced condensates are capable of silencing CBX2 target genes to suppress neural differentiation.

3.6 The FUS^{IDR} rescues transcriptional repression of neurogenesis genes

We next tested if the phenotypic rescue by the FUS^{IDR} in the neurite assay is reflected in transcriptional changes at PRC1 target genes. To this end, we performed genome-wide transcription and binding analysis in the wildtype N2a, $CBX2_{WT}$, $CBX2_{KRA}$, FUS^{IDR}, and $CBX2^{IDR}$ cell lines. First, we analyzed gene expression by RNA-seq for three biological replicates of the differentiation across three time points (0hr, 12hr, 24hr). The MDS plot of the RNA-seq samples (**Fig. 3.5C**) shows differentiation driving the y-axis variation, separating all 0hr samples from later timepoints. The FUS^{IDR} replicates cluster towards the N2a, $CBX2_{WT}$, and $CBX2^{IDR}$ samples and away from the $CBX2_{KRA}$ samples. The shift of the FUS^{IDR} replicates away

from the $CBX2_{KRA}$ suggests a nearly complete rescue of gene expression by fusion of the CBX2 phase separation mutant to the FUS^{IDR}.

When genome-wide expression was analyzed, we found that CBX2 overexpression downregulates neurogenesis genes as compared to wildtype N2a cells. At the 24hr timepoint, the majority of Differentially Expressed Genes (DEGs) are downregulated in the $CBX2_{WT}$ line (**Fig. 3.5D**). Many of the downregulated genes are represented in the Gene Ontology (GO) term for neurogenesis pathways (**Fig. 3.5D**).[57,58,59] The *Npr2* gene, a neurogenesis gene involved in neuron projection during differentiation,[60,61] is one of the most significantly affected genes (**Fig. 3.5D**). To determine whether impaired phase separation impacts transcriptional regulation, we compared gene expression in the $CBX2_{WT}$ and $CBX2_{KRA}$ lines. The majority of DEGs were upregulated in the $CBX2_{KRA}$ line and these genes are enriched for neurogenesis related pathways including *Npr2* (**Fig. 3.5E**). The upregulation of neurogenesis genes is consistent with the observed increase in neural differentiation of the $CBX2_{KRA}$ line. We compared gene expression patterns in the $CBX2_{KRA}$ line and the FUS^{IDR} line to determine the effect of the heterologous IDR on genome-wide transcription. A volcano plot of the DEGs (**Fig. 3.5F**) demonstrates that neurogenesis genes were also downregulated in the FUS^{IDR} line, when compared to the $CBX2_{KRA}$ line. The rescue of neurogenesis gene repression by the FUS^{IDR}, including *Npr2*, is consistent with the phenotypic result from the neurite assay. A comparison of the genes changing in the $CBX2_{WT}$ line vs. the $CBX2_{KRA}$ line and those changing in the FUS^{IDR} line vs. the $CBX2_{KRA}$ line demonstrates strong correlation (Pearson correlation coefficient = 0.768) (**Fig. 3.6A**). Together, these data suggest that the FUS^{IDR} rescues transcriptional repression of neurogenesis genes.

We examined individual Polycomb target genes; *Lhx2* and *Npr2* are shown in **Fig. 3.5G**. *Lhx2* is a known Polycomb target and homeobox transcription factor involved in regulating neuron development and differentiation from ES cells.[62] For *Lhx2*, RING1B occupancy is similar to the wildtype N2a line for all overexpression cell lines including $CBX2_{WT}$, $CBX2_{KRA}$, FUS^{IDR}, and $CBX2^{IDR}$ (**Fig. 3.5G**). However, *Lhx2* is misexpressed in the $CBX2_{KRA}$ line. The FUS^{IDR} fusion reduces *Lhx2* transcription to similar levels as the $CBX2_{WT}$ or $CBX2^{IDR}$ lines. Another example of FUS^{IDR} rescue is the *Npr2* gene. In accordance with their respective differentiation phenotypes, *Npr2* is more highly expressed in the N2a and $CBX2_{KRA}$ lines relative to the $CBX2_{WT}$, FUS^{IDR} or $CBX2^{IDR}$ lines (**Fig. 3.5G**). Therefore, the FUS^{IDR} restores the repressive nature of cPRC1 bound at the *Npr2* TSS directly impacting the neurite extension phenotype. Additional examples of Polycomb target genes, including *Bex4*, *Ccnd2*, *Nrbp2*, and *Foxf1*, also demonstrate this rescue effect (**Fig. 3.6B**).

FUS^{IDR} phenotypically rescues the CBX2 phase separation mutant similarly to the CBX2^{IDR} and represses neurogenesis genes. **(A)** Quantification of the percentage of neurite-forming cells in the N2a, CBX2$_{WT}$, CBX2$_{KRA}$, FUS^{IDR}, and CBX2^{IDR} lines after 24hr differentiation. *P*-values were calculated by two-tailed t-test for three independent differentiations (~200 cells per replicate): N2a vs. CBX2$_{WT}$ (**, *P*-value = 0.0022), N2a vs. CBX2$_{KRA}$ (**, *P*-value = 0.0049), N2a vs. FUS^{IDR} (**, *P*-value = 0.0031), N2a vs CBX2^{IDR} (**, *P*-value = 0.0021), CBX2$_{KRA}$ vs. CBX2$_{WT}$ (***, *P*-value = 0.0003), CBX2$_{KRA}$ vs. FUS^{IDR} (***, *P*-value = 0.0006), and CBX2$_{KRA}$ vs. CBX2^{IDR} (***, *P*-value = 0.0002). **(B)** Comparison of CBX2 IDR sequence and FUS IDR sequence amino acid composition highlighting serine residues, hydrophobic residues, and positively charged residues. **(C)** RNA-seq MDS plot of 0, 12, and 24hr differentiation timepoints for three biological replicates of the N2a, CBX2$_{WT}$, CBX2$_{KRA}$, FUS^{IDR}, and CBX2^{IDR} cell lines. **(D)** RNA-seq Differentially Expressed Genes (DEGs) at the 24hr timepoint are shown in a volcano plot for CBX2$_{WT}$ vs. N2a with highlighted neurogenesis genes. **(E)** RNA-seq DEGs at the 24hr timepoint are shown in a volcano plot for CBX2$_{WT}$ vs. CBX2$_{KRA}$ with highlighted neurogenesis genes. **(F)** RNA-seq DEGs at the 24hr timepoint are shown in a volcano plot for FUS^{IDR} vs. CBX2$_{KRA}$ with highlighted neurogenesis genes. **(G)** IGV browser tracks for *Lhx2* and *Npr2* showing RING1B normalized CUT&RUN signal and RNA-seq transcription for N2a, CBX2$_{WT}$, CBX2$_{KRA}$, FUS^{IDR}, and CBX2^{IDR} lines at the 24hr differentiation timepoint.

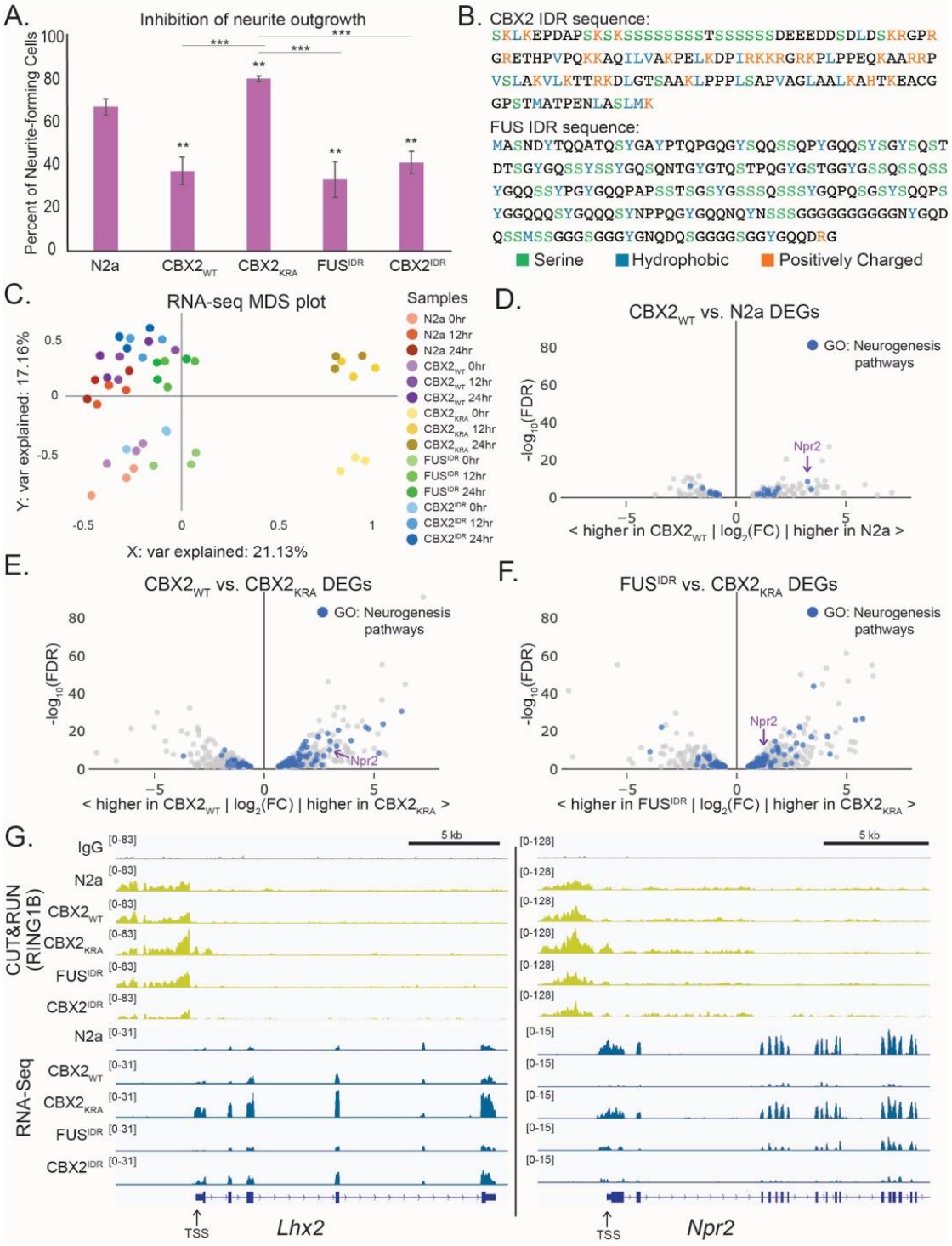


Figure 3.5 Heterologous phase separation rescues gene repression

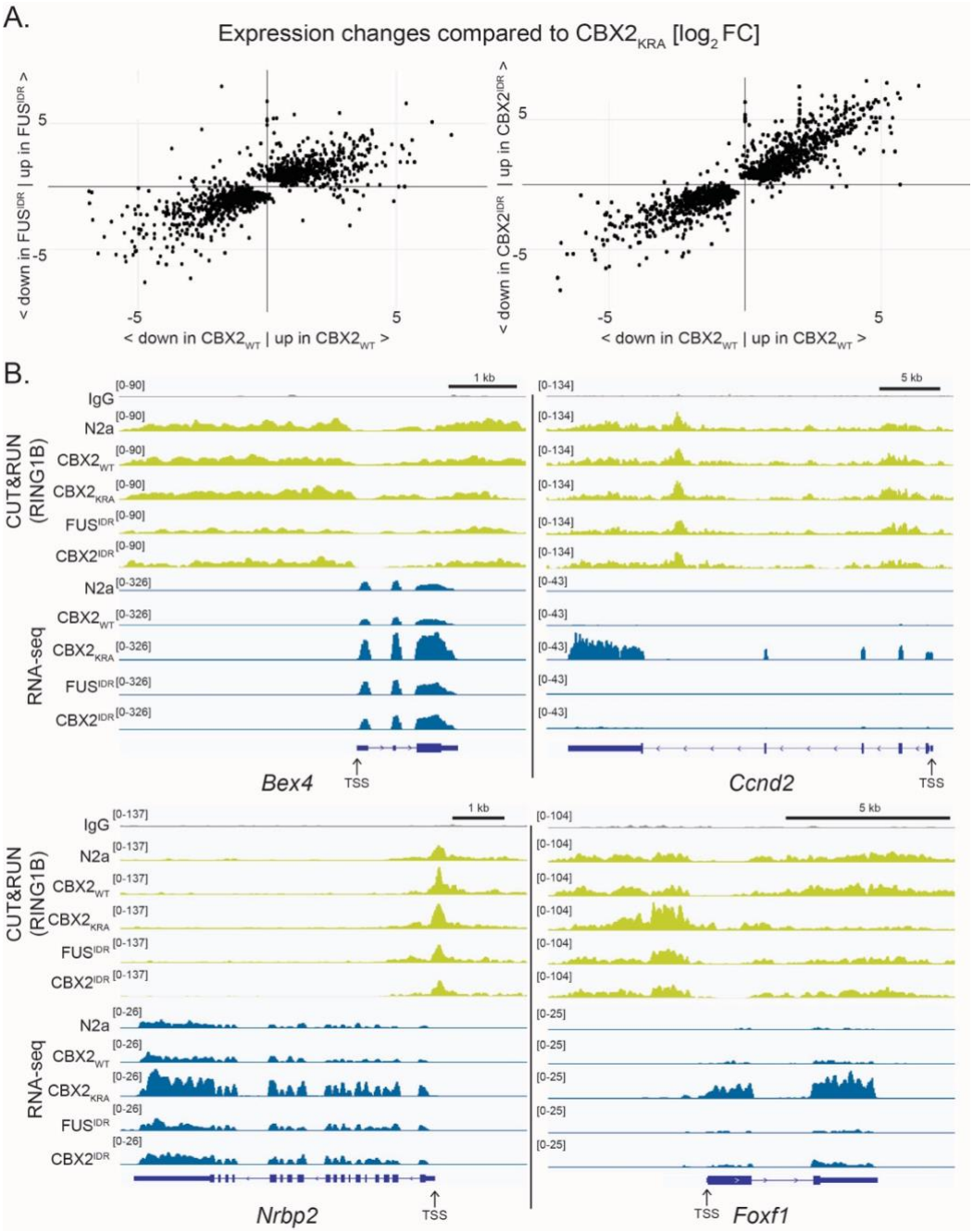


Figure 3.6 The FUS^{IDR} and CBX2^{IDR} both silence transcription

Correlation of FUS^{IDR} and CBX2^{IDR} effects on gene expression and additional IGV browser tracks of Polycomb target genes. **(A)** Scatterplots showing the correlation between DEGs of Fus^{IDR} vs. CBX2$_{KRA}$ against the DEGs of CBX2 vs. CBX2$_{KRA}$ (Pearson correlation coefficient = 0.768) and correlation between DEGs of CBX2^{IDR} vs. CBX2$_{KRA}$ against the DEGs of CBX2 vs. CBX2$_{KRA}$ (Pearson correlation coefficient = 0.901). **(B)** IGV browser tracks of RING1B normalized CUT&RUN and RNA-seq for Polycomb targets *Bex4*, *Ccnd2*, *Nrbp2*, and *Foxf1* for the N2a, CBX2$_{WT}$, CBX2$_{KRA}$, FUS^{IDR}, and CBX2^{IDR} cell lines at the 24hr timepoint.

3.7 PRC1 target loci are mostly identical in the wildtype, mutant and rescue cell lines

Gene expression changes in cell lines with various altered versions of CBX2 might be caused by differences in cPRC1-mediated gene silencing, in cPRC1 localization, or in cPRC1 bound occupancy. To determine the relationship between binding by PcG components and transcriptional changes, CUT&RUN was performed with antibodies for GFP, CBX2, RING1B, H3K27me3, and H2AK119ub1. The GFP signal tracks the localization of the fusion constructs. There were approximately 11,000 RING1B peaks that showed significant fold enrichment over IgG in the wildtype N2a cell line. This number increased slightly with overexpression of CBX2, and the GFP peaks in the $CBX2_{WT}$ line almost completely overlap with RING1B peaks and CBX2 peaks (**Fig. 3.7A**). Overexpression of the $CBX2_{KRA}$ or FUS^{IDR} construct led to fewer significant cPRC1 binding sites as indicated by the modest reductions in GFP, CBX2, and RING1B signal demonstrated in the binding graphs (**Fig. 3.7A**). The positively charged lysine and arginine residues in CBX2's IDR might help mediate CBX2 binding to the negatively charged DNA backbone, which is consistent with the recovery of GFP, CBX2, and RING1B signal in the $CBX2^{IDR}$ line (**Fig. 3.7A**). The overlap of GFP and RING1B binding sites or GFP and CBX2 binding sites was highly significant for each cell line, confirming that the fusion constructs have similar binding to endogenous CBX2 and normal recruitment of cPRC1 (**Fig. 3.7A**). Analysis of GFP peaks and RING1B peaks demonstrates that nearly all peaks in the $CBX2_{KRA}$ and FUS^{IDR} lines are a subset of the peaks in the $CBX2_{WT}$ line (**Fig. 3.7B-C**). This confirms that there are very few loci where the $CBX2_{KRA}$ or FUS^{IDR} constructs have mislocalized (**Fig. 3.7B-C**). This is also consistent when looking at the overlap of GFP peaks in the $CBX2_{KRA}$ and FUS^{IDR} lines with CBX2 peaks from the wildtype N2A line (**Fig. 3.8A**). Localization and levels of H3K27me3 and H2AK119ub1 in all cell lines are largely similar, consistent with

retained function of the Polycomb system and an absence of global disruption of these covalent modifications (**Fig. 3.8B**). We conclude that the rescue of gene regulation by the FUS^{IDR} fusion was not caused by mislocalization of fusion constructs to non-CBX2 targets, by changes in ability to recruit cPRC1, or by increased cPRC1 occupancy at target genes. Instead, we observe a trend to decreased binding in both the $CBX2_{KRA}$ and FUS^{IDR} lines.

3.8 FUS^{IDR} fusion rescues gene repression of bound genes

A volcano plot of DEGs in the FUS^{IDR} line vs. the $CBX2_{KRA}$ line shows common RING1B target genes are more highly expressed in the $CBX2_{KRA}$ line (**Fig. 3.7D**). This is also true for GFP target genes as demonstrated by a scatterplot showing an increased number of significantly altered targets skewed towards lower expression in the FUS^{IDR} line (**Fig. 3.8C**). Additionally, changes in H3K27me3 signal strongly correlate with changes in gene expression, supporting a Polycomb-mediated mechanism of repression (**Fig. 3.8D**). Thus, though the FUS^{IDR} does not recover cPRC1 occupancy to the same extent as the positively charged $CBX2^{IDR}$, it represses cPRC1 genes where it is bound to functionally rescue gene expression, resulting in the phenotypic rescue of neurite suppression (**Fig. 3.7E**).

Similar Polycomb targets are bound in each cell line and the FUS^{IDR} represses cPRC1 bound genes. **(A)** Binding graphs for CUT&RUN enrichment of GFP, CBX2, and RING1B over IgG at TSSs demonstrate co-binding of overexpression constructs in the CBX2$_{WT}$, CBX2$_{KRA}$, FUS^{IDR}, and CBX2^{IDR} lines at RING1B bound loci and CBX2 bound loci. **(B)** Venn diagram showing overlap of GFP CUT&RUN peaks in the CBX2$_{WT}$, CBX2$_{KRA}$, and FUS^{IDR} lines. **(C)** Venn diagram showing overlap of RING1B CUT&RUN peaks in the CBX2$_{WT}$, CBX2$_{KRA}$, and FUS^{IDR} lines. **(D)** RNA-seq DEGs at the 24hr timepoint are shown in a volcano plot for FUS^{IDR} vs. CBX2$_{KRA}$ highlighting RING1B target genes bound in both cell lines. **(E)** Table summarizing a comparison of the FUS^{IDR} and the CBX2^{IDR} in cellular condensate formation, neural differentiation, breast cancer transwell invasion, transcriptional repression, and Polycomb target gene binding.

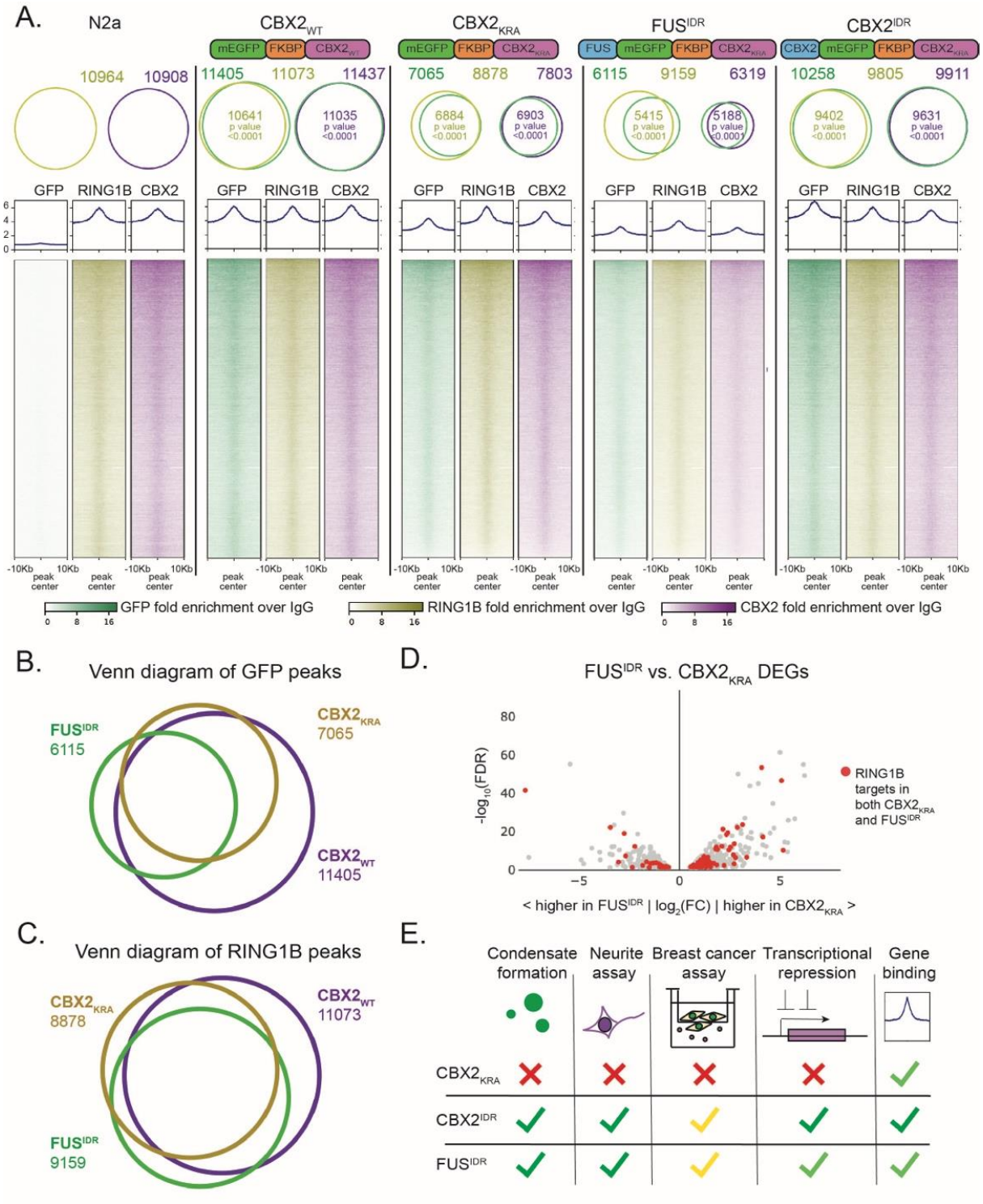


Figure 3.7 The FUS^{IDR} rescues bound Polycomb target genes

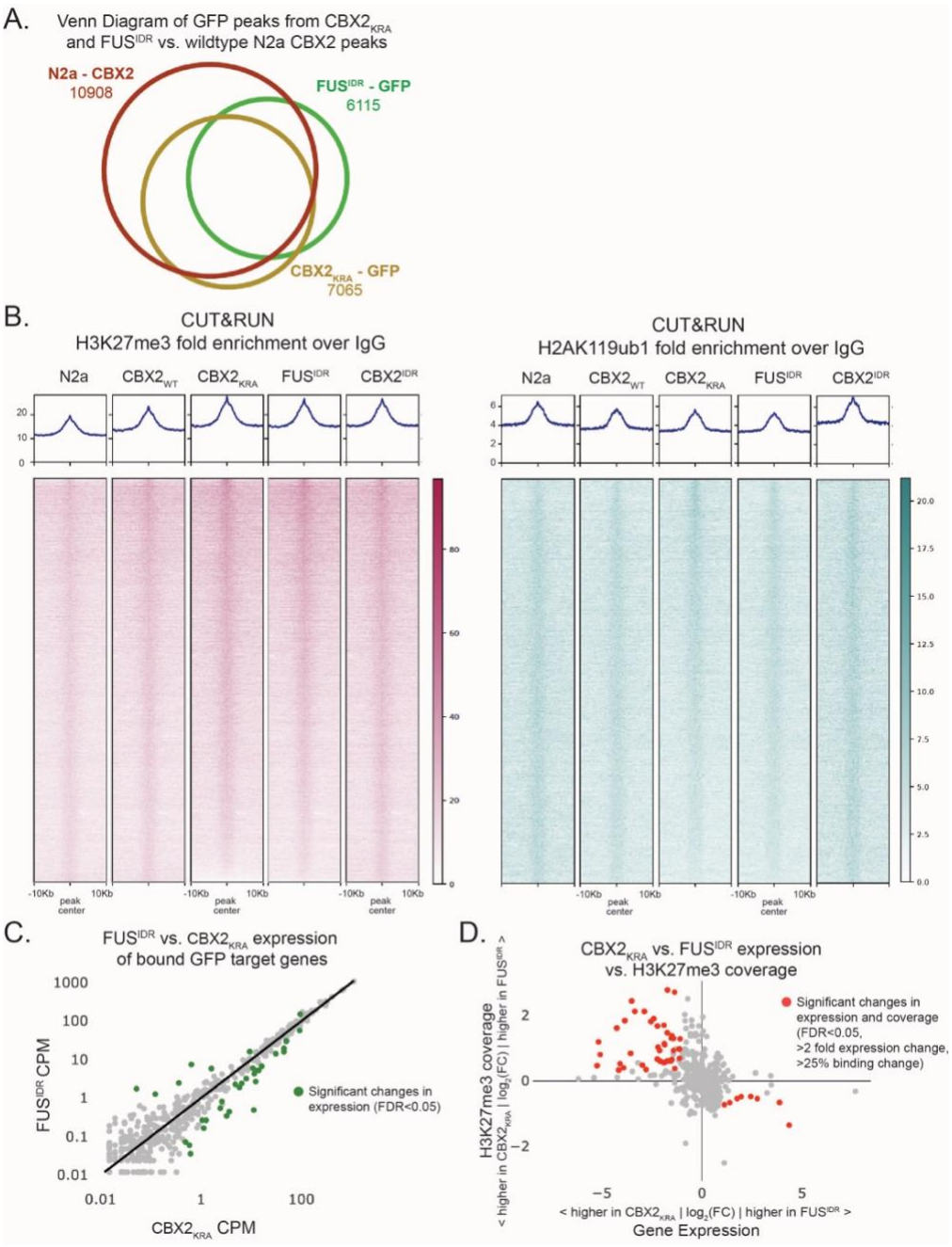


Figure 3.8 Additional CUT&RUN data for H3K27me3 and H2AK119ub1

Additional Venn diagram, H3K27me3 and H2AK119Ub1 binding graphs, and correlation of GFP and H3K27me3 peaks with gene expression changes. **(A)** Venn diagram showing the overlap of CBX2 peaks from the wildtype N2a cells with GFP peaks from the $CBX2_{KRA}$ and FUS^{IDR} cell lines. **(B)** Binding graphs showing H3K27me3 and H2AK119Ub1 enrichment over IgG in N2a, $CBX2_{WT}$, $CBX2_{KRA}$, FUS^{IDR}, and $CBX2^{IDR}$ cell lines at the 24hr timepoint. **(C)** Scatterplot of GFP target genes bound in both the $CBX2_{KRA}$ and the FUS^{IDR} lines with significant changes in expression highlighted. **(D)** Graph of DEGs for $CBX2_{KRA}$ vs. FUS^{IDR} plotted against changes in H3K27me3 coverage. Significant changes in both gene expression and H3K27me3 coverage are highlighted in red.

3.9 FUS^{IDR} fusion rescues a CBX2-dependent breast cancer phenotype

To confirm the function of CBX2 phase separation in a different model system, we studied the effects of induced phase separation in the 4T1 mouse breast cancer cell line. CBX2 is known to have a role in shaping the aggressive phenotype of triple negative breast cancer (TNBC).[63,64] A transwell invasion protocol was used to measure the aggressiveness of 4T1 cell lines expressing the mEGFP-FKBP(F36V) fusion constructs used in the neurite assay: $CBX2_{WT}$, $CBX2_{KRA}$, FUS^{IDR}- $CBX2_{KRA}$, and $CBX2^{IDR}$- $CBX2_{KRA}$ (**Fig. 3.9A**). These constructs showed similar ability to form condensates in the breast cancer cell line (**Fig. 3.9A**) as compared to the neuroblastoma cell line (**Fig. 3.1B**). In the transwell invasion assay, the FUS^{IDR} line was significantly more aggressive than the $CBX2_{KRA}$ line and had increased transwell invasion similar to the $CBX2^{IDR}$ line, although both IDR fusion lines were not as invasive as the wildtype 4T1 or the $CBX2_{WT}$ line (**Fig. 3.9B-D**). Overall, these data show that CBX2 phase separation activity also plays a role in promoting invasiveness in a breast cancer model system, demonstrating that this effect is not limited to neuroblastoma.

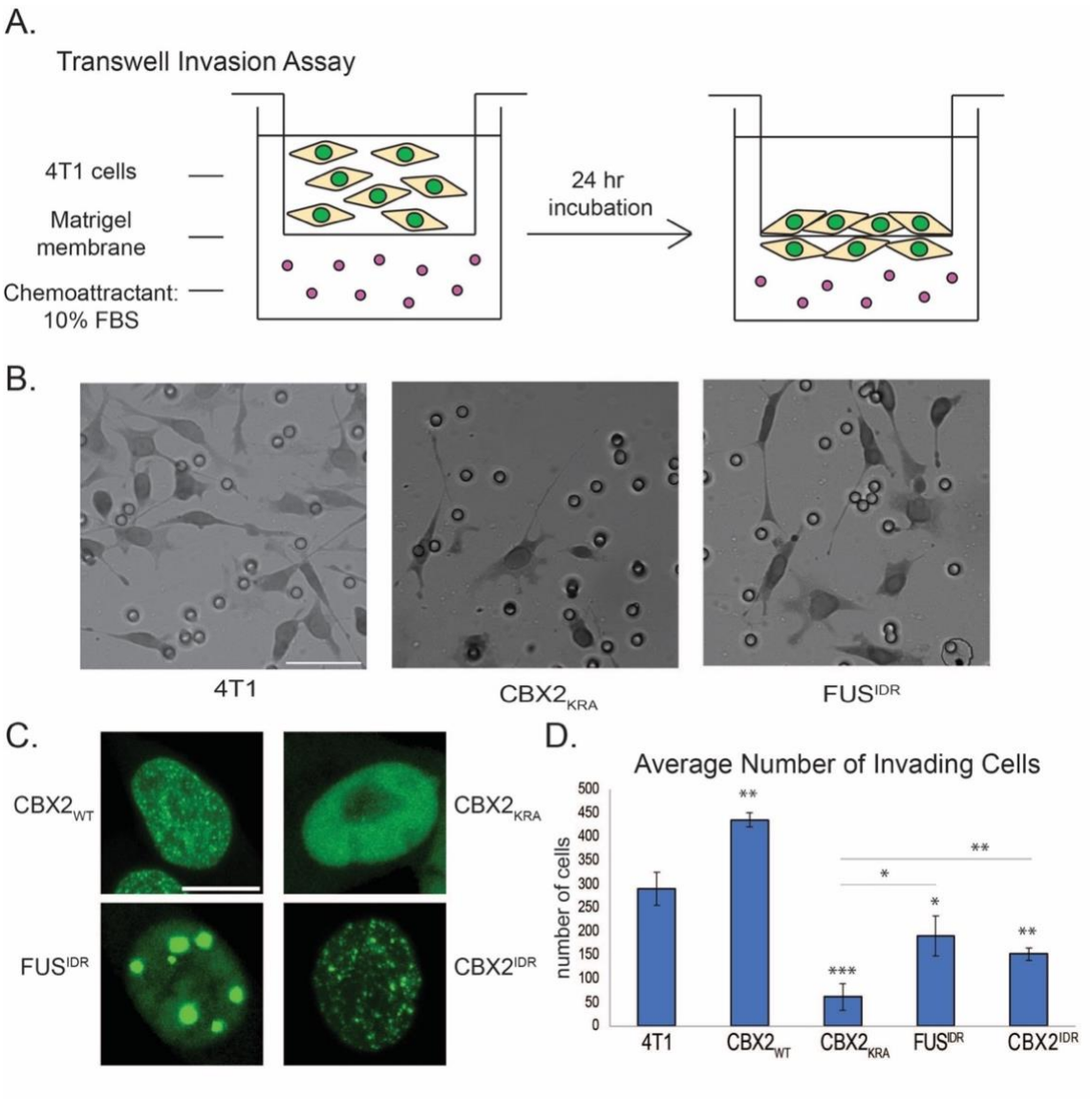


Figure 3.9 The FUS^{IDR} can rescue a CBX2 breast cancer phenotype

The FUS^{IDR} rescues the CBX2 phase separation mutant in a breast cancer transwell invasion assay. **(A)** Schematic demonstrating the experimental setup. **(B)** Sample images of invading cells for the wildtype 4T1, CBX2$_{KRA}$, and FUS^{IDR} cell lines after 24hr incubation and staining. Scale bar is 50 μm. Small round shapes are the holes of the membrane. **(C)** Z-projection of live cell images for the CBX2$_{WT}$, CBX2$_{KRA}$, FUS^{IDR}, and CBX2^{IDR} overexpressing 4T1 cell lines. The scale bar is 10 μm. **(D)** Quantification of the average number of invading cells for the 4T1, CBX2$_{WT}$, CBX2$_{KRA}$, FUS^{IDR}, and CBX2^{IDR} lines for three replicates. *P*-values calculated by two-tailed t-test are as follows: 4T1 vs. CBX2$_{WT}$ (**, 0.0026), 4T1 vs. CBX2$_{KRA}$ (**, 0.009), 4T1 vs. FUS^{IDR} (*, 0.0326), 4T1 vs. CBX2^{IDR} (**, 0.0029), CBX2$_{KRA}$ vs. FUS^{IDR} (*, 0.0117), CBX2$_{KRA}$ vs. CBX2^{IDR} (**, 0.0074).

3.10 Conclusions

The primary conclusion from this work is that phase separation makes a functional contribution to cPRC1-mediated gene repression. A heterologous IDR from the FUS protein, with distinct amino acid composition and a distinct biological regulatory role from CBX2's IDR, rescues function of a CBX2 phase separation deficient mutant at both phenotypic and transcriptional levels. This functional complementation with the FUS^{IDR} is surprising, because it is increasingly appreciated that despite their disorder, IDRs have specific functions in their native protein context.[28,37,65] Consistent with these observations, we found that the $Med1^{IDR}$ and the $HNRNPA1^{IDR}$ did not rescue the neural phenotype to the same extent as the FUS^{IDR}. Additionally, the $CBX2^{IDR}$ seems to uniquely facilitate enhanced occupancy at Polycomb target genes. This suggests that the positive charges in the $CBX2^{IDR}$ contribute to binding of Polycomb targets through electrostatic interactions with the negatively charged DNA backbone. We conclude that the unique properties of the FUS^{IDR} happen to be compatible with generating functional Polycomb condensates within this context, while the biological specificity of IDRs remains a truism. Serine phosphorylation of the IDR is known to regulate both CBX2 and FUS condensate formation, which might help to explain their compatibility in this study.[19,20,66] We will compare the dynamics of CBX2 and FUS^{IDR} condensates by performing Fluorescence After Photobleaching (FRAP). Perhaps similarities in condensate dynamics help to facilitate functional rescue by the FUS^{IDR}. We note that cPRC1 family complexes have multiple components and at least two distinct subunits, one from the CBX and one from the PHC families of proteins, are known to contribute to phase separation.[67,68] Overall, the simplest explanation for our observations is that phase separation *per se* plays a key role in generating the repressive function

of PRC1, potentially by compacting chromatin, concentrating repressive components, or excluding activating factors from the condensate.[38,39,40,69]

Rescue of transcriptional repression by the FUS^{IDR} is nearly complete (**Fig. 3.5C**). The binding localization of PRC1 on the genome is not altered substantially by any of the overexpressed CBX2 constructs examined in this study (**Fig. 3.7A-C, Fig. 3.8A-B**). The lack of complete transcriptional rescue by the FUS^{IDR} might therefore be explained by decreased occupancy and/or decreased repressive function. There is a trend toward lower levels of binding in both the $CBX2_{KRA}$ and the FUS^{IDR} lines (**Fig. 3.7A**). The significance of this is unclear, due to both the modest change observed and a poor understanding of the binding threshold required for full repression. We show that the condensates formed by the FUS^{IDR} restore the function of cPRC1 enabling significant gene repression despite decreased binding occupancy. To determine if FUS^{IDR} driven condensates sequester nonfunctional cPRC1 off the genome as an alternative mechanism of rescue, we will knockdown endogenous CBX2 in these lines with an shRNA targeting the 3'UTR of the endogenous mRNA. This will confirm that the observed rescue comes from the FUS^{IDR} driving repression rather than from endogenous CBX2 present in the cell functioning more effectively when $CBX2_{KRA}$ occupancy decreases.

These data further confirm the importance of CBX2 in neuroblastoma and breast cancer pathogenesis. Phase separation activity is necessary for CBX2 function in driving a cancer stem cell-like, tumor-promoting phenotype. This suggests that targeting CBX2 phase separation in CBX2-dependent cancers, such as these, might be a successful therapeutic strategy. Overall, this

work demonstrates the functional significance of phase separation in Polycomb epigenetic regulation, thereby providing new perspectives on CBX2 mechanism.

3.11 Materials and methods

Cell Culture

Neuro2a cells were obtained from ATCC (https://www.atcc.org/products/ccl-131) and cultured in sterile filtered DMEM with 10% Fetal Bovine Serum (FBS) and penicillin/streptomycin (100U/mL).

HEK293T cells were a gift from the lab of Prof. Paul Lehner (CITIID, Cambridge, UK). They were cultured in sterile filtered IMDM with 10% FBS and penicillin/streptomycin (100 U/mL).

4T1 cells were obtained from ATCC (https://www.atcc.org/products/crl-2539) and cultured in sterile filtered RPMI with 10% FBS and pen/strep (100U/mL). The 4T1 cell line was single cell cloned to generate all cell lines in a consistent genetic background. All cell lines were grown at 37^{o} C with 5% CO_2.

Lentiviral Cell Line Generation

Second generation lentiviral plasmids, pCMVΔR8.91 and pMD.G were transfected into HEK293T cells alongside the lentiviral vector containing the transgene of interest using the TransIT-293 Reagent from Mirus bio. The viral supernatant was harvested and filtered after 48 hours. The viral supernatant was used fresh or stored at -80^{o}C. Neuro2a or 4T1 cells were seeded at 0.5*10^5 cells/mL in a 24 well plate. After 24 hours, the media was aspirated from the 24 well plate and 1mL of the desired viral supernatant was used to transduce the Neuro2a or 4T1 cells. The cells were incubated with virus for 4-6 hours at 37^{o} C. Then, the virus was removed and replaced with the appropriate media. After 48 hours, the cells were selected in puromycin at 1ug/mL for several days. For the cell lines transduced with a GFP-containing transgene, the selected lines were sorted on a BD FACS Melody cell sorter for high GFP expression. Full

hairpin sequences for constitutive CBX2 shRNA knockdown are as follows: gatccGGAACATGAGAAGGAGGTTTTCAAGAGAAACCTCCTTCTCATGTTCCTTTTTTg, gatccAGGCTATTCCTGCTACCAATTCAAGAGATTGGTAGCAGGAATAGCCTTTTTTTg. Full hairpin sequence for the scrambled shRNA control sequence is as follows: gatccAGGCTATTCCTGCTACCAATTCAAGAGATTGGTAGCAGGAATAGCCTTTTTTTg. These sequences were cloned into the pSIREN vector backbone. Lentiviral plasmids or cell lines available upon reasonable request.

Induced Phase Separation System and Live Cell Imaging

An IDR (FUS^{IDR}, $HNRNPA1^{IDR}$, $MED1^{IDR}$, $CBX2^{IDR}$) was fused to the FKBP12 protein variant, that has a F36V mutation known to cause homodimerization. This was fused to mEGFP and the N-terminus of the CBX2(23KRA) mutant. A construct lacking any IDR was also engineered. An sv40 NLS sequence was added to the $CBX2^{IDR}$ construct to ensure nuclear localization. These constructs were integrated by lentiviral transduction into the Neuro2a or 4T1 cell lines as described above. From there, the AP20187 rapalog (Sigma-Aldrich SML2838-1MG) was used at 25nM concentration to induce homodimerization of FKBP(F36V).

Cell lines were seeded on round glass bottom dishes from Mattek at $0.5*10^5$ cells/mL and incubated overnight in imaging media, DMEM without phenol red supplemented with 10% FBS. The samples were imaged on a confocal Nikon A1R with a live cell chamber at 63x magnification. Z-stack images were generated with 0.25uM intervals. Collapsed z-stacks using Max intensity settings were scored for puncta area and number with FIJI.

Western Blots

Cells were pelleted and resuspended in RIPA buffer. They were incubated at 4^{o}C for 15 minutes. The samples were centrifuged at 14,000G for 15 minutes and the supernatant was removed to a new Eppendorf tube. The Pierce BCA Protein Assay Kit was used to measure the concentration of total protein in each sample. 60-100 ug of protein was run on a Nupage 4-12% gel in MOPS running buffer at 90V using 1ul of BioRad's protein plus ladder. A wet transfer was done overnight onto a PVDF membrane in 1X NuPage transfer buffer with 10% methanol. The membrane was activated for 5 minutes in 100% methanol prior to the transfer. The transfer was run at 100 milliAmps for 8 hours. The membrane was blocked for an hour in 5% Milk (w/v) PBS with 0.2% Tween-20 and incubated in primary antibody overnight. Antibodies used included Rabbit Polyclonal anti-CBX2 by Bethyl (A302-524A) at 1:1000 dilution and Rabbit Polyclonal anti-Histone 3 by Abcam (ab1791) at 1:5000 dilution. The membrane was washed three times for 30 minutes in total with PBS-T. The membrane was incubated with fluorescent secondary antibodies overnight at 1:10,000 dilution in 5% Milk (w/v) PBS-T and washed three times with PBS-T before imaging on an Amersham Typhoon biomolecular imager.

Immunofluorescence

Neuro2a cell lines were plated overnight on coverslips in a 12 well plate at 0.5*10^5 cells/mL with 25nM of AP20187. The media was removed and the coverslips were washed once with PBS. The cells were fixed in PBS with 1% formaldehyde for 10 minutes. The solution was removed and the coverslips were washed with PBS for 5 minutes. The coverslips were incubated in PBS with 0.2% Triton for 10 minutes. The solution was removed and the coverslips were blocked in PBS with 3% BSA and 0.05% Triton for 30 minutes. The blocking solution was

removed and primary antibody solutions in PBS with 3% BSA and 0.05% Triton were added to the coverslips. The coverslips were incubated overnight at 4ºC. From there, the primary antibody solution was removed and the coverslips were washed once with PBS with 0.05% Triton for 5 minutes. Secondary antibody solutions in PBS with 3% BSA and 0.05% Triton were added to the coverslips and incubated overnight at 4ºC. The next day, the coverslips were washed twice with PBS with 0.05% Triton and mounted on glass slides using Vectashield Mounting Medium with Dapi. The slides were imaged on a Nikon Eclipse 90i at 100x magnification. Antibodies used included GFP Invitrogen (A11120), GFP Invitrogen (A11122), CBX2 Bethyl (A302-524A) , CBX4 Cell Signaling Technology (44268), CBX8 Cell Signaling Technology (D2O8C), CBX7 Abcam (ab21873), RING1B Cell Signaling Technology (D22F2), PHC1 Active Motif (39723), PHC2 Active Motif (39661), BMI-1 Abcam (ab54897) , EZH2 Cell Signaling Technology (D2C9), EED Cell Signaling Technology (E4L6E), SUZ12 Cell Signaling Technology (D39F6), RYBP Abcam (ab5976), H3K27me3 Cell Signaling Technology (C36B11), H2AK119Ub1 Cell Signaling Technology (D27C4), H3K4me3 Cell Signaling Technology (C42D8), RPB1 Cell Signaling Technology (D8L4Y), H3K27ac Active Motif (39133), and SMARCC1/BAF155 Cell Signaling Technology (D7F8S).

Neurite Assay

Neuro2a cells were plated on coverslips in a 12 well plate at 0.5*10^5 cells/mL in 1% FBS supplemented DMEM with 10uM of trans-retinoic acid. Rapalog AP20187 was dissolved in DMSO and added to the appropriate wells at 25nM concentration. Cells were incubated for 24 hours to induce differentiation. Media was removed and the coverslips were washed once with PBS. Cells were stained with the 1:10,000 dilution of the cell membrane stain from Invitrogen's

Neurite Outgrowth Staining Kit and a 1:10,000 dilution of Hoechst 3342 in 1% formaldehyde PBS for 20 minutes. The solution was removed and the cells were washed twice with PBS. Coverslips were dried and mounted on to glass slides using 20 ul of Vectashield PLUS mounting solution. Slides were imaged on a Nikon Eclipse 90i at 20x magnification. Neurites were traced and quantified using the NeuronJ plugin for ImageJ.[70]

RNA Sequencing

The following Neuro2a cell lines were differentiated with 1% FBS DMEM and 10 uM Retinoic acid: wildtype N2a, $CBX2_{WT}$, $CBX2_{KRA}$, FUS^{IDR}, and $CBX2^{IDR}$ as defined in **Fig. 1A**. Timepoints were taken at 0, 12, and 24hrs. Cells were trypsinized and spun down at 500g for 5 minutes. They were washed once with PBS and stored at -80^{o}C until further processing. Total RNA was extracted with Trizol reagent and chloroform. The mRNA was isolated using the NEBNext Poly(A) mRNA Magnetic Isolation Module. 1 ug of mRNA was used as input for the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina. Adaptors and primers from the NEBNext Multiplex Oligos for Illumina (Dual Index Primer Set 1) were used and the libraries were sequenced on a Nextseq 2000.

RNA Sequencing Analysis

Prior to analysis, the sequencing reads underwent quality control with FastQC v0.11.9,[71] whose reports indicated that, apart from the first bases of the 5' end, the sequencing data were of high quality, with no significant issues in base quality or composition. For further processing, the initial 10 bp of all (61 bp long) reads were clipped off the 5' end. Adapter sequences were then trimmed using Trimmomatic v0.39.[72] Ribosomal reads were filtered out by alignment with

Bowtie2 v2.4.5[73] to mouse ribosomal sequences retrieved from NCBI.[74] Remaining reads were then aligned splicing-aware with STAR v2.7.10a[75] to the mouse genome mm39 using RefSeq annotations from NCBI.[76] A database lookup of unaligned reads with Kraken2[77] confirmed that all samples were free of contamination. Alignments were mapped to genomic feature annotations using featureCounts v2.0.3.[78] The thereby generated gene expression data was further analyzed and visualized in R. Differential gene expression analysis was based on the exact test provided by edgeR[79] using TMM normalization. Differentially expressed genes were determined by a significant coverage change of FDR < 0.05 (Benjamini-Hochberg) between the compared sample group triplicates and a median coverage of TPM > 1 among the compared samples. The neurogenesis reference gene set is presented in GO term GO:0022008.[57-59] MDS analysis was performed on CPM and quantile-normalized counts using the limma package v3.58.1 for R.[80] Volcano and MDS plots were generated with plotly v4.10.3.[81]

CUT&RUN

CUT&RUN experiments were performed on the N2a, $CBX2_{WT}$, $CBX2_{KRA}$, FUS^{IDR}, and $CBX2^{IDR}$ cell lines at 0hr and 24hr differentiation timepoints. Approximately 1.2 million cells were used per sample. Human cells, HEK293ET cells were spiked-in for each sample at 2% (24,000 cells). The method was based on the Henikoff protocol with the following alterations.[82] Concanavalin A beads (20 ul per sample) were prepped in binding buffer (1 mM $MnCl_2$, 1mM $CaCl_2$, 10 mM KCl, 20mM HEPES) with three washes over ~15 minutes and resuspended in 40 ul of binding buffer per sample. All cells were fixed with 1% formaldehyde for 10 minutes and fixing was quenched with 0.5M Tris (pH 8). After two PBS washes, the cells were resuspended in 100 ul of PBS per sample. To this was added 100 ul of wash buffer (20mM HEPES (pH 7.9),

150 mM NaCl, 2mM EDTA (pH 8), 0.5% BSA, and 0.1% Triton) and 40 ul of the prepped Con A beads. The cells were incubated with the beads on a rotator for 10 minutes at room temperature. A magnet was used to collect the beads and they were washed once with wash buffer before being resuspended in a primary antibody solution (150 ul per sample) and incubated overnight at 4°C. The following antibodies were used: H3K27me3 Cell Signaling Technology (C36B11), H2AK119Ub1 Cell Signaling Technology (D27C4), CBX2 custom-made serum (JK856)[83], RING1B Cell Signaling Technology (D22F2), GFP Invitrogen (A-11122), and Normal Rabbit IgG Cell Signaling Technology (#2729). Antibodies were diluted at 1:100 in wash buffer.

The next day samples were washed twice with wash buffer and each was resuspended in 50 ul of wash buffer containing 0.04 ug/ul concentration of pA-Mnase. Samples were incubated at room temperature for an hour, then washed three times with wash buffer. Samples were resuspended in 100 ul of wash buffer containing 10mM $CaCl_2$ and incubated on ice at 4°C for 2 hours. 100 ul of STOP Buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.1 ug/ul Rnase A) was added to each sample and they were incubated at 37°C for 15 minutes. The supernatant was removed to new tubes and 2 ul of 10% SDS and 2 ul of Proteinase K (19 mg/mL) were added to each sample. The samples were incubated overnight at 50°C in a hybridization oven.

The next day the samples were treated with a standard ethanol precipitation and resuspended in 30 ul of TE buffer. The NEBNext Ultra II DNA Library Prep Kit for Illumina was used to prep libraries. Adaptors and primers from the NEBNext Multiplex Oligos for Illumina (Dual Index Primer Set 1) were used and the libraries were sequenced on a Nextseq 2000.

CUT&RUN Sequencing Analysis

The quality of the sequenced libraries was again confirmed by an analysis with FastQC.[71] Trimmomatic[72] was then used to trim the first 5 bases from the 5' end of all sequenced reads and remove sequencing adapters. Mitochondrial reads were removed by aligning sample libraries to the mitochondrial sequence from the mouse genome mm39 with Bowtie2.[73] Furthermore, human spike-in reads were removed by alignment to the human genome hg38 and were not used for normalization in further analysis because of their fluctuation between replicates. The remaining reads were aligned to mm39. Reads with multiple alignments alongside with distant or poorly aligned pairs were discarded, and unaligned reads were analyzed with Kraken2[77] and the nt database to verify the absence of contaminants. Duplicate reads were marked by Picard's *MarkDuplicates*[84] function and excluded from further analysis. Peaks were called using the *findPeaks* function provided by Homer v4.11[85] using condition- and time-point-specific IgG controls as input and empirically determined parameters. Details are presented in the source code uploaded to the repository Zenodo (Table S1). Peak sets were merged for further analysis, first, by antibody, requiring peak regions to be called as peaks by Homer in multiple samples. Subsequently, the merged peak sets for CBX2, GFP, and RING1B were merged again into a single peak set to effectively compare the differential occupancy by the different proteins throughout conditions and time points. The occupancy tracks for the shown screenshots were generated from the merged replicates using the *bamCoverage* function from the deepTools v3.5.2 package[86] normalizing all samples to 1x coverage in non-peak (background) regions. For differential peak analysis, occupancy was determined with featureCounts, and the results were further processed in R. Peak presence was denoted as at least 2-fold enrichment over IgG with an FDR<0.05 as computed by edgeR's exact test using library-size normalization. Binding

heatmaps were generated with *computeMatrix* and *plotHeatmap* functions from deepTools.[86] A direct comparison of peak coverage with an exact test between the wildtype N2A line and other lines showed that CBX2 and Ring1b peaks were mainly lost in the $CBX2_{KRA}$ and FUS^{IDR} lines and peak gains represented only 3% of all peaks. We therefore sorted peaks by their occupancy in the FUS^{IDR} and $CBX2_{KRA}$ samples in a descending order. The bigWigs used for the binding heatmap convey the enrichment over IgG and were generated by *bigWigCompare* using the 'ratio' parameter.[86] Repositories for the raw and processed data is shown in Table S1.

Breast Cancer Transwell Invasion Assay

Corning Matrigel Invasion Chambers were thawed and rehydrated with serum-free RPMI media for several hours with 500ul in the well and 500 ul in the insert. The 4T1 cell lines were trypsinized and resuspended in serum-free RPMI at $1*10^5$ cells/mL concentration. The media was removed from the top of the inserts and the inserts were transferred to wells with 750 ul of serum-free RPMI (one triplicate of wildtype 4T1 cells as a control) or full RPMI media with 10% FBS. 500 ul of the 4T1 cell line solutions were seeded in the inserts in triplicated. The plates were incubated overnight for 22 hours.

The Tek-Select Diff Stain Kit from IMEB was used to stain the cells that invaded the matrigel. Media was removed from the top of the inserts and they were moved into a 24-well plate containing 1 mL of PBS. 500 ul of PBS was added to the top of the insert. A cotton swab was used to remove the cells from the top of the insert scraping gently. The media on the top of the insert was removed and a second cotton swab was used to remove residual media. Then, the insert was moved to a 24-well plate containing 1 mL of the fixative, Solution A, from the Diff

Stain Kit. The plate was incubated at RT on a shaker for 4 minutes. Solution A was removed from the top of the insert and the insert was moved to a new 24-well plate containing 1 mL of Solution B from the Diff Stain Kit. 100 ul of Solution B was added to the top of the insert and the plate was incubated at RT on a shaker for 4 minutes. Solution B was removed and the insert was moved to a well containing 1 mL of Solution C. 100 ul of Solution C was added to the top of the insert and the plate was incubated at RT on a shaker for 4 minutes. Solution C was removed and the insert was moved to a 24 well plate containing 1 mL of Milli-Q water. 500 ul of water was added to the top of the insert. After 5 minutes, the water was removed and the insert was moved to a new well with 1 mL of water for additional destaining. 500 ul of water was added to the top of the insert. After 5 minutes, the water was removed and the insert was dried upside down overnight. The next day, the inserts were cut out and mounted on glass slides for imaging using Vectashield PLUS mounting media. The slides were imaged on a Nikon Eclipse 90i using the bright field setting at 20x magnification. Cell counting was performed using FIJI.

References

1. Schuettengruber, B., Bourbon, H. M., Di Croce, L., and Cavalli, G. (2017). Genome regulation by Polycomb and Trithorax: 70 years and counting. *Cell, 171*(1), 34-57. https://doi.org/10.1016/j.cell.2017.08.002

2. Whitcomb, S. J., Basu, A., Allis, C. D., and Bernstein, E. (2007). Polycomb Group proteins: an evolutionary perspective. *Trends in Genetics*, 23, 494–502. https://doi.org/10.1016/j.tig.2007.08.006

3. Blackledge, N. P., and Klose, R. J. (2021). The molecular principles of gene regulation by Polycomb repressive complexes. *Nature Reviews Molecular Cellular Biology*, *22*(12), 815–833. https://doi.org/10.1038/s41580-021-00398-y

4. Loh, C. H. and Veenstra, G. J. C. (2022). The Role of Polycomb Proteins in Cell Lineage Commitment and Embryonic Development. *Epigenomes*, *6*(3):23. https://doi.org/10.3390/epigenomes6030023

5. Kim, J. J. and Kingston, R. E. (2022). Context-specific Polycomb mechanisms in development. *Nature Reviews Genetics*, *23*(11), 680–695. https://doi.org/10.1038/s41576-022-00499-0

6. Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., Wutz, A., Vidal, M., Elderkin, S., and Brockdorff, N. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at Polycomb target sites independently of PRC2 and H3K27me3. *Cell*, *148*(4), 664–678. https://doi.org/10.1016/j.cell.2011.12.029

7. Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature*, *431*(7010), 873–878. https://doi.org/10.1038/nature02985

8. Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature*, *469*(7330), 343–349. https://doi.org/10.1038/nature09784

9. Guo, Y., Zhao, S., and Wang, G. G. (2021). Polycomb Gene Silencing Mechanisms: PRC2 Chromatin Targeting, H3K27me3 'Readout', and Phase Separation-Based Compaction. *Trends in Genetics*, *37*(6), 547–565. https://doi.org/10.1016/j.tig.2020.12.006

10. Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J. T., Sadreyev, R. I., Dekker, J., and Kingston, R. E. (2017). Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Molecular Cell*, *65*(3), 432-446. https://doi.org/10.1016/j.molcel.2017.01.009

11. Parreno, V., Martinez, A. M., and Cavalli, G. (2022). Mechanisms of Polycomb group protein function in cancer. *Cell Research*, *32*, 231–253. https://doi.org/10.1038/s41422-021-00606-6

12. Jangal, M., Lebeau, B., and Witcher, M. (2019). Beyond EZH2: is the Polycomb protein CBX2 an emerging target for anti-cancer therapy? *Expert Opinion on Therapeutic Targets*, *23*(7), 565–578. https://doi.org/10.1080/14728222.2019.1627329

13. Zhou, H., Xiong, Y., Liu, Z., Hou, S., and Zhou, T. (2021). Expression and prognostic significance of CBX2 in colorectal cancer: database mining for CBX family members in malignancies and vitro analyses. *Cancer Cell International, 21,* 402. https://doi.org/10.1186/s12935-021-02106-4

14. Piqué, D. G., Montagna, C., Greally, J. M., and Mar, J. C. (2019). A novel approach to modelling transcriptional heterogeneity identifies the oncogene candidate CBX2 in invasive breast carcinoma. *British Journal of Cancer, 120,* 746–753. https://doi.org/10.1038/s41416-019-0387-8

15. Grau, D. J., Chapman, B. A., Garlick, J. D., Borowsky, M., Francis, N. J., and Kingston, R. E. (2011). Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes and Development*, *25*(20), 2210–2221. https://doi.org/10.1101/gad.17288211

16. Francis, N. J., Kingston, R. E., and Woodcock, C. L. (2004). Chromatin compaction by a polycomb group protein complex. *Science (New York, N.Y.)*, *306*(5701), 1574–1577. https://doi.org/10.1126/science.1100576

17. Guo, Y. and Wang, G. G. (2022). Modulation of the high-order chromatin structure by Polycomb complexes. *Frontiers in Cell and Developmental Biology*, *10*, 1021658. https://doi.org/10.3389/fcell.2022.1021658

18. Lau, M. S., Schwartz, M. G., Kundu, S., Savol, A. J., Wang, P., Marr, S. K., Grau, D. J., Schorderet, P., Sadreyex, R. I., Tabin, C. J., and Kingston, R. E. (2017). Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. *Science, 355*(6329), 1081-1084. https://doi.org/10.1126/science.aah5403

19. Plys, A., Davis, C. P., Kim, J., Rizki, G., Keenen, M. M., Marr, S. K., and Kingston, R. E. (2019). Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes and Development, 33*, 799-813. https://doi.org/10.1101/gad.326488.119

20. Tatavosian, R., Kent, S., Brown, K., Yao, T., Duc, H. N., Hunyh, T. N., Zhen, C. Y., Ma, B., Wang, H., and Ren, X. (2019). Nuclear condensates of the Polycomb protein chromobox2 (CBX2) assemble through phase separation. *Journal of Biological Chemistry, 294*, 1451-1463. https://doi.org/10.1074/jbc.RA118.006620

21. Saurin, A. J., Shiels, C., Williamson, J., Satijn, D. P., Otte, A. P., Sheer, D., and Freemont, P. S. (1998). The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. *Journal of Cell Biology*, *142*(4), 887-98. https://doi.org/10.1083/jcb.142.4.887

22. Pirrotta, V. and Li, H. B. (2011). A view of nuclear Polycomb bodies. *Current Opinion in Genetics and Development, 22*(2): 101-9. https://doi.org/10.1016/j.gde.2011.11.004

23. Hnisz, D., Shrinivas, K., Young R. A., Chakraborty, A. K., and Sharp, P. A. (2017). A phase separation model for transcriptional control. *Cell, 169*(1), 13-23. https://doi.org/10.1016/j.cell.2017.02.007

24. Sabari, B. R., Dall'Agnese, A., and Young, R. A. (2020). Biomolecular Condensates in the nucleus. *Trends in Biochemical Sciences, 45*(11), 961-977. https://doi.org/10.1016/j.tibs.2020.06.007

25. Boeynaems, S., Alberti, S., Fawzi, N. L., Mittag, T., Polymenidou, M., Rousseau, R., Schymkowitz, J., Shorter, J., Wolozin, B., Van Den Bosch, L., Tompa, P., and Fuxreiter, M. (2018). Protein Phase Separation: A New Phase in Cell Biology. *Trends in Cell Biology, 28*(6), 420-435. https://doi.org/10.1016/j.tcb.2018.02.004

26. Cho, W. K., Spille, J. H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I. I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science, 361*(6400), 412–415. https://doi.org/10.1126/science.aar4199

27. 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., Li, C. H., Guo, Y. E., Day, D. S., Schuijers, J., Vasile, E., Malik, S., Hnisz, D., Lee, T. I., Cisse, I. I., Roeder, R. G., Sharp, P. A., Chakraborty, A.K., and Young, R. A. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science*, *361*(6400), eaar3958. https://doi.org/10.1126/science.aar3958

28. Lyons, H., Veettil, R. T., Pradhan, P., Fornero, C., De La Cruz, N., Ito, K., Eppert, M., Roeder, R. G., and Sabari, B. R. (2023). Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell*, *186*(2), 327–345.e28. https://doi.org/10.1016/j.cell.2022.12.013

29. Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S., and Narlikar, G. J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature*, *547*(7662), 236–240. https://doi.org/10.1038/nature22822

30. Strom, A. R., Emelyanov, A. V., Mir, M., Fyodorov, D. V., Darzacq, X., and Karpen, G. H. (2017). Phase separation drives heterochromatin domain formation. *Nature*, *547*(7662), 241–245. https://doi.org/10.1038/nature22989

31. Keenen, M. M., Brown, D., Brennan, L. D., Renger, R., Khoo, H., Carlson, C. R., Huang, B., Grill, S. W., Narlikar, G. J., and Redding, S. (2021). HP1 proteins compact DNA into mechanically and positionally stable phase separated domains. *eLife, 10*, e64563. https://doi.org/10.7554/eLife.64563

32. Sanulli, S., Trnka, M. J., Dharmarajan, V., Tibble, R. W., Pascal, B. D., Burlingame, A. L., Griffin, P. R., Gross, J. D., and Narlikar, G. J. (2019). HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature, 575*, 390–394. https://doi.org/10.1038/s41586-019-1669-2

33. Tortora, M. M. C., Brennan, L. D., Karpen, G., and Jost, D. (2023). HP1-driven phase separation recapitulates the thermodynamics and kinetics of heterochromatin condensate formation. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(33), e2211855120. https://doi.org/10.1073/pnas.2211855120

34. Gibson, B. A., Doolittle, L. K., Schneider, M. W. G., Jensen, L. E., Gamarra, N., Henry, L., Gerlich, D. W., Redding, S., and Rosen, M. K. (2019). Organization of chromatin by intrinsic and regulated phase separation. *Cell, 179*(2), 470-484. https://doi.org/10.1016/j.cell.2019.08.037

35. Shakya, A., Park, S., Rana, N., and King, J. T. (2019). Liquid-liquid phase separation of histone proteins in cells: Role in chromatin organization. *Biophysical Journal, 118*(3), 753-764. https://doi.org/10.1016/j.bpj.2019.12.022

36. Zhang, M., Díaz-Celis, C., Onoa, B., Cañari-Chumpitaz, C., Requejo, K. I., Liu, J., Vien, M., Nogales, E., Ren, G., and Bustamante, C. (2022). Molecular organization of the early stages of nucleosome phase separation visualized by cryo-electron tomography. *Molecular Cell*, *82*(16), 3000–3014.e9. https://doi.org/10.1016/j.molcel.2022.06.032

37. Patil, A., Strom, A. R., Paulo, J. A., Collings, C. K., Ruff, K. M., Shinn, M. K., Sankar, A., Cervantes, K. S., Wauer, T., St Laurent, J. D., Xu, G., Becker, L. A., Gygi, S. P., Pappu, R. V., Brangwynne, C. P., and Kadoch, C. (2023). A disordered region controls cBAF activity via condensation and partner recruitment. *Cell*, *186*(22), 4936–4955.e26. https://doi.org/10.1016/j.cell.2023.08.032

38. Brown, K., Chew, P. Y., Ingersoll, S., Espinosa, J. R., Aguirre, A., Espinoza, A., Wen, J., Astatike, K., Kutateladze, T. G., Collepardo-Guevara, R., and Ren, X. (2023). Principles of assembly and regulation of condensates of Polycomb repressive complex 1 through phase separation. *Cell Reports*, *42*(10), 113136. https://doi.org/10.1016/j.celrep.2023.113136

39. Eeftens, J. M., Kapoor, M., Michieletto, D., and Brangwynne, C. P. (2021). Polycomb condensates can promote epigenetic marks but are not required for sustained chromatin compaction. *Nature Communications*, *12*(1), 5888. https://doi.org/10.1038/s41467-021-26147-5

40. Williamson, I., Boyle, S., Grimes, G. R., Friman, E. T., and Bickmore, W. A. (2023). Dispersal of PRC1 condensates disrupts Polycomb chromatin domains and loops. *Life Science Alliance*, *6*(10), e202302101. https://doi.org/10.26508/lsa.202302101

41. Shin, Y., Berry, J., Pannucci, N., Haataja, M. P., Toettcher, J. E., and Brangwynne, C. P. (2017). Spatiotemporal control of intracellular phase transitions using light-activated optodroplets. *Cell, 168*(1-2), 159-171. https://doi.org/10.1016/j.cell.2016.11.054

42. Schwartz, J. C., Ebmeier, C. C., Podell, E. R., Heimiller, J., Taatjes, D. J., and Cech, T. R. (2012). FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. *Genes and Development, 26*(24), 2690–2695. https://doi.org/10.1101/gad.204602.112

43. Rogelj, B., Easton, L., Bogu, G. K., Stanton, L. W., Rot, G., Curk, T., Zupan, B., Sugimoto, Y., Modic, M., Haberman, N., Tollervey, J., Fujii, R., Takumi, T., Shaw, C. E., and Ule J. (2012). Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Science Reports, 2*, 603. https://doi.org/10.1038/srep00603

44. Levone, B. R., Lenzken, S. C., Antonaci, M., Maiser, A., Rapp, A., Conte, F., Reber, S., Mechtersheimer, J., Ronchi, A. E., Mühlemann, O., Leonhardt, H., Cardoso, M. C., Ruepp, M. D., and Barabino, S. M. L. (2021). FUS-dependent liquid-liquid phase separation is important for DNA repair initiation. *Journal of Cell Biology*, *220*(5), e202008030. https://doi.org/10.1083/jcb.202008030

45. Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M. Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T. M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L. A., Weigert, M., Myers, E. W., Grill, S., Drechsel, D., Hyman, A. A., and Alberti, S. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell, 162*(5), 1066–1077. https://doi.org/10.1016/j.cell.2015.07.047

46. Clarke, J. P., Thibault, P. A., Salapa, H. E., and Levin, M. C. (2021). A Comprehensive Analysis of the Role of hnRNP A1 Function and Dysfunction in the Pathogenesis of Neurodegenerative Disease. *Frontiers in Molecular Biosciences*, *8*, 659610. https://doi.org/10.3389/fmolb.2021.659610

47. Richter, W. F., Nayak, S., Iwasa, J., and Taatjes, D. J. (2022). The Mediator complex as a master regulator of transcription by RNA polymerase II. *Nature Reviews Molecular Cell Biology*, *23*(11), 732–749. https://doi.org/10.1038/s41580-022-00498-3

48. Janssen, A. F. J., Katrukha, E.A., van Straaten, W., Verlhac, P., Reggiori, F., and Kapitein, L.C. (2018). Probing aggrephagy using chemically-induced protein aggregates. *Nature Communications, 9*(1), 4245. https://doi.org/10.1038/s41467-018-06674-4

49. Jaensch, E. S., Zhu, J., Cochrane, J. C., Marr, S. K., Oei, T. A., Damle, M., McCaslin, E. Z., and Kingston, R. E. (2021). A Polycomb domain found in committed cells impairs differentiation when introduced into PRC1 in pluripotent cells. *Molecular Cell, 81*(22), 4677–4691.e8. https://doi.org/10.1016/j.molcel.2021.09.018

50. Zhen, C. Y., Duc, H. N., Kokotovic, M., Phiel, C. J., and Ren, X. (2014). Cbx2 stably associates with mitotic chromosomes via a PRC2- or PRC1-independent mechanism and is

needed for recruiting PRC1 complex to mitotic chromosomes. *Molecular Biology of the Cell*, *25*(23), 3726–3739. https://doi.org/10.1091/mbc.E14-06-1109

51. Kent, S., Brown, K., Yang, C. H., Alsaihati, N., Tian, C., Wang, H., and Ren, X. (2020). Phase-Separated Transcriptional Condensates Accelerate Target-Search Process Revealed by Live-Cell Single-Molecule Imaging. *Cell Reports*, *33*(2), 108248. https://doi.org/10.1016/j.celrep.2020.108248

52. Brown, K., Andrianakos, H., Ingersoll, S., and Ren, X. (2021). Single-molecule imaging of epigenetic complexes in living cells: insights from studies on Polycomb group proteins, *Nucleic Acids Research, 49*(12), 6621–6637. https://doi.org/10.1093/nar/gkab304

53. Huseyin, M.K. and Klose, R.J. (2021). Live-cell single particle tracking of PRC1 reveals a highly dynamic system with low target site occupancy. *Nature Communications, 12*(1), 887. https://doi.org/10.1038/s41467-021-21130-6

54. Doyle, E. J., Morey, L., and Conway, E. (2022). Know when to fold 'em: Polycomb complexes in oncogenic 3D genome regulation. *Frontiers in Cell and Developmental Biology*, *10*, 986319. https://doi.org/10.3389/fcell.2022.986319

55. Gu, X., Wang, X., Su, D., Su, X., Lin, L., Li, S., Wu, Q., Liu, S., Zhang, P., Zhu, X., and Jiang, X. (2018). CBX2 inhibits neurite development by regulating neuron-specific genes expression. *Frontiers in Molecular Neuroscience*, *11*, 46. https://doi.org/10.3389/fnmol.2018.00046

56. Tsherniak, A., Vazquez, F., Montgomery, P. G., Weir, B.A., Kryukov, G., Cowley, G. S., Gill, S., Harrington, W. F., Pantel, S., Krill-Burger, J. M., Meyers, R. M., Ali, L., Goodale, A., Lee, Y., Jiang, G., Hsiao, J., Gerath, W. F. J., Howell, S., Merkel, E., Ghandi, M., Garraway, L. A., Root, D. E., Golub, T. R., Boehm, J. S., and Hahn, W. C. (2017). Defining a cancer dependency map. *Cell, 170*(3), 564-576. https://doi.org/10.1016/j.cell.2017.06.010

57. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics, 25*(1), 25–29. https://doi.org/10.1038/75556

58. Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., and Lewis, S. (2009). AmiGO Hub, web presence working group. AmiGO: online access to ontology and annotation data. *Bioinformatics, 25*(2), 288-289. https://doi.org/10.1093/bioinformatics/btn615

59. Blake, J., Bug, W., Chisholm, R., Clark, J., Feltrin, E., Finger, J., Hill, D., Harris, M., Hayamizu, T., Howe, D., Martone, M., Millen, K., and Sele, F. (2006). Improving the representation of contral nervous system development in the biological process ontology. *AmiGO2.* Available at https://amigo.geneontology.org/amigo/term/GO:0022008.

60. Schmidt, H., Stonkute, A., Jüttner, R., Schäffer, S., Buttgereit, J., Feil, R., Hofmann, F., and Rathjen, F. G. (2007). The receptor guanylyl cyclase Npr2 is essential for sensory axon bifurcation within the spinal cord. *Journal of Cell Biology*, *179*(2), 331–340. https://doi.org/10.1083/jcb.200707176

61. Schmidt, H., Dickey, D. M., Dumoulin, A., Octave, M., Robinson, J. W., Kühn, R., Feil, R., Potter, L. R., and Rathjen, F. G. (2018). Regulation of the Natriuretic Peptide Receptor 2 (Npr2) by Phosphorylation of Juxtamembrane Serine and Threonine Residues Is Essential for Bifurcation of Sensory Axons. *Journal of Neuroscience*, *38*(45), 9768–9780. https://doi.org/10.1523/JNEUROSCI.0495-18.2018

62. Hou, P. S., Chuang, C. Y., Kao, C. F., Chou, S. J., Stone, L., Ho, H. N., Chien, C. L., and Kuo, H. C. (2013). LHX2 regulates the neural differentiation of human embryonic stem cells via transcriptional modulation of PAX6 and CER1. *Nucleic Acids Research*, *41*(16), 7753–7770. https://doi.org/10.1093/nar/gkt567

63. Bilton, L. J., Warren, C., Humphries, R. M., Kalsi, S., Waters, E., Francis, T., Dobrowinski, W., Beltran-Alvarez, P., and Wade, M. A. (2022). The epigenetic regulatory protein CBX2 promotes mTORC1 signalling and inhibits DREAM complex activity to drive breast cancer cell growth. *Cancers*, *14*(14), 3491. https://doi.org/10.3390/cancers14143491

64. Zheng, S., Lv, P., Su, J., Miao, K., Xu, H., and Li, M. (2019). Overexpression of CBX2 in breast cancer promotes tumor progression through the PI3K/AKT signaling pathway. *American Journal of Translational Research*, *11*(3), 1668–1682. https://pubmed.ncbi.nlm.nih.gov/30972192

65. Cermakova, K. and Hodges, H. C. (2023). Interaction modules that impart specificity to disordered protein. *Trends in Biochemical Sciences*, *48*(5), 477–490. https://doi.org/10.1016/j.tibs.2023.01.004

66. Monahan, Z., Ryan, V. H., Janke, A. M., Burke, K. A., Rhoads, S. N., Zerze, G. H., O'Meally, R., Dignon, G. L., Conicella, A. E., Zheng, W., Best, R. B., Cole, R. N., Mittal, J., Shewmaker, F., and Fawzi, N. L. (2017). Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *The EMBO journal*, *36*(20), 2951–2967. https://doi.org/10.15252/embj.201696394

67. Seif, E., Kang, J. J., Sasseville, C., Senkovich, O., Kaltashov, A., Boulier, E. L., Kapur, I., Kim, C. A., and Francis, N. J. (2020). Phase separation by the polyhomeotic sterile alpha motif compartmentalizes Polycomb Group proteins and enhances their activity. *Nature Communications*, *11*(1), 5609. https://doi.org/10.1038/s41467-020-19435-z

68. Kapur, I., Boulier, E. L., and Francis, N. J. (2022). Regulation of polyhomeotic condensates by intrinsically disordered sequences that affect chromatin binding. *Epigenomes*, *6*(4), 40. https://doi.org/10.3390/epigenomes6040040

69. Boettiger, A. N., Bintu, B., Moffitt, J. R., Wang, S., Beliveau, B. J., Fudenberg, G., Imakaev, M., Mirny, L. A., Wu, C. T., and Zhuang, X. (2016). Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature, 529*(7586), 418–422. https://doi.org/10.1038/nature16496

70. Meijering, E., Jacob, M., Sarria, J. C. F., Steiner, P., Hirling, H., and Unser M. (2004). Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. *Cytometry, 58*(2), 167-176. https://doi.org/10.1002/cyto.a.20022

71. Andrews, S. (2010). FastQC: A quality control tool for high throughput sequencing data (v0.11.9). *Babraham Bioinformatics*, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

72. Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

73. Langmead, B., Wilks, C., Antonescu, V., and Charles, R. (2019). Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics, 35(3)*, 421-432. https://doi.org/10.1093/bioinformatics/bty648

74. Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, P., Skripchenko, Y., Tse, T., Wang, J., Williams, R., Trawick, B. W., Pruitt, K. D., and Sherry, S. T. (2022). Database resources of the national center for biotechnology information. *Nucleic Acids Research, 50(D1)*, D20-D26. https://doi.org/10.1093/nar/gkab1112

75. van de Geijn, B., McVicker, G., Gilad, Y., and Pritchard, J. K. (2015). WASP: allele-specific software for robust molecular quantitative trait locus discovery. *Nature methods, 12*(11), 1061-1063. https://doi.org/10.1038/nmeth.3582

76. O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C. M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V. S., Kodali, V. K., Li, W., Maglott, D., Masterson, P., McGarvey, K. M., Murphy, M. R., O'Neill, K., Pujar, S., Rangwala, S. H., Rausch, D., Riddick, L. D., Schoch, C., Shkeda, A., Storz, S. S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R. E., Vatsan, A. R., Wallin, C., Webb, D., Wu, W., Landrum, M. J., Kimchi, A., Tatusova, T., DiCuccio, M., Kitts, P., Murphy, T. D., and Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic acids research*, *44*(D1), D733–D745. https://doi.org/10.1093/nar/gkv1189

77. Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome biology, 20*(1), 257. https://doi.org/10.1186/s13059-019-1891-0

78. Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930. https://doi.org/10.1093/bioinformatics/btt656

79. Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, *Bioinformatics, 26(*1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

80. Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., and Smyth, G. K. (2015). *limma* powers differential expression analyses for RNA-sequencing and microarray studies, *Nucleic Acids Research, 43*(7), e47. https://doi.org/10.1093/nar/gkv007

81. Plotly Technologies Inc. (2015). Collaborative data science (v4.10.3). *Plotly Technologies Inc.*, https://plot.ly/

82. Skene, P. J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife*, *6*, e21856. https://doi.org/10.7554/eLife.21856

83. Kim, J. J., Steinson, E. R., Lau, M. S., de Rooij, D. G., Page, D. C., and Kingston, R. E. (2023). Cell type-specific role of CBX2 and its disordered region in spermatogenesis. *Genes and development*, *37*(13-14), 640–660. https://doi.org/10.1101/gad.350393.122

84. Broad Institue. Picard Tools. Available at https://broadinstitute.github.io/picard/

85. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell*, *38*(4), 576–589. https://doi.org/10.1016/j.molcel.2010.05.004

86. Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research*, *44*(W1), W160–W165. https://doi.org/10.1093/nar/gkw257

Chapter 4 : Conclusions and future directions

4.1 CBX2, CBX4, and CBX8 drive phase separation of Polycomb Repressive Complex 1

The overarching conclusion from this dissertation is the importance of phase separation by CBX2, CBX4, and CBX8 in cPRC1 function. While CBX2 had been previously studied, the phase separation activity of CBX4 and CBX8 in addition to CBX2 demonstrates the importance of condensate formation as a general mechanism for this family of paralogs.[1,2] CBX4 and CBX8 form droplets *in vitro* and condensates in cells. FRAP dynamics demonstrate the liquid-like properties of CBX4 and CBX8 condensates within the nucleus. Polycomb bodies containing CBX4 and CBX8 in NIH-3T3 cells are highly dynamic. It is proposed that these dynamics are part of a system where the diffusion of cPRC1 into and out of Polycomb bodies maintains repression at target genes. CBX4 and CBX8 are also capable of driving Polycomb condensate formation in synthetic systems. The OptoDroplet experiments discussed in Chapter 2 suggest self-interactions facilitated by CRY2 oligomerization drove condensate formation for CBX4 and CBX8 as well as CBX2. Mutation of positive charges in the IDRs of CBX4 and CBX8 attenuate their ability to form condensates. Therefore, the IDRs of CBX4 and CBX8 act in a similar manner to the IDR of CBX2 to facilitate Polycomb body formation in cells.

We propose that the differences in amino acid composition and liquid-like dynamics amongst CBX2, CBX4, and CBX8 may have evolved as a mechanism to tune Polycomb body formation or stability. The exact dynamics of the Polycomb bodies may be altered by the ratio of different CBX paralogs in these condensates. CBX2, CBX4, and CBX8 seem to have slightly different phase separation behavior although more studies are needed to differentiate their dynamics in cells. It seems the ability to tune phase separation through the inclusion of a different paralog in the complex provides a potential mechanism of modulating gene repression (**Fig. 4.1**). Single

molecule studies tagging the endogenous CBX2, CBX4, and CBX8 proteins in differentiated cells will give us a deeper understanding of their respective dynamics within heterogeneous Polycomb bodies. Experiments to quantify the differences in dynamics of Polycomb bodies that contain multiple CBX proteins in different quantities may also help to determine the synergistic or prohibitive effect each CBX might have on condensate formation or stability.

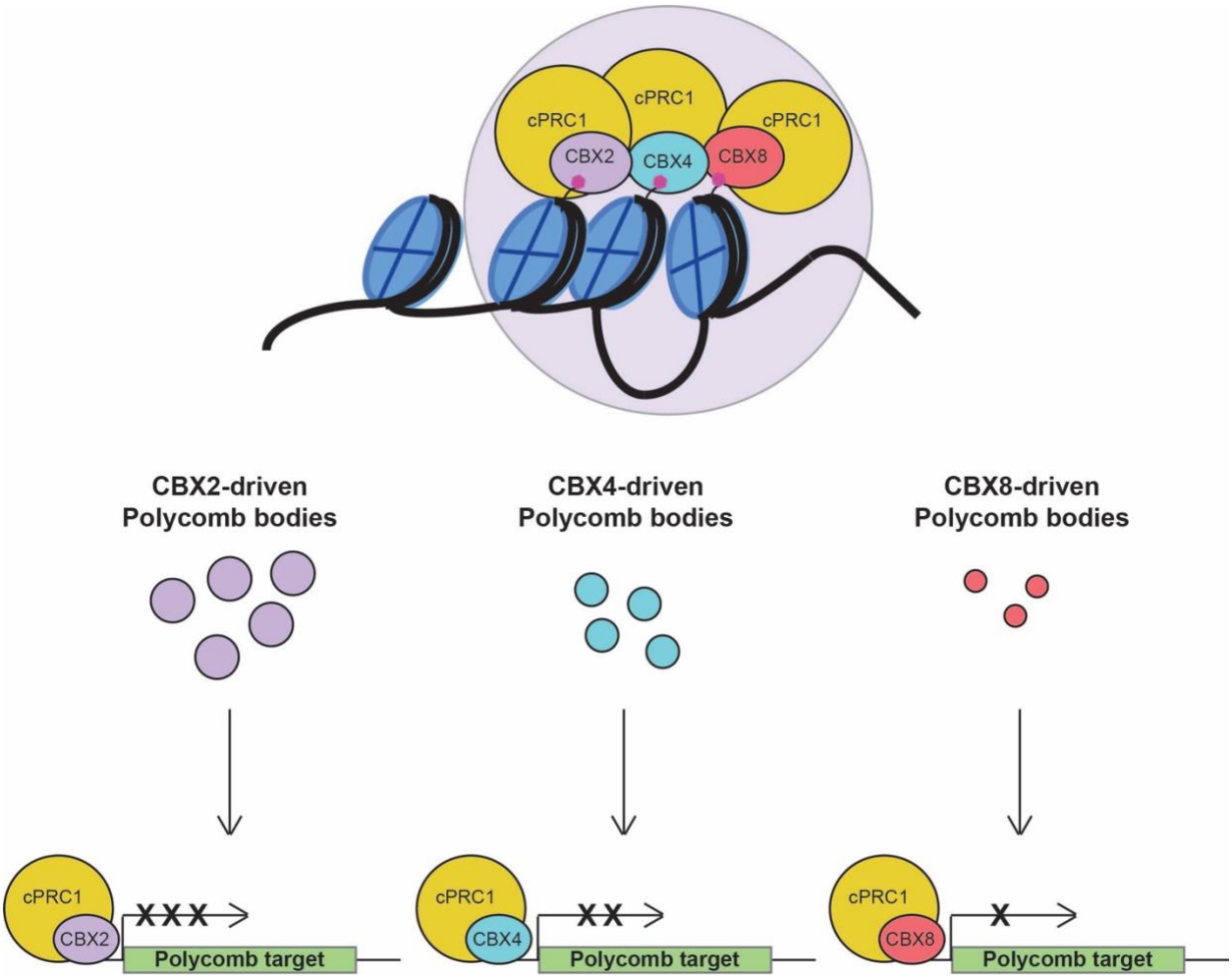


Figure 4.1 Model for CBX phase separation

CBX2, CBX4, and CBX8 condensate dynamics may tune repression of Polycomb target genes.

The expression of CBX paralogs with phase separation activity is an important inflection point in development. The main transition during mouse embryonic stem cell (mESC) differentiation occurs between inclusion of CBX7, a non-phase separating paralog, in cPRC1, to expression and inclusion of CBX2, CBX4, and CBX8 in cPRC1.[3] In mESCs, the inclusion of a non-condensate forming CBX protein is proposed to be important for maintaining bivalent chromatin.[4] Bivalent chromatin includes both activating and repressive histone marks and exists in a poised state ready for either expression or silencing during subsequent differentiation.[5] Therefore, a non-phase separating CBX protein might be a feature of bivalent chromatin that enables greater plasticity.[4] When a CBX7 mutant containing the IDR of CBX2 is expressed in mESCs, differentiation programs cannot be properly activated.[4] As development progresses, lineage genes are expressed while pluripotency and non-lineage genes are turned down to maintain cell identity.[6] Meanwhile, CBX2, CBX4, and CBX8 increase in expression to different levels based on cell type with both overlapping and distinct target genes.[3,7,8] This suggests that CBX2, CBX4, and CBX8 driven PRC1 condensate formation is important for the maintenance of a differentiated cell identity.

The tuning of gene repression by Polycomb body composition may explain the variable levels of each CBX in distinct cell lineages. The ratio of CBX2, CBX4, and CBX8 may create a more complex system of Polycomb repression via condensate behavior. Although some compensation has been observed between paralogs, there are also lineages that are dependent on one paralog over the others.[3,7,9] CBX2, CBX4, and CBX8 are often found at similar genes throughout the genome as determined by genome-wide binding analysis such as ChIP-seq.[9,10] We propose that phase separation properties of these paralogs might add another layer of gene regulation. Indeed,

previous work from our lab and others has shown that the composition of Polycomb bodies is important for the regulation of their properties.[11,12] Inclusion of CBX2 or CBX7 in cPRC1 influences the threshold for condensation of nucleosomal arrays and purified protein, while inclusion of PHC1 or PHC2 influences the gel-like or liquid-like properties of the condensates by way of their polymerization activity.[11] The composition of Polycomb bodies also influences the enrichment or exclusion of other PRC1 subunits.[12] For instance, CBX2 condensates can concentrate CBX4, CBX7 and CBX8 containing PRC1 in Polycomb bodies, while excluding CBX6 containing PRC1.[12] Additionally, phase separation may provide a non-stoichiometric means of concentrating Polycomb proteins. Single molecule studies recently demonstrated that only 3 CBX2 molecules are necessary to concentrate ~70 RING1B molecules in ES cells suggesting not all PcG protein located at a target gene are within a PRC1 complex.[13] Additionally, they show that the number of CBX2 molecules per condensate is cell-type dependent.[13] When the mESCs were differentiated to NPCs, there was an increase to ~15 CBX2 molecules per condensate.[13] However, other data has shown that the concentration of condensates is at least somewhat controlled by stoichiometry amongst the different subunits.[12] CBX2 condensates are enriched with PHC1 until a high concentration of PHC1 actually begins to dissolve the CBX2 driven condensates.[12] This suggests that the ratio of PRC1 subunits within Polycomb bodies can influence the stability of the condensates formed. Further studies will be needed to fully understand the composition of Polycomb bodies *in vivo* including the ratio of CBX proteins and the stoichiometry of cPRC1 components for different cell types.

Besides regulating condensate composition, the IDRs of the CBX proteins may also regulate cPRC1 binding to the genome. While the chromodomain of the CBX proteins has traditionally

thought to be the main mechanism of chromatin binding for PRC1, it is known that Polycomb condensates are not dependent on H3K27me3 and the deposition of H3K27me3 is not sufficient for gene repression.[2,14-16] The chromodomain of the CBX proteins certainly plays a part in targeting to the genome. For example, a chromodomain mutant of CBX2 shows a decrease in condensate overlap with H3K27me3 compared to wildtype CBX2.[17] Additionally, the AT-Hook region of the CBX2 protein can bind to DNA.[18] CBX4 and CBX8 have AT-hook like motifs that may similarly facilitate DNA binding.[18] Therefore, chromatin binding of PRC1 is almost certainly regulated by domains beyond the chromodomain. The charged amino acids in the IDRs of CBX2, CBX4, and CBX8 are important for phase separation activity and may also contribute to DNA binding.[8] As shown in Chapter 2, mutation of lysine and arginine residues in the IDRs of CBX2, CBX4, and CBX8 disrupts the ability of these proteins to phase separate *in vitro* and form nuclear foci in cells.[1] In Chapter 3, the overexpression of the $CBX2_{KRA}$ mutant led to a global decrease in the binding of cPRC1 to CBX2 genomic targets. Binding was restored when the mutant was fused to the positively charged $CBX2^{IDR}$, but was unchanged by fusion to the uncharged and hydrophobic FUS^{IDR}. The importance of acidic amino acids is unsurprising given that the charges likely enable the protein to interact more strongly with the negative DNA backbone as is the case with HP1.[19] Therefore, we hypothesize that the positive charges in the IDRs of the CBX paralogs may play a dual role in the multivalent interactions required for phase separation and the interactions required for occupancy at genomic targets. Work from the Ren lab has shown that phase separation may contribute to CBX2's target search for its binding sites.[20] The formation of condensates is thought to increase the local concentration of CBX2 and its chromatin occupancy and therefore speed up CBX2's target search.[20] They found that while H3K27me3 was not necessary for the binding of CBX2 to the genome, the mark did facilitate

CBX2's target search.[20] Once CBX2 binds to its genomic targets, it is thought to recruit the rest of the PRC1 complex.[20] The positive charges in the IDRs of the CBX proteins may play an important role in maintaining genomic occupancy at target genes even while they facilitate the concentrating and recruitment of Polycomb complexes via condensate formation.

4.2 Model for Polycomb body phase separation

Given the initiation of cPRC1 localization to the genome by CBX2, it has been proposed that condensation occurs through a client-scaffold mechanism.[2,12] Client-scaffold models of phase separation have been described for other nuclear bodies.[21] For example, P bodies have RNA molecules that act as scaffolds.[21] Scaffold interactions are typically stronger and drive the formation of the condensate.[21] Scaffolds then recruit client molecules.[21] These interactions are more transient and the localization of client molecules is dependent on the presence of the scaffold molecules.[21] In the case of Polycomb bodies, CBX2 has been proposed to act as a scaffold binding to genomic target genes and then recruiting and concentrating other PcG proteins as clients.[12] The heterotypic multivalent interactions of CBX2 with other PRC1 molecules or nucleic acids are more transient and these molecules are thought to depend on CBX2 for assembly to Polycomb bodies.[12] Supporting data for this theory includes knockout of other PRC1 subunits demonstrating that CBX2 will still form condensates even when it is not part of a complex.[12] Additionally, *in vitro* experiments demonstrate that CBX2 has the strongest phase separation activity compared to the other CBX paralogs or the other cPRC1 subunits.[12] While CBX2 drives the formation of the most stable condensates, this does not necessarily mean it is the only scaffold for Polycomb body formation. It is possible that the family of CBX proteins may be capable of acting as scaffolds, or perhaps the general scaffold-client model is too

simplistic to explain the dynamics of Polycomb body formation. Previous studies have shown that the CBX proteins may have a weaker or stronger K_d depending on the cellular context.[22,23] Therefore, different CBXs may act as the main scaffold in different cell lineages. Another possibility is that the multivalent interactions amongst PRC1 subunits may provide additional specificity to seed condensate formation.

The mechanism of client recruitment or partitioning of components into CBX2 condensates is mediated by the structured C box domain, which binds to RING1B, and by the multivalent interactions of CBX2's disordered region: both heterotypic interactions with other cPRC1 subunits or nucleic acids, and homotypic interactions with other CBX2 proteins.[24,12] It may be that CBX2 recruits the rest of the cPRC1 complex together as a heterotrimeric unit made up of the PHC, PCGF, and RING1 proteins by direct binding to the RING1 subunit.[2] Crystal structures reveal a RAWUL ubiquitin-like binding domain of RING1B that binds to the CBX protein to generate the heterodimer.[24] The N-terminus of RING1B has a ring finger domain that associates with the ring finger domain of PCGF4, also known as BMI1.[25] The PCGF subunit interacts with the PHC subunit to complete the core complex.[26] Interestingly, some single molecule tracking studies have shown that the four main cPRC1 subunits are not always present in Polycomb bodies in 1:1:1:1 stoichiometry.[13] As described above, in mESCs where CBX2 is lowly expressed, three CBX2 molecules are capable of recruiting ~70 RING1B molecules.[13] However, CBX7 is the most highly expressed CBX in mESCs[4], so it is possible that the "extra" RING1B molecules are actually in complex with CBX7. The stoichiometry of cPRC1 within Polycomb bodies *in vivo* requires additional study.

Multiple CBX paralogs are often present in the cell and determining their functional contribution to Polycomb-mediated gene repression will be important. A recent preprint has shown that in diffuse midline glioma (DMG), CBX4 has lower expression than CBX2 or CBX8 and only makes up ~5% of the cPRC1 complexes in the cell as demonstrated by IP-mass spec for RING1B.[27] However, CBX4 functionally drives the oncogenic repression that is important for DMG tumor progression and is more highly bound at Polycomb target genes.[27] Therefore, the abundance of CBX paralogs within a cell or the ratio of their inclusion in cPRC1 may not be predictive of functional contributions to Polycomb-mediated gene silencing. The factors controlling CBX paralog inclusion in cPRC1 are not clear. Perhaps the binding of CBX to RING1B via the C box domain is competitive amongst the paralogs, or perhaps different post-translational modifications in distinct cell types dictate protein stability or affinity for the complex. Regardless, the functional contribution of a CBX protein to gene repression seems to be independent of paralog inclusion in cPRC1.[27] This raises the question of whether the other ~95% of cPRC1 complexes are included in Polycomb bodies in DMG.[27] It would be interesting to study Polycomb bodies in DMG or other cancers with a known CBX4 dependency to determine if functional repression by CBX4 is reflected in Polycomb body composition. Even if CBX4 makes up a minority of cPRC1 in DMG cells, non-stoichiometric CBX4 proteins may be included in Polycomb bodies to tune gene repression or CBX4 might have the highest affinity for Polycomb targets. Determining the mechanism of repression when multiple CBX paralogs are present in a cell type will entail a deeper understanding of the relationship between Polycomb body composition, the ratio of CBX proteins present at target genes, and subsequent levels of transcription.

4.3 Mechanism of Polycomb-mediated gene repression by phase separation

The work from Chapter 3 suggests that phase separation itself is important for the functional repression of genes by Polycomb complexes. While the FUS^{IDR} cannot recover the binding of cPRC1 mediated by the $CBX2^{IDR}$, it nearly completely rescues repression of neurogenesis genes as demonstrated by the phenotypic rescue and RNA-sequencing data. Highlighting RING1B targets in a volcano plot of the Differentially Expressed Genes (DEGs) between the $CBX2_{KRA}$ and FUS^{IDR} lines demonstrates that the inclusion of a condensate forming CBX2, even if phase separation activity is conferred by a heterologous IDR, makes cPRC1 more repressive at target genes. This suggests that phase separation activity, independently of the specific sequence of the $CBX2^{IDR}$, is an important part of Polycomb-mediated gene repression. However, the mechanism of repression by phase separation is still unclear.

Since the positive charges in the $CBX2^{IDR}$ facilitate both compaction and phase separation, it has been proposed that the presence of CBX condensates is inextricable linked to nucleosome compaction.[1,28,29] The lysine and arginine residues that drive CBX2 phase separation also drive compaction of nucleosomal arrays *in vitro* as demonstrated by electron microscopy and Restriction Enzyme Accessibility assays, where nucleosome compaction prevents the cutting of DNA bound to nucleosomal arrays.[1,28] Driving synthetic CBX2 phase separation with the corelet system, results in chromatin compaction in cells as represented by increased variance in fluorescent signal of labelled H2B nucleosomes.[17] Treatment of cells with hexanediol, a compound known to disrupt the multivalent interactions necessary for phase separation, disperses Polycomb condensates in cells.[30] Simultaneously, the HoxD locus, a known Polycomb target, is decompacted as the distance between Fluorescence In Situ Hybridization (FISH) probes

bound to the DNA locus increases significantly.[30] This decompaction is reversible and can be recovered by the removal of hexanediol.[30] Altogether, these data support the ability of Polycomb condensate formation to drive local compaction of chromatin as a potential model for gene repression (**Fig. 4.2**).

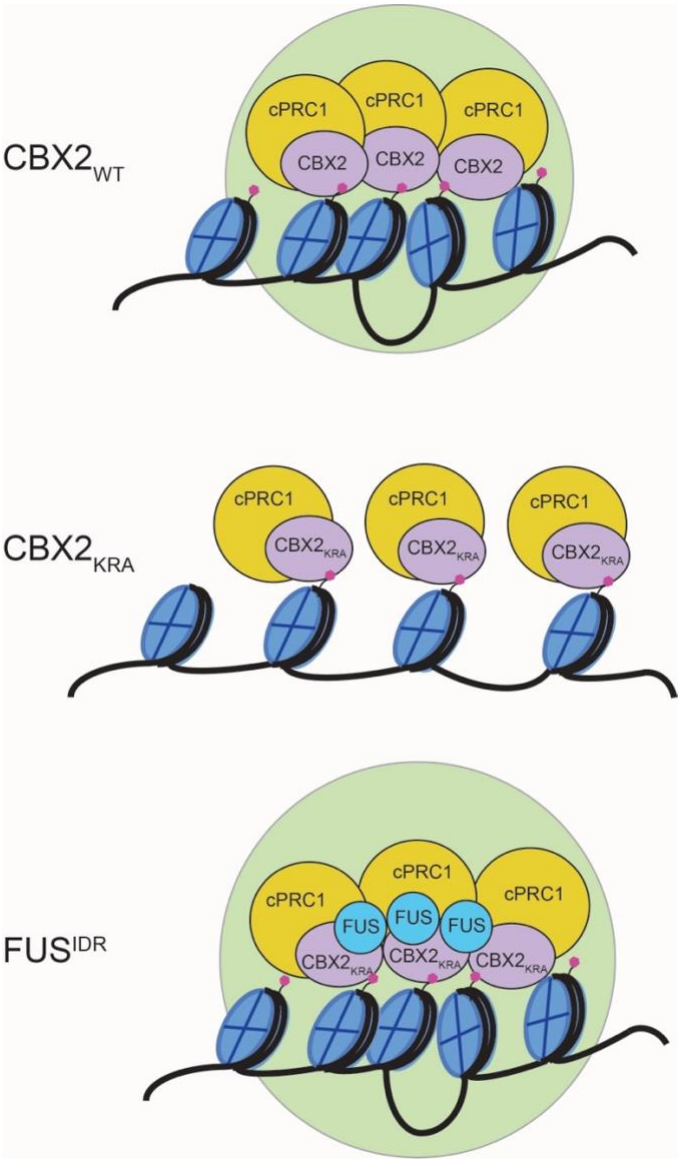


Figure 4.2 Model for the FUS^{IDR} rescue of the CBX2$_{KRA}$ mutant

Heterologous phase separation by the FUS^{IDR} rescues phase separation and compaction of CBX2$_{KRA}$ to restore gene repression.

This model is not without controversy. Some argue that compaction itself rather than phase separation is responsible for driving gene repression. Work from the Brangwynne lab has demonstrated that Polycomb phase separation is important for establishing repressive histone marks, but the formation of condensates is dispensable for sustained compaction and repression.[17] In this case, gene repression is attributed to the compacted state rather than the formation of membraneless compartments in the nucleus.[17] In contrast, others argue that neither compaction nor phase separation is primarily responsible for gene repression by Polycomb. Recent work from Alistair Boettiger's group argued that repression is generated by spatial feedback, transient interactions that propagate memory of the repressed state.[31] They use ORCA, Optical Reconstruction of Chromatin Architecture, to record the behavior of the HoxA cluster.[31] Dynamic simulations show intermixing of the HoxA locus with both open and closed chromatin even while the locus is uniformly repressed.[31] They interpret this as a refutation of a compacted globule compartmentalized by liquid liquid phase separation and propose a spatial feedback mechanism as a model for repression.[31]

In order to determine the impact of phase separation on compaction, one could measure compaction *in vivo* by changes in FISH probe distances to see the changes in locus compaction at genes repressed by heterologous phase separation in the FUS_{IDR} line such as *Lhx2* or *Npr2*. In a more controlled experiment, the heterologous IDR fusion constructs could be recruited directly to a target locus by fusion to dCas9. Then, FISH could be performed to measure compaction and qPCR could be used to measure the transcriptional output from the locus. This experimental design could also be used to compare the CBX paralogs, which display different levels of compaction *in vitro*.[28] One caveat is that it may be challenging to discern if compaction is a

primary effect of phase separation or a secondary effect of transcriptional changes at the locus. Therefore, *in vitro* studies of the purified FUS^{IDR} construct compared to the $CBX2_{KRA}$ construct when incubated with nucleosomal arrays could contribute to understanding the relationship between phase separation and compaction. If compaction by phase separation is the main mechanism of gene repression, it would be hypothesized that the FUS^{IDR} could induce nucleosomal array compaction and prevent cutting in the REA assay. Although *in vitro* studies could be an important first step in establishing this mechanism, validation in cells will be required at Polycomb target genes.

Chromatin compaction is one feature of higher order genomic architecture within the cell that provides another layer of transcriptional regulation.[32,33] Changes in 3D genome structure occur during development corresponding with changes in transcriptional states regulated, at least in part, by Polycomb group complexes.[34,35] Generally, chromatin can be divided into open and transcribed euchromatin and compacted, silent heterochromatin.[36,37] Euchromatin, also known as Compartment A, is gene rich and accessible to transcription factors, whereas heterochromatin, or Compartment B, is inaccessible and densely packaged.[36,37,38,39] Heterochromatin can be divided into two types: constitutive and facultative.[39] Constitutive heterochromatin is guarded by proteins such as HP1 in stable silent domains, while Polycomb proteins are enriched in facultative heterochromatin that represents an in-between state with the "facultas," Latin for "opportunity", to become open or closed depending on the context.[39] Facultative heterochromatin includes genes that are temporally constricted in expression to a certain stage of development.[39] Facultative heterochromatin is known to occupy both A and B compartments and Polycomb coated regions of the genome regularly interact with both A and B compartments in molecular dynamic

simulations.[31] A and B compartments occur on the scale of several megabases and within these compartments are Topologically Associated Domains (TADs), which are insulated by CTCF boundaries and can span hundreds of kilobases.[35] Chromatin within TADs is self-interacting and occupies a similar transcriptional state.[40] Within facultative heterochromatin, Polycomb Associated Domains (PADs), which are smaller than TADs at 20-140 kilobases, are coated with H3K27me3 and preferentially interact to form higher order structure via oligomerization of the PHC subunit of PRC1.[35,41,42] Knockout of *Ring1A* and *Ring1B* strongly reduces the strength of PAD boundaries in mouse oogenesis.[34] Previous work from our lab demonstrated that knockout of *Phc1* impacts PAD boundaries and leads to decompaction of *Hox* loci as seen by Stochastic Optical Reconstruction Microscopy (STORM).[41] PADs were shown to change during differentiation reflecting changes in binding of PRC1.[34,41] PADs were independent of ubiquitylation, suggesting ncPRC1 is not important for PAD formation.[41] Knockout of *Eed* did not impact established PADs in mouse oocytes, but did prevent the reestablishment of PADs in early embryonic development.[34] Therefore, PRC2 and PRC1 are thought to work together to establish PADs in higher order chromatin architecture as an important layer of PcG regulation.

As previously discussed, the PHC subunit of PRC1 can contribute to the formation of Polycomb condensates.[11] Therefore, we speculate that phase separation may play a role in higher order chromatin structure driven by PADs. Local nucleosome compaction and long-range contacts may both contribute to the formation of facultative 3D heterochromatin structure. While it is known that oligomerization of the PHC subunit of PRC1 has a major role in driving higher order structure,[41,42] it is unclear if the phase separation activity driven by the CBX subunit can also contribute to the formation of PADs. Future work will study the heterologous IDR lines from

Chapter 3 in Hi-C experiments to understand the potential influence of phase separation on genome structure, particularly at genes where the FUS^{IDR} rescued repression. Advances in Hi-C methods have improved the resolution of 3D genome mapping and consequently it may be possible to resolve shorter range Polycomb compaction as well.[43] Understanding the relationship between the higher order structure of PADs, chromatin compaction, and Polycomb phase separation will be important to develop a model for the mechanism of Polycomb-mediated gene repression.

The early models of transcriptional control by phase separation proposed membraneless compartments demixed from the surrounding chromatin as sites of activation or repression.[44] In the case of Polycomb, these nuclear bodies might be thought to concentrate repressive PcG components and occlude transcription factors or activating machinery.[1,2] However, this model is likely too simplistic. For Polycomb bodies, the dynamic nature of condensates seems to be an important feature and as described above Polycomb Associated Domains are intermixed with active and repressive TADs.[31] Additionally, our genome-wide analysis of Polycomb bound genes shows that many RING1B and CBX2 targets are not completely repressed. In these cases, the presence of Polycomb condensates may be tuning down transcriptional output without making the locus inaccessible to transcriptional machinery. For example, the *Npr2* gene is bound in all cell lines, but the overexpression of $CBX2_{WT}$ decreases gene expression whereas the $CBX2_{KRA}$ does not. This demonstrates that overexpression of $CBX2_{WT}$ is modulating transcription at this gene even while the gene is still expressed. In contrast, the $CBX2_{KRA}$ is incapable of tuning down gene expression at this locus, presumably due to its inability to form condensates. Additionally, if Polycomb condensates concentrate repressive factors, it would be expected see an increase in

RING1B and CBX2 binding at Polycomb target genes in the FUS^{IDR} line. However, we tend to see a decrease in RING1B and CBX2 occupancy upon induced phase separation in the CUT&RUN data. It is possible that condensates recruit more cPRC1 that is not bound but diffusing freely and therefore the local increase in cPRC1 concentration due to condensate formation is not reflected in the CUT&RUN signal. It also could be that condensate formation reduces the amount of PRC1 required for repression of a target gene. Alternative methods to measure whether Polycomb condensates increase recruitment of PRC1 to the target gene might include the use of proximity labelling methods such as TurboID.[44] This type of experiment might determine if there is an increase in PRC1 or PRC2 complexes in the general vicinity of the gene, when the FUS^{IDR} construct binds near the TSS. Conversely, the exclusion of activating factors such as chromatin remodelers or transcription factors could also be examined. Since facultative heterochromatin can be engaged by activating factors, it provides a unique opportunity to study the tug-of-war between repression and activation at genes.[39] Designing experiments to understand these dynamics will require precise temporal control to measure the transition at these sites as they moved from active to silent or silent to active rather than measuring an end state. Even a recent preprint demonstrating the porosity of CBX8 condensates suggests that these condensates are still impermeable to larger proteins such as transcription factors.[45] Thus, even if the edges of condensates are not completely impermeable boundaries, they still function to control access to bound genes.[45] Overall, understanding the dynamics of modulating gene repression by cPRC1 will be an important aspect of establishing a model for Polycomb condensates.

It will also be important to remember that distinct subsets of Polycomb bound target genes may exist. These subsets may behave differently with regards to regulation by Polycomb complex binding, repressive histone marks, or condensate formation. They also may vary in different cell lineages. For example, Hox clusters are often studied as typical Polycomb targets and must be stably turned off to maintain cell identity.[46] These loci are also heavily coated with H3K27me3 and H2AK119ub1 to reinforce gene silencing.[46] In the Neuro2a cell lines, the Hox genes were not affected by the presence of the $CBX2_{KRA}$ or the induced phase separation by the FUS^{IDR} and remained transcriptionally silent, even as other Polycomb target genes were modulated by the variation in Polycomb complexes. The Hox clusters may represent one extreme of a spectrum of Polycomb target genes, where repression is strongly reinforced and therefore these genes are insensitive to changes in one isolated aspect or another. There is also evidence in plants that changes in repression at Polycomb loci may occur stochastically over the course of many cell divisions.[47] In *Arabidopsis*, the *FLC* locus is a Polycomb target that behaves in a stochastic, switch-like manner digitally transitioning from an ON state to an OFF state upon cold exposure with positive feedback reinforcing the OFF state over time. Therefore, it might require studying Polycomb on longer timescales and at a single-cell level to observe changes in the maintenance of gene expression over time.

4.4 Polycomb condensates in maintenance and memory

It has long been known that Polycomb proteins play a role in epigenetic memory to maintain cell identity.[46] Further understanding of the function of Polycomb bodies in facilitating epigenetic memory will be essential to establishing a comprehensive model. Condensate formation could be a potential mechanism for the quick reestablishment of silencing after cell division. Work from

the Ren lab has shown that ~96% of CBX2 remains bound to mitotic chromosomes as shown by live cell imaging of fluorescently tagged proteins in mESCs.[48] The other CBX proteins were bound to a lesser extent: ~22% of CBX4, ~44% of CBX7, and ~40% of CBX8. CBX2 was required for the association of other cPRC1 subunits with mitotic chromosomes and directly recruited them.[48] In contrast, the recruitment of CBX2 to mitotic chromosomes was independent of PRC1 or PRC2 proteins.[48] CBX2's stability on chromatin throughout cell division might make it responsible for regenerating Polycomb bodies, promoting histone marks, and reestablishing silent domain of heterochromatin to maintain repression of lineage non-specific genes (**Fig. 4.3**). It has recently been shown that compaction of chromatin by the H1 histone is important for reestablishing H3K27me3 and H2AK119ub1 after DNA replication and cell division.[49] Compaction by H1 helps to propagate the H2AK119ub1 mark.[49] There is a positive feedback loop whereby binding RYBP or YAF2, subunits of ncPRC1, to H2AK119ub1 recruits ncPRC1 and thereby spreads the H2AK119ub1 mark to neighboring nucleosomes.[49] H2AK119ub1 is also capable of recruiting PRC2 in a hierarchical model of Polycomb complexes to promote establishment of H3K27me3 as well.[50] In a more direct fashion, another study has shown that H1 compaction helps to restore H3K27me3 at dense repressive regions of the genome.[51] Knockout of H1 strongly reduced H3K27me3 levels particularly at regions of the genome that were highly repressed and typically recover H3K27me3 most quickly after replication.[51] Therefore, H1's role in compaction may work synergistically with CBX2 compaction and phase separation to promote repressive histone marks. Study of histone turnover during replication demonstrates symmetric dilution of histones into daughter strands.[52] Histones with repressive marks such as H3K27me3 are preferentially recycled at the same loci in daughter cells, but histones in active regions of the genome are not.[53] This local redeposition of repressive histones signifies the particular

importance of epigenetic memory for repression. Local redeposition may facilitate the quick reestablishment of PADs by stimulating the deposition of H3K27me3 and recruiting Polycomb complexes.[53]

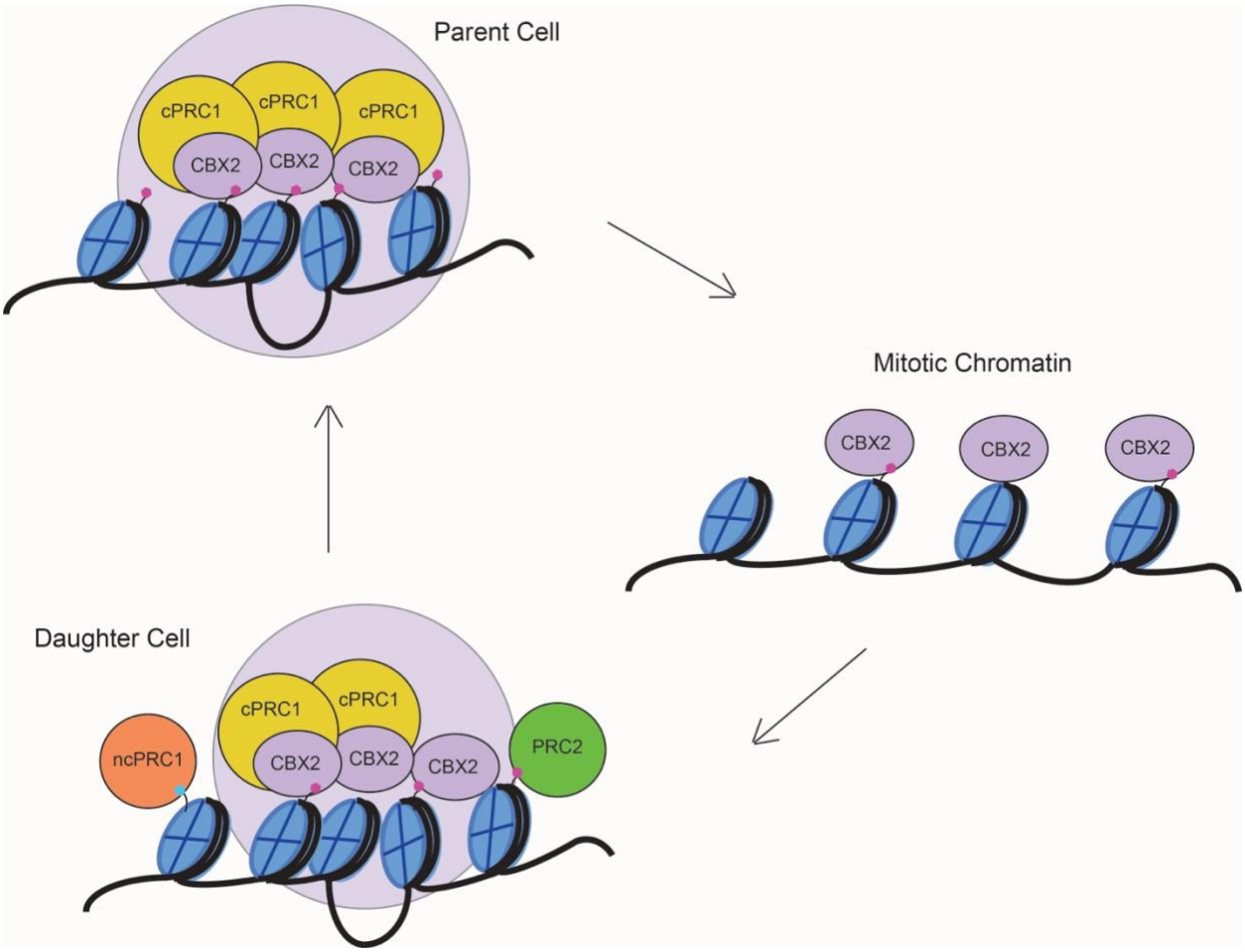


Figure 4.3 Model for CBX2 phase separation in epigenetic memory

CBX2 phase separation as a mechanism of epigenetic memory during cell division to reestablish Polycomb silencing.

Another potential mechanism for repression by cPRC1 condensates may be the spreading of repressive histone marks stimulated by compaction and phase separation. H3K27me3 spreading may reestablish PADs after cell division in epigenetic memory, but it may also be a more general mechanism of reinforcing gene repression.[54] PRC1 complexes are the main component of Polycomb bodies, however PRC2 might also partition dynamically into Polycomb condensates. EED has been shown to directly interact with cPRC1 helping to recruit cPRC1 to sites of H3K27me3.[55,56] PRC2 deposits H3K27me3 and this may create a positive feedback loop to recruit more PRC1 for compaction and silencing. Recruitment experiments show that heritable silencing requires both PRC2 deposition of H3K27me3 and cPRC1 function.[57] Therefore, PRC1 condensates may help to concentrate PRC2 or directly stimulate PRC2 activity to promote H3K27me3 spreading. Phase separation has been shown to catalyze other histone marks such as H2B ubiquitination.[58] The rescue of gene repression in the FUS^{IDR} line was highly correlated with changes in H3K27me3, suggesting phase separation might directly stimulate H3K27me3 to facilitate repression. To further understand the relationship between Polycomb condensates and H3K27me3, PRC2 inhibitors such as UNC1999 or GSK343 could be used to reduce global levels of H3K27me3.[59,60] If H3K27me3 is necessary for the mechanism of silencing by phase separation, then the rescue of gene repression in the FUS^{IDR} line will be compromised in the inhibitor treated sample. This experiment would help to elucidate the role of H3K27me3 in the proposed Polycomb condensation-driven mechanism of gene repression. In contrast to H3K27me3, H2AK119ub1 was less reliably correlated with gene repression in the Neuro2a cell lines. Work from the Brangwynne lab has demonstrated the writing of H2AK119ub1 by induction of CBX2 phase separation using the synthetic corelet system.[17] It is known that ncPRC1 and the H2AK119ub1 mark play important roles in mESCs to regulate developmental

genes.[61,62] However, the role for ncPRC1 in differentiated cell lineages is unclear. Recent work has shown that H2AK119ub1 can increase transcription at Polycomb target genes by inhibiting cPRC1 binding to nucleosomes and preventing condensation.[63] In line with these observations, immunofluorescence experiments in the FUS^{IDR} line did not show enrichment of ncPRC1 subunit, RYBP, in the induced Polycomb bodies. The use of a PRC1 inhibitor, such as RB-3 to deplete H2AK119ub1 would be useful to test if this histone mark is inhibiting cPRC1 binding in Neuro2a cells.[64] Upon treatment, we would expect see an increase in CUT&RUN signal for RING1B and CBX2 at target genes if H2AK119ub1 is interfering with the binding of cPRC1. Additionally, we would expect to see a decrease in gene expression as cPRC1 binding increased. It would be interesting to note if the phase separation induced by the FUS^{IDR} altered gene expression in the absence of H2AK119ub1. The use of PRC2 or ncPRC1 catalytic inhibitors could be a useful tool in teasing out the relationship between condensate formation, the spreading of histone marks, and gene repression.

4.5 Polycomb condensates in disease

The experiments performed in Chapter 3 demonstrate the importance of Polycomb function in cancer pathology. CBX2 plays a tumor promoting role in both neuroblastoma and breast cancer by supporting an aggressive, stem cell-like phenotype.[65-67] CBX2's phase separation activity helps to drive the dependence these cancer cell lines have on Polycomb function. The $CBX2_{KRA}$ line is less stem-like and differentiates more in the neural differentiation assay compared to the $CBX2_{WT}$ overexpression line. Fusion of the FUS^{IDR} to $CBX2_{KRA}$ not only rescues CBX2 condensate formation and gene repression, but also rescues a cancer stem cell phenotype. In a similar fashion, the FUS^{IDR} rescues the aggressive transwell migration phenotype of the 4T1

breast cancer cell line. As discussed in Chapter 1, it is increasingly clear that PcG proteins play important roles in cancer, however the tumor-promoting mechanisms often vary with cell lineage.[68] The overexpression of CBX2 is found in several cancer types, and this likely increases Polycomb body formation and stability.[65-68] Our work confirms the importance of CBX2 epigenetic regulation in maintaining the cancer stem cell niche and supports a role for increased condensate formation in driving cancer pathology.

Beyond cancer, PcG proteins likely play an important role in aging. The CBX paralogs have been implicated in cellular senescence and play important roles in maintaining adult stem cell populations.[69-71] Loss of CBX2 has been shown to induce genome instability that leads to premature senescence in fibroblasts.[72] The PRC2 clock has even been proposed as a companion to the DNA methylation clock, which has been widely discussed as a biomarker of aging.[73] CpG islands have low levels of DNA methylation and are known to recruit PRC2.[73] Over the course of aging, CpG islands enriched for PRC2 increase in methylation levels, likely reducing PRC2 binding and destabilizing Polycomb regulation.[73,74] Rejuvenation treatments such as calorie restriction for mice resulted in a reversal of the PRC2 clock.[73] Thus, there seems to be evidence of a direct relationship between DNA methylation and Polycomb silencing. It would be interesting to examine the changes in Polycomb body formation in models of aging and cellular senescence. It seems likely that CBX2 condensate formation plays an important role in regulating cellular health by stabilizing heterochromatin and maintaining gene repression.

The significant role of CBX proteins in cancer and aging suggests the potential for a therapeutic strategy targeting Polycomb condensate formation. The development of small molecules

targeting IDRs has been challenging, particularly because the lessons from targeting structured domains of proteins may be inappropriate for designing small molecules against disordered regions.[75,76] For example, drug development targeting the disordered N-terminus of the androgen receptor was reported to only be possible with natural product libraries.[76] Chromatin interacting proteins are often disordered.[77] Therefore, modulating gene regulation will likely take developing new strategies for targeting IDRs for drug development. Advances in the development of heterobifunctional degraders may be a good place to start for this class of "undruggable" proteins.[78,79] In these cases, weak binders to the IDR of the protein of interest could be leveraged for degradation by PROTACs or molecular glues. As the diversity and function of IDRs are discerned, the development of strategies for drug development will be more and more relevant. One exciting example is the development of a potent degrader targeting MYC a white whale of an oncogene that is highly disordered.[80] We are just at the tip of the iceberg when it comes to understanding the biology of disordered proteins and developing targeted therapeutics. Several biotech companies are tackling the challenges of drugging disordered proteins highlighting the innovation and promise of targeting IDRs in many disease contexts.

As this dissertation comes to a close, I can only hope to have made a small contribution to the march of scientific advancement. At the beginning of my graduate studies, I was captivated by the Kuhnian paradigm shift that phase separation represented for the field of gene regulation. Hopefully, this work has deepened our understanding of the mechanisms at play in Polycomb-mediated gene repression as we continue to discover the function of intrinsically disordered proteins in chromatin biology and disease.

References

1. Plys, A. J., Davis, C. P., Kim, J., Rizki, G., Keenen, M. M., Marr, S. K., and Kingston, R. E. (2019). Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes & development*, *33*(13-14), 799–813. https://doi.org/10.1101/gad.326488.119

2. Tatavosian, R., Kent, S., Brown, K., Yao, T., Duc, H. N., Huynh, T. N., Zhen, C. Y., Ma, B., Wang, H., and Ren, X. (2019). Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. *The Journal of biological chemistry*, *294*(5), 1451–1463. https://doi.org/10.1074/jbc.RA118.006620

3. Morey, L., Pascual, G., Cozzuto, L., Roma, G., Wutz, A., Benitah, S. A., and Di Croce, L. (2012). Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. *Cell stem cell*, *10*(1), 47–62. https://doi.org/10.1016/j.stem.2011.12.006

4. Jaensch, E. S., Zhu, J., Cochrane, J. C., Marr, S. K., Oei, T. A., Damle, M., McCaslin, E. Z., and Kingston, R. E. (2021). A Polycomb domain found in committed cells impairs differentiation when introduced into PRC1 in pluripotent cells. *Molecular cell*, *81*(22), 4677–4691.e8. https://doi.org/10.1016/j.molcel.2021.09.018

5. Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., and Lander, E. S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, *125*(2), 315–326. https://doi.org/10.1016/j.cell.2006.02.041

6. Lee, T. I., Jenner, R. G., Boyer, L. A., Guenther, M. G., Levine, S. S., Kumar, R. M., Chevalier, B., Johnstone, S. E., Cole, M. F., Isono, K., Koseki, H., Fuchikami, T., Abe, K., Murray, H. L., Zucker, J. P., Yuan, B., Bell, G. W., Herbolsheimer, E., Hannett, N. M., Sun, K., Odom, D. T., Otte, A. P., Volkert, T. L., Bartel, D. P., Melton, D. A., Gifford, D. K., Jaenisch, R., and Young, R. A. (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell*, *125*(2), 301–313. https://doi.org/10.1016/j.cell.2006.02.043

7. Ma, R. G., Zhang, Y., Sun, T. T., and Cheng, B. (2014). Epigenetic regulation by polycomb group complexes: focus on roles of CBX proteins. *Journal of Zhejiang University. Science. B*, *15*(5), 412–428. https://doi.org/10.1631/jzus.B1400077

8. Vincenz, C. and Kerppola, T. K. (2008). Different polycomb group CBX family proteins associate with distinct regions of chromatin using nonhomologous protein sequences. *Proceedings of the National Academy of Sciences*, *106*(43), 16572-16577. https://doi.org/10.1073/pnas.0805317105

9. Klauke, K., Radulović, V., Broekhuis, M., Weersing, E., Zwart, E., Olthof, S., Ritsema, M., Bruggeman, S., Wu, X., Helin, K., Bystrykh, L., and de Haan, G. (2013). Polycomb Cbx

family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. *Nature cell biology*, *15*(4), 353–362. https://doi.org/10.1038/ncb2701

10. Pemberton, H., Anderton, E., Patel, H., Brookes, S., Chandler, H., Palermo, R., Stock, J., Rodriguez-Niedenführ, M., Racek, T., de Breed, L., Stewart, A., Matthews, N., and Peters, G. (2014). Genome-wide co-localization of Polycomb orthologs and their effects on gene expression in human fibroblasts. *Genome biology*, *15*(2), R23. https://doi.org/10.1186/gb-2014-15-2-r23

11. Niekamp, S., Marr, S. K., Oei, T. A., Subramanian, R., and Kingston, R. E. (2024). Modularity of PRC1 composition and chromatin interaction define condensate properties. *Molecular cell*, S1097-2765(24)00181-3. Advance online publication. https://doi.org/10.1016/j.molcel.2024.03.001

12. Brown, K., Chew, P. Y., Ingersoll, S., Espinosa, J. R., Aguirre, A., Espinoza, A., Wen, J., Astatike, K., Kutateladze, T. G., Collepardo-Guevara, R., and Ren, X. (2023). Principles of assembly and regulation of condensates of Polycomb repressive complex 1 through phase separation. *Cell reports*, *42*(10), 113136. https://doi.org/10.1016/j.celrep.2023.113136

13. Ingersoll, S., Trouth, A., Luo, X., Espinoza, A., Wen, J., Tucker, J., Astatike, K., Phiel, C. J., Kutateladze, T. G., Wu, T. P., Ramachandran, S., and Ren, X. (2024). Sparse CBX2 nucleates many Polycomb proteins to promote facultative heterochromatinization of Polycomb target genes. *bioRxiv*, 2024.02.05.578969. https://doi.org/10.1101/2024.02.05.578969

14. Dobrinić, P., Szczurek, A. T., and Klose, R. J. (2021). PRC1 drives Polycomb-mediated gene repression by controlling transcription initiation and burst frequency. *Nature structural & molecular biology*, *28*(10), 811–824. https://doi.org/10.1038/s41594-021-00661-y

15. Henikoff, S., and Shilatifard, A. (2011). Histone modification: cause or cog?. *Trends in genetics*, *27*(10), 389–396. https://doi.org/10.1016/j.tig.2011.06.006

16. O'Geen, H., Ren, C., Nicolet, C. M., Perez, A. A., Halmai, J., Le, V. M., Mackay, J. P., Farnham, P. J., and Segal, D. J. (2017). dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic acids research*, *45*(17), 9901–9916. https://doi.org/10.1093/nar/gkx578

17. Eeftens, J. M., Kapoor, M., Michieletto, D., and Brangwynne, C. P. (2021). Polycomb condensates can promote epigenetic marks but are not required for sustained chromatin compaction. *Nature communications*, *12*(1), 5888. https://doi.org/10.1038/s41467-021-26147-5

18. Senthilkumar, R., and Mishra, R.K. (2009). Novel motifs distinguish multiple homologues of *Polycomb* in vertebrates: expansion and diversification of the epigenetic toolkit. *BMC Genomics*, *10*, 549. https://doi.org/10.1186/1471-2164-10-549

19. Phan, T. M., Kim, Y. C., Debelouchina, G. T., and Mittal, J. (2024). Interplay between charge distribution and DNA in shaping HP1 paralog phase separation and localization. *eLife, 12,* RP90820. https://doi.org/10.7554/eLife.90820.3

20. Kent, S., Brown, K., Yang, C. H., Alsaihati, N., Tian, C., Wang, H., and Ren, X. (2020). Phase-Separated Transcriptional Condensates Accelerate Target-Search Process Revealed by Live-Cell Single-Molecule Imaging. *Cell reports*, *33*(2), 108248. https://doi.org/10.1016/j.celrep.2020.108248

21. Banani, S. F., Rice, A. M., Peeples, W. B., Lin, Y., Jain, S., Parker, R., and Rosen, M. K. (2016). Compositional Control of Phase-Separated Cellular Bodies. *Cell*, *166*(3), 651–663. https://doi.org/10.1016/j.cell.2016.06.010

22. Connelly, K. E., Weaver, T. M., Alpsoy, A., Gu, B. X., Musselman, C. A., and Dykhuizen, E. C. (2019). Engagement of DNA and H3K27me3 by the CBX8 chromodomain drives chromatin association. *Nucleic acids research*, *47*(5), 2289–2305. https://doi.org/10.1093/nar/gky1290

23. Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Molecular and cellular biology*, *26*(7), 2560–2569. https://doi.org/10.1128/MCB.26.7.2560-2569.2006

24. Bezsonova, I., Walker, J. R., Bacik, J. P., Duan, S., Dhe-Paganon, S., and Arrowsmith, C. H. (2009). Ring1B contains a ubiquitin-like docking module for interaction with Cbx proteins. *Biochemistry*, *48*(44), 10542–10548. https://doi.org/10.1021/bi901131u

25. Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M., and Sixma, T. K. (2006). Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. *The EMBO journal*, *25*(11), 2465–2474. https://doi.org/10.1038/sj.emboj.7601144

26. Gray, F., Cho, H. J., Shukla, S., He, S., Harris, A., Boytsov, B., Jaremko, Ł., Jaremko, M., Demeler, B., Lawlor, E. R., Grembecka, J., and Cierpicki, T. (2016). BMI1 regulates PRC1 architecture and activity through homo- and hetero-oligomerization. *Nature communications*, *7*, 13343. https://doi.org/10.1038/ncomms13343

27. Gannon, D., Lagan, E., Silva, A. J., Bibawi, P., Doherty, A. M., Nimmo, D., McCole, R., Monger, C., Genesi, G., Dibekeme, A. V., Innes, J. A., Yang, L., Chen, B., van Mierlo, G., Jansen, P. W. T. C., Wynne, K., Sanchez-Rivera, F. J., Soto-Feliciano, Y. M., Vermeulen, M., Oliverio, G., Chen, C., Philips, R. E., Bracken, A. P., and Brien, G. L. (2023). Specific cPRC1 complexes are co-opted to mediate oncogenic gene repression in diffuse midline glioma. *bioRxiv*, 2023.12.02.569451. https://doi.org/10.1101/2023.12.02.569451

28. Grau, D. J., Chapman, B. A., Garlick, J. D., Borowsky, M., Francis, N. J., and Kingston, R. E. (2011). Compaction of chromatin by diverse Polycomb group proteins requires localized

regions of high charge. *Genes & development*, *25*(20), 2210–2221. https://doi.org/10.1101/gad.17288211

29. Lau, M. S., Schwartz, M. G., Kundu, S., Savol, A. J., Wang, P. I., Marr, S. K., Grau, D. J., Schorderet, P., Sadreyev, R. I., Tabin, C. J., and Kingston, R. E. (2017). Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. *Science*, *355*(6329), 1081–1084. https://doi.org/10.1126/science.aah5403

30. Williamson, I., Boyle, S., Grimes, G. R., Friman, E. T., and Bickmore, W. A. (2023). Dispersal of PRC1 condensates disrupts polycomb chromatin domains and loops. *Life science alliance*, *6*(10), e202302101. https://doi.org/10.26508/lsa.202302101

31. Murphy, S. E., and Boettiger, A. N. (2024). Polycomb repression of Hox genes involves spatial feedback but not domain compaction or phase transition. *Nature genetics*, *56*(3), 493–504. https://doi.org/10.1038/s41588-024-01661-6

32. Mansisidor, A. R., and Risca, V. I. (2022). Chromatin accessibility: methods, mechanisms, and biological insights. *Nucleus*, *13*(1), 236–276. https://doi.org/10.1080/19491034.2022.2143106

33. Poonperm, R., and Hiratani, I. (2021). Formation of a multi-layered 3-dimensional structure of the heterochromatin compartment during early mammalian development. *Development, growth & differentiation*, *63*(1), 5–17. https://doi.org/10.1111/dgd.12709

34. Du, Z., Zheng, H., Kawamura, Y. K., Zhang, K., Gassler, J., Powell, S., Xu, Q., Lin, Z., Xu, K., Zhou, Q., Ozonov, E. A., Véron, N., Huang, B., Li, L., Yu, G., Liu, L., Au Yeung, W. K., Wang, P., Chang, L., Wang, Q., He, A., Sun, Y., Na, J., Sun, Q., Sasaki, H., Tachibana, K., Peters, A. H. F. M., and Xie, W. (2020). Polycomb Group Proteins Regulate Chromatin Architecture in Mouse Oocytes and Early Embryos. *Molecular cell*, *77*(4), 825–839.e7. https://doi.org/10.1016/j.molcel.2019.11.011

35. Guo, Y., and Wang, G. G. (2022). Modulation of the high-order chromatin structure by Polycomb complexes. *Frontiers in cell and developmental biology*, *10*, 1021658. https://doi.org/10.3389/fcell.2022.1021658

36. Berger, F. (2019). Emil Heitz, a true epigenetics pioneer. *Nat Rev Mol Cell Biol*, *20*(572). https://doi.org/10.1038/s41580-019-0161-z

37. Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., and Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*(5950), 289–293. https://doi.org/10.1126/science.1181369

38. Penagos-Puig, A., and Furlan-Magaril, M. (2020). Heterochromatin as an Important Driver of Genome Organization. *Frontiers in cell and developmental biology*, *8*, 579137. https://doi.org/10.3389/fcell.2020.579137

39. Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? *Molecular cell*, *28*(1), 1–13. https://doi.org/10.1016/j.molcel.2007.09.011

40. Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., and Aiden, E. L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, *159*(7), 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021

41. Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J. T., Sadreyev, R. I., Dekker, J., and Kingston, R. E. (2017). Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Molecular cell*, *65*(3), 432–446.e5. https://doi.org/10.1016/j.molcel.2017.01.009

42. Isono, K., Endo, T. A., Ku, M., Yamada, D., Suzuki, R., Sharif, J., Ishikura, T., Toyoda, T., Bernstein, B. E., and Koseki, H. (2013). SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. *Developmental cell*, *26*(6), 565–577. https://doi.org/10.1016/j.devcel.2013.08.016

43. Bouwman, B. A. M., Crosetto, N., and Bienko, M. (2022). The era of 3D and spatial genomics. *Trends in genetics*, *38*(10), 1062–1075. https://doi.org/10.1016/j.tig.2022.05.010

44. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K., and Sharp, P. A. (2017). A Phase Separation Model for Transcriptional Control. *Cell*, *169*(1), 13–23. https://doi.org/10.1016/j.cell.2017.02.007

45. Uckelmann, M., Levina, V., Taveneau, C., Han Ng, X., Pandey, V., Martinez, J., Mendiratta, S., Houx, J., Boudes, M., Venugopal, H., Trépout, S., Zhang, Q., Flanigan, S., Li, M., Sierecki, E., Gambin, Y., Das, P. P., Bell, O., de Marco, A., and Davidovich, C. (2024). Dynamic PRC1-CBX8 stabilizes a porous structure of chromatin condensates. *bioRxiv*, 2023.05.08.539931. https://doi.org/10.1101/2023.05.08.539931

46. Schuettengruber, B., Bourbon, H. M., Di Croce, L., and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell*, *171*(1), 34–57. https://doi.org/10.1016/j.cell.2017.08.002

47. Menon, G., Schulten, A., Dean, C., and Howard, M. (2021). Digital paradigm for Polycomb epigenetic switching and memory. *Current opinion in plant biology*, *61*, 102012. https://doi.org/10.1016/j.pbi.2021.102012

48. Zhen, C. Y., Duc, H. N., Kokotovic, M., Phiel, C. J., and Ren, X. (2014). Cbx2 stably associates with mitotic chromosomes via a PRC2- or PRC1-independent mechanism and is

needed for recruiting PRC1 complex to mitotic chromosomes. *Molecular biology of the cell*, *25*(23), 3726–3739. https://doi.org/10.1091/mbc.E14-06-1109

49. Zhao, J., Wang, M., Chang, L., Yu, J., Song, A., Liu, C., Huang, W., Zhang, T., Wu, X., Shen, X., Zhu, B., and Li, G. (2020). RYBP/YAF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2AK119ub1 during cell division. *Nature cell biology*, *22*(4), 439–452. https://doi.org/10.1038/s41556-020-0484-1

50. van Kruijsbergen, I., Hontelez, S., and Veenstra, G. J. (2015). Recruiting polycomb to chromatin. *The international journal of biochemistry & cell biology*, *67*, 177–187. https://doi.org/10.1016/j.biocel.2015.05.006

51. Liu, C., Yu, J., Song, A., Wang, M., Hu, J., Chen, P., Zhao, J., and Li, G. (2023). Histone H1 facilitates restoration of H3K27me3 during DNA replication by chromatin compaction. *Nat Commun*, 14(4081). https://doi.org/10.1038/s41467-023-39846-y

52. Wenger, A., Biran, A., Alcaraz, N., Redó-Riveiro, A., Sell, A. C., Krautz, R., Flury, V., Reverón-Gómez, N., Solis-Mezarino, V., Völker-Albert, M., Imhof, A., Andersson, R., Brickman, J. M., and Groth, A. (2023). Symmetric inheritance of parental histones governs epigenome maintenance and embryonic stem cell identity. *Nature genetics*, *55*(9), 1567–1578. https://doi.org/10.1038/s41588-023-01476-x

53. Escobar, T. M., Oksuz, O., Saldaña-Meyer, R., Descostes, N., Bonasio, R., and Reinberg, D. (2019). Active and Repressed Chromatin Domains Exhibit Distinct Nucleosome Segregation during DNA Replication. *Cell*, *179*(4), 953–963.e11. https://doi.org/10.1016/j.cell.2019.10.009

54. Gaydos, L. J., Wang, W., and Strome, S. (2014). Gene repression. H3K27me and PRC2 transmit a memory of repression across generations and during development. *Science*, *345*(6203), 1515–1518. https://doi.org/10.1126/science.1255023

55. Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., 3rd, Voigt, P., Martin, S. R., Taylor, W. R., De Marco, V., Pirrotta, V., Reinberg, D., and Gamblin, S. J. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature*, *461*(7265), 762–767. https://doi.org/10.1038/nature08398

56. Cao, Q., Wang, X., Zhao, M., Yang, R., Malik, R., Qiao, Y., Poliakov, A., Yocum, A. K., Li, Y., Chen, W., Cao, X., Jiang, X., Dahiya, A., Harris, C., Feng, F. Y., Kalantry, S., Qin, Z. S., Dhanasekaran, S. M., and Chinnaiyan, A. M. (2014). The central role of EED in the orchestration of polycomb group complexes. *Nature communications*, *5*, 3127. https://doi.org/10.1038/ncomms4127

57. Moussa, H. F., Bsteh, D., Yelagandula, R., Pribitzer, C., Stecher, K., Bartalska, K., Michetti, L., Wang, J., Zepeda-Martinez, J. A., Elling, U., Stuckey, J. I., James, L. I., Frye, S. V., and Bell, O. (2019). Canonical PRC1 controls sequence-independent propagation of Polycomb-mediated gene silencing. *Nature communications*, *10*(1), 1931. https://doi.org/10.1038/s41467-019-09628-6

58. Gallego, L. D., Schneider, M., Mittal, C., Romanauska, A., Gudino Carrillo, R. M., Schubert, T., Pugh, B. F., and Köhler, A. (2020). Phase separation directs ubiquitination of gene-body nucleosomes. *Nature*, *579*(7800), 592–597. https://doi.org/10.1038/s41586-020-2097-z

59. Konze, K. D., Ma, A., Li, F., Barsyte-Lovejoy, D., Parton, T., Macnevin, C. J., Liu, F., Gao, C., Huang, X. P., Kuznetsova, E., Rougie, M., Jiang, A., Pattenden, S. G., Norris, J. L., James, L. I., Roth, B. L., Brown, P. J., Frye, S. V., Arrowsmith, C. H., Hahn, K. M., Wang, G. G., Vedadi, M., and Jin, J. (2013). An orally bioavailable chemical probe of the Lysine Methyltransferases EZH2 and EZH1. *ACS chemical biology*, *8*(6), 1324–1334. https://doi.org/10.1021/cb400133j

60. Yu, T., Wang, Y., Hu, Q., Wu, W., Wu, Y., Wei, W., Han, D., You, Y., Lin, N., and Liu, N. (2017). The EZH2 inhibitor GSK343 suppresses cancer stem-like phenotypes and reverses mesenchymal transition in glioma cells. *Oncotarget*, *8*(58), 98348–98359. https://doi.org/10.18632/oncotarget.21311

61. Tamburri, S., Lavarone, E., Fernández-Pérez, D., Conway, E., Zanotti, M., Manganaro, D., and Pasini, D. (2020). Histone H2AK119 Mono-Ubiquitination Is Essential for Polycomb-Mediated Transcriptional Repression. *Molecular cell*, *77*(4), 840–856.e5. https://doi.org/10.1016/j.molcel.2019.11.021

62. Blackledge, N. P., Fursova, N. A., Kelley, J. R., Huseyin, M. K., Feldmann, A., and Klose, R. J. (2020). PRC1 Catalytic Activity Is Central to Polycomb System Function. *Molecular cell*, *77*(4), 857–874.e9. https://doi.org/10.1016/j.molcel.2019.12.001

63. Zhao, J., Lan, J., Wang, M., Liu, C., Fang, Z., Song, A., Zhang, T., Wang, L., Zhu, B., Chen, P., Yu, J., and Li, G. (2024). H2AK119ub1 differentially fine-tunes gene expression by modulating canonical PRC1- and H1-dependent chromatin compaction. *Molecular cell*, *84*(7), 1191–1205.e7. https://doi.org/10.1016/j.molcel.2024.02.017

64. Shukla, S., Ying, W., Gray, F., Yao, Y., Simes, M. L., Zhao, Q., Miao, H., Cho, H. J., González-Alonso, P., Winkler, A., Lund, G., Purohit, T., Kim, E., Zhang, X., Ray, J. M., He, S., Nikolaidis, C., Ndoj, J., Wang, J., Jaremko, Ł., Jaremko, M., Ryan, R. J. H., Guzman, M. L., Grembecka, J. and Cierpicki, T. (2021). Small-molecule inhibitors targeting Polycomb repressive complex 1 RING domain. *Nature chemical biology*, *17*(7), 784–793. https://doi.org/10.1038/s41589-021-00815-5

65. Gu, X., Wang, X., Su, D., Su, X., Lin, L., Li, S., Wu, Q., Liu, S., Zhang, P., Zhu, X., and Jiang, X. (2018). CBX2 Inhibits Neurite Development by Regulating Neuron-Specific Genes Expression. *Frontiers in molecular neuroscience*, *11*, 46. https://doi.org/10.3389/fnmol.2018.00046

66. Zheng, S., Lv, P., Su, J., Miao, K., Xu, H., and Li, M. (2019). Overexpression of CBX2 in breast cancer promotes tumor progression through the PI3K/AKT signaling pathway. *American journal of translational research*, *11*(3), 1668–1682.

67. Bilton, L. J., Warren, C., Humphries, R. M., Kalsi, S., Waters, E., Francis, T., Dobrowinski, W., Beltran-Alvarez, P., and Wade, M. A. (2022). The Epigenetic Regulatory Protein CBX2 Promotes mTORC1 Signalling and Inhibits DREAM Complex Activity to Drive Breast Cancer Cell Growth. *Cancers*, *14*(14), 3491. https://doi.org/10.3390/cancers14143491

68. Parreno, V., Martinez, A. M., and Cavalli, G. (2022). Mechanisms of Polycomb group protein function in cancer. *Cell research*, *32*(3), 231–253. https://doi.org/10.1038/s41422-021-00606-6

69. Dietrich, N., Bracken, A. P., Trinh, E., Schjerling, C. K., Koseki, H., Rappsilber, J., Helin, K., and Hansen, K. H. (2007). Bypass of senescence by the polycomb group protein CBX8 through direct binding to the INK4A-ARF locus. *The EMBO journal*, *26*(6), 1637–1648. https://doi.org/10.1038/sj.emboj.7601632

70. Chen, Y. H., Zhang, X., Ko, K. Y., Hsueh, M. F., and Kraus, V. B. (2022). CBX4 Regulates Replicative Senescence of WI-38 Fibroblasts. *Oxidative medicine and cellular longevity*, *2022*, 5503575. https://doi.org/10.1155/2022/5503575

71. Klauke, K., Radulović, V., Broekhuis, M., Weersing, E., Zwart, E., Olthof, S., Ritsema, M., Bruggeman, S., Wu, X., Helin, K., Bystrykh, L., and de Haan, G. (2013). Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. *Nature cell biology*, *15*(4), 353–362. https://doi.org/10.1038/ncb2701

72. Baumann, C., Zhang, X., and De La Fuente, R. (2020). Loss of CBX2 induces genome instability and senescence-associated chromosomal rearrangements. *The Journal of cell biology*, *219*(11), e201910149. https://doi.org/10.1083/jcb.201910149

73. Moqri, M. , Cipriano, A. , Nachun, D. , Murty, T. , de Sena Brandine, G. , Rasouli, S. , Tarkhov, A. , Aberg, K. A. , van den Oord, E. , Zhou, W., and Smith, A. (2022). PRC2 clock: a universal epigenetic biomarker of aging and rejuvenation. *bioRxiv*, 2022.06.03.494609.

74. Mozhui, K., and Pandey, A. K. (2017). Conserved effect of aging on DNA methylation and association with EZH2 polycomb protein in mice and humans. *Mechanisms of ageing and development*, *162*, 27–37. https://doi.org/10.1016/j.mad.2017.02.006

75. Ruan, H., Sun, Q., Zhang, W., Liu, Y., and Lai, L. (2019). Targeting intrinsically disordered proteins at the edge of chaos. *Drug discovery today*, *24*(1), 217–227. https://doi.org/10.1016/j.drudis.2018.09.017

76. Sadar M. D. (2020). Discovery of drugs that directly target the intrinsically disordered region of the androgen receptor. *Expert opinion on drug discovery*, *15*(5), 551–560. https://doi.org/10.1080/17460441.2020.1732920

77. Musselman, C. A., and Kutateladze, T. G. (2021). Characterization of functional disordered regions within chromatin-associated proteins. *iScience*, *24*(2), 102070. https://doi.org/10.1016/j.isci.2021.102070

78. Samarasinghe, K. T. G., and Crews, C. M. (2021). Targeted protein degradation: A promise for undruggable proteins. *Cell chemical biology*, *28*(7), 934–951. https://doi.org/10.1016/j.chembiol.2021.04.011

79. Schreiber S. L. (2021). The Rise of Molecular Glues. *Cell*, *184*(1), 3–9. https://doi.org/10.1016/j.cell.2020.12.020

80. Xu, Y., Yu, Q., Wang, P., Wu, Z., Zhang, L., Wu, S., Li, M., Wu, B., Li, H., Zhuang, H., Zhang, X., Huang, Y., Gan, X., and Xu, R. (2022). A Selective Small-Molecule c-Myc Degrader Potently Regresses Lethal c-Myc Overexpressing Tumors. *Advanced science*, *9*(8), e2104344. https://doi.org/10.1002/advs.202104344