

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Recognition of Histone 3 Lysine 9 Dimethylation Regulates Its Production

Permalink

<https://escholarship.org/uc/item/3s04f72q>

Author

Simental, Eric Alexander

Publication Date

2024

Peer reviewed|Thesis/dissertation

Recognition of Histone 3 Lysine 9 Dimethylation Regulates Its Production

by
Eric Simental

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in
Biochemistry and Molecular Biology

in the
GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:



399BB46C754548E...

Abigail Cool

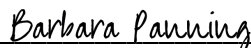
Chair

DocuSigned by:



DocuSigned by: 4 CA...

Bassem Al-Sady



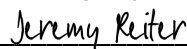
Barbara Panning

DocuSigned by: 17...



Bo Huang

DocuSigned by: 4 F3...



Jeremy Reiter

F07C889D1B164B3...

Committee Members

Acknowledgments

If you know me even a little bit, you know that I can sometimes be quite dramatic. As such, my internal working title for this thesis, taken from a favorite album of mine, was *Años en Infierno*. And if you know me a little bit more, you'll know immediately why I can sometimes refer to the last decade as such. Therefore, in the spirit of gratitude and positive reflection, I must start this thesis by acknowledging everyone who helped me make it through the last decade.

I'd first like to express the deepest and sincerest thank you to my parents, Alejandro and Myriam. I know I was a weird, difficult little guy. I also know that for quite a while I failed to make clear that this negativity was neither directed at nor a consequence of y'all. Thus, let me state as a matter of public record that y'all are *excellent* parents. To my dad, I remember my earliest years were spent trying to be like you. So much of who I am, from my interests to my work ethic, to my sense of confidence, to the way I dress, is a result of all those years I spent at your side. To my mom, I would not care about anyone or be a scientist without you. You were literally my first teacher. You taught me to read and spell and draw and express myself. You were the first person to ever show me love, and the love I receive from you to this day exceeds that of any normal human interaction I've yet encountered. As a son, I am so incredibly lucky to have been raised by both of you.

I next want to acknowledge my siblings, Vanessa and Mark. To my big sister, I don't have a single memory of you ever being mad at or negative towards me. I've always looked up to you, your sociality, your intelligence, your work ethic, and how you literally can make friends with anyone. It has been probably the single best experience

of my life watching you succeed our parents as an amazing parent to Carter, Rae, and Ezra. To my little brother, you are literally the original homie. Every foundational memory I have as a kid was by your side and it kills me that we spent so many formative years apart. I am so proud of your self-confidence, your work ethic, your intelligence, and your sense of humor. There is truly no one funnier! The come-up you've had while I've been on my journey has filled me with so much pride and has been such a grounding source of inspiration for me as I've navigated my battles.

I'd like to next acknowledge the educators and mentors who have been instrumental in my journey. Mrs. Truman, you were the first person outside my family to believe in me and push me to expand my knowledge. Mr. Davis, my seventh-grade math teacher, you were the first POC educator I ever had and showed me what it was to represent my community as a professional. To my ninth-grade math teacher DK, you taught me how to recover after failure. My English teacher Joanna, you treated me like an actual person. Dr. Gordon, senior year English, you taught me how to and how not to express myself authentically. You also never let me get away with anything. To Adam, my freshman-year chemistry TA, this is all your fault. You were the first person to put me onto the possibility of a Ph.D., and it is your early inspirations that I went back to when I needed motivation. To the MARC Program at the University of Arizona, specifically Cindy Neal, Megan McEvoy, and Mark Tischler. This is also your fault. I find it highly dubious that I'd have ever gotten into a lab let alone a Ph.D. program were it not for the care you put into my training and journey. To the Kens in my first undergrad lab: you were my first scientific colleagues. You always made time for both training and mentorship. Lauren Cote, you were the best mentor I had as an undergraduate. You

were patient with me, let me be myself, and I credit you with helping me bridge the gap from undergraduate science baby to undergraduate science teenager.

In graduate school, I must first thank my advisors, Bassem and Barbara. Bassem, you gave me the confidence I needed to take the work in new directions. Barbara, you've been an advocate for my individuality since day one. To both of you, I acknowledge, and my parents will affirm, that I can be a handful. You both allowed me the room to grow into myself and gave me the space and support I needed to make it through these past few years. To my advisory team, Abby, Bo, and Jeremy. Abby, there was never a moment where you let me speak poorly of myself. Not only that, but as someone I truly look up to scientifically and professionally, the encouragement you consistently gave me was instrumental in building my confidence. Bo, you've always put me on and have been an advocate for me as a scientist and individual since I rotated. Jeremy, I feel like every time we interact, I gain some new insight that reminds me that everything's going to be okay. To the science and lab homies, I must first and foremost thank Farzad Yousefi. Farzad, you have been so kind to me and worked so hard on the project, I truly don't think I could have done this without you. Eric Martin, you bring such energy and warmth to the lab. Ahmed, you invited me into your home and fed me so we are legally siblings. Nathan, Perla, Can, Kif, Ann, Lena, and Liv, y'all are infinitely better scientists than I am and that was apparent when y'all were techs in the lab. It has been such a joy to watch you all grow. Nick and Rachel, fellow biochemist and genomicist, I look up to both of you scientifically and professionally, and I was lucky to follow in your footsteps. Lucy Brennan, you are probably the smartest person I've ever met and your kindness and generosity towards me are one of the things that give me hope for the

future of academia. To David Brown, my mentor in the Huang lab, you're one of the best dudes I ever met. I'd also like to shout out Harold Marin. It has truly been sick getting to contribute to a scientific story with a fellow Latino scientist. I'd also like to give extra special thanks to Joanna Cai, Sophia Jia, and Daniel Darling. I had the opportunity to mentor the former two scientists over intensive years, as well as Daniel during his rotation. Interacting with and getting to teach you all, and seeing your passion for science, is the highlight of my time as a graduate student at UCSF.

As a highly socially anxious person, I have *battled* the feeling that I don't belong, especially in this space. When dealing with negative emotions, the easiest thing to do is succumb to and dig into them. I would therefore like to extend the sincerest thank you to all the homies I've made, whose love and support have helped me overcome the tendency to dig in and shown me that I do belong. First and foremost, I must shout out Donovan Trinidad. The perfect CoStar compatibility doesn't lie. This man is my best friend, roommate, and brother for real. There's no way I could have done this without your friendship. I'd also like to thank Elise Muñoz. We've grown so much closer as friends since we both made the totally independent unrelated decision to attend UCSF. Kyle Lopez is another dear homie from the University of Arizona days who I must thank. We've known each other since the true gutter days and have shown absolutely no growth. To the homie Santi, you are one of my best friends and I truly always feel warm and fuzzy inside thinking about our parallel experiences. You have been the sanity check I needed. I must also shout out my friend, roommate, and lab mate, Luke Strauskulage. You helped convince me to come to UCSF, showed me my favorite bars, and have been such a positive presence to be around for the last seven years. When I

arrived at UCSF, the first person I ever felt like myself around was Chase Webb. Chase is an essential component of my graduate school experience and there is truly no one else who supports me as I exist in the world as forcefully as he does. Early on, I met Dana Kennedy and Francesca Del Frate, two critical relationships I value to this day. Dana joined the lab at the same time as me and we've been through so much together. Francesca is a Southwest Mexican genomics nerd like me, and she dresses cooler than anyone else I know. They are both incredible people and scientists and I feel lucky to be their friend. To Chris Carlson, Varun Bhadkamkar, Danny Conrad, CJ San Felipe, Niles Robertson, and Liz Bond, my life is better because of knowing you each individually. UCSF is also where I got to meet Ady Steinbach. Ady, you have been there for me like no one else has and we have grown so much together. I love you so very much.

Last but certainly not least, I thank Po Cat. Believe it or not, this little idiot was the mechanism by which I learned that I can form stable attachments. You've been by my side for nine years and I look forward to gaining an immortal form with you and heading into deep space where we'll spend an eternity together amongst the stars.

Contributions

Chapter 2 was a partnership with Farzad Yousefi. Experiments were conducted in the lab of Bassem Al-Sady, with advisement from Barbara Panning. We obtained additional experimental support from Sophia Jia and Daniel Darling. Sam Whedon and Phil Cole at Harvard produced trapping nucleosome substrates. Mike Trnka performed Crosslinking Mass Spectrometry. This work comprises a future manuscript.

Chapter 3 was a collaborative effort led by Harold Marin in Abigail Buchwalter's lab with assistance from Charlie Allen. Harold's initial thesis work focused on identifying the proteins that tethered H3K9me2 to the nuclear periphery in mESCs. My initial thesis work focused on tracking H3K9me2 spreading in mESCs during differentiation using CUT&RUN and microscopy. Following a combination of efforts, I contributed experimental and analytical support on CUT&RUN, and knowledge/analytical support on microscopy analysis, differentiations, and tissue culture. An additional contributor to this work is Eric Martin. This work also comprises a future manuscript.

Recognition of Histone 3 Lysine 9 Dimethylation Regulates Its Production

Eric Simental

Abstract

Specifying cellular identity requires silencing cell-type inappropriate genes. In eukaryotes, this is accomplished by varying the packaging density of nucleosomes, segmenting chromosomes into regions of tight (heterochromatin) and loose (euchromatin) packing. Packaging is directed in part by dimethylation of Histone 3 lysine 9 (H3K9me₂). H3K9me₂ is catalyzed by the G9a-GLP heterodimer, which, like other heterochromatic lysine methyltransferases, possess product-recognition, or 'reading' domains. We first explore how reading by G9a-GLP affects nucleosome methylation *in vitro*. We unveil that a *cis* composition of heterotypic product-recognition domains functions to regulate nucleosome dimethylation but not monomethylation. Our findings illuminate the pivotal role of methyl nucleosome binding by these domains in facilitating dimethylation. We then turn our attention to the cell, where H3K9me₂ is enriched at the nuclear periphery. We find that knocking out putative H3K9me₂ readers, which releases the mark from the nuclear periphery, affects the genomic abundance and distribution of H3K9me₂ specifically during differentiation. These insights underscore the complexity of the methylation landscape and open doors to understanding the interplay of molecular recognition and enzymatic activity in chromatin modification and ultimately, cell identity.

Table of Contents

1: Introduction	1
2: Product recognition by G9a-GLP is required for nucleosome dimethylation	7
2.1 Introduction	8
2.2 Results	10
2.3 Discussion	24
2.4 Methods	29
3: H3K9me2 propagation requires association with the nuclear periphery	33
3.1 Introduction	34
3.2 Results	36
3.3 Discussion	47
3.4 Methods	49
Concluding Thoughts	52
References	55

List of Figures

Figure 2.1: The ANK domains of G9a-GLP are required for nucleosome dimethylation	18
Figure 2.2: Methyl recognition via the ANKs promotes nucleosome dimethylation	19
Figure 2.3: A cis composition of heterotypic ANKs is required for nucleosome dimethylation	20
Figure 2.4: Nucleosome methylation induces ANK-nucleosome contacts	22
Figure 3.1. H3K9me2 CUT&RUN in mESCs	40
Figure 3.2. H3K9me2 maintenance does not depend on its association with the nuclear periphery	41
Figure 3.3. Peripheral association of H3K9me2 limits its production in EpiLCs	43
Figure 3.4. H3K9me2 propagation depends on its association with the nuclear periphery	45

1: Introduction

Life is defined as a quality that distinguishes matter that has biological processes, organisms, from matter that does not. To be an organism is to exhibit five traits: 1) growth 2) reaction to stimuli 3) metabolism 4) energy transformation and 5) reproduction. The individual building blocks of organisms are called cells. Indeed, cells are the smallest individual organisms, and in most cases, they are quite small. Cells are so central to life that most scientists include cellularity as a requirement for being living. Cells are membrane-enclosed droplets of mostly water that house all the biomolecules that enable growth, reaction to stimuli, metabolism, energy transformation, and reproduction. These biomolecules are the building blocks of cells in the same way cells are the building blocks of organisms. As a molecular biologist, I seek to understand how the interactions among and actions of these biomolecules, which themselves are non-living, contribute to the formation of organisms, and more importantly, how we can exploit the actions of these biomolecules to our benefit.

The first major class of these biomolecules are called proteins. Proteins are long chains of chemically diverse amino acids that fold to form 3D shapes based on interactions of these chemical groups with each other and the surrounding environment. These 3D shapes act as tiny machines that perform tasks in the cell that enable the aforementioned traits. These tasks range from breaking down glucose to contracting muscle fibers, to processing light in our eyes. The unique characteristics and functions a cell performs determine what we will refer to as the cell's identity, and are a direct consequence of the suite of proteins a cell makes.

While a cell must express the proper proteins at the proper times, it must also ensure that the wrong proteins aren't made at the wrong times. Cells accomplish this dynamic control, or regulation, of protein production, using the second major class of biomolecules: nucleic acids. In general, the role of nucleic acids is to support the formation of proteins. DNAs, or deoxyribonucleic acids, store the amino acid sequences for all the proteins a cell makes using unique combinations of four nucleic acids: adenine, thymidine, guanine, and cytosine. Each protein is encoded in the DNA as a gene. The sum of an organism's genes is termed the genome. Gene expression, then, is the process of making a protein from DNA and requires two steps. The first step of gene expression is transcription. During transcription, proteins read a sequence of DNA and transcribe it into a corresponding mRNA sequence. mRNA, or ribonucleic acid, is like DNA except thymidine is replaced by uracil. The RNA sequence is translated into a corresponding sequence of amino acids in the second step of protein production, termed translation. A cell therefore maintains its identity through the precise regulation of transcription and translation.

Positive regulation, or gene expression, and negative regulation, gene repression, establish two ends of a wide spectrum of protein expressions. And because cell identity is a direct parallel of the suite of proteins a cell produces, it too, exists on a spectrum. And perhaps this is not surprising. Consider single-celled organisms. They must be able to respond to changes in their environment: heat, pH, light, etc. A yeast switches its mating type to reproduce, enabling future generations to succeed. And there is almost no better example of the dynamic nature of cell identity than our own embryonic development. Sperm and egg meet, and that begins a cascade of events

that involves the rapid and controlled generation of new cell types, a whole organism, from a single progenitor. This thesis is focused on the dynamic nature of cell identity, and more specifically a single process that enables these dynamics: limiting the access of transcriptional machinery to DNA.

Eukaryotic genomes comprise a protein component called histones, an octamer of which wraps ~150 base pairs (bp) of DNA around itself. The genome is punctuated by repeat units of DNA-histone complexes, termed nucleosomes, along its entire length as beads on a string. Each string is called a chromosome. The dominantly discussed role of histone octamers is to package DNA, six meters long in humans, into the 10-micrometer nucleus, the membrane-bound organelle that houses the genome in eukaryotes. But not all packaging is equal. Chromosomes are partitioned into dense or loose packaging in part through the selective modification of nucleosomes. We call these dense and loose regions heterochromatin and euchromatin, respectively. Histones, like all proteins, are chemically modifiable by other proteins especially suited for this task. In general, post-translational modifications (PTMs) modulate a protein's function. The function of histones is to package the genome; thus, modification of histones affects the packaging of the genome. Here, we focus on the post-translational methylation of Histone 3 lysines (H3Kme), which promotes heterochromatin formation and therefore a stronger occlusion of transcriptional machinery from DNA¹⁻⁴. This occurs in part through the recruitment of repressive protein complexes that specifically recognize, or read, H3Kme and help aid in gene repression^{5,6}. So, while H3Kme heterochromatin is not itself gene-repressive, it acts as the bedrock for gene-repressive

activities. As cell identity changes, so too does the genomic pattern of heterochromatin in a process we refer to as heterochromatin spreading⁴.

Heterochromatin assembly and spreading share several conserved features across eukaryotes. Heterochromatin is nucleated at specific DNA sequences through the recruitment of repressive enzymes that catalyze H3Kme⁷⁻⁹. These methyltransferases, or writers, contain or work with H3Kme readers, which bind H3Kme on the nucleosome. If the reader is contained in the writer molecule, the reader binding to the methylated nucleosome places a writer in the proximity of the adjacent, unmethylated nucleosome. The writer is then free to methylate that adjacent nucleosome, and the read-write cycle continues until a domain boundary is reached. This is how heterochromatin spreading is proposed to operate for most if not all forms of heterochromatin^{1,4,10-15}.

Heterochromatic H3Kme comes in three distinct forms. The first class is H3K27 trimethylated (H3K27me₃). H3K27me₃ is catalyzed by the Polycomb Repressive Complex 2 (PRC2), primarily at developmental genes. PRC2 contains the catalytic subunit, enhancer of zeste (Ezh). Ezh belongs to the SET domain-containing family of enzymes, which is conserved among methyltransferases across taxa, so named for the catalytic Su(var)3-9, Enhancer-of-zeste and Trithorax (SET) domain. PRC2 also contains a reader protein, EED¹⁶. EED binding to H3K27me stimulates catalytic activity, enabling heterochromatin spreading¹⁷⁻¹⁹. The second class of heterochromatin is trimethylated at lysine 9 (H3K9me₃). In metazoans, which includes mammals and humans, H3K9me₃ is catalyzed primarily by Setdb1 and Suv39h1/h2, predominantly repressing sequences that might cause genomic instability if expressed²⁰. Like PRC2,

Setdb1 and Suv39h1/h2 belong to the class of SET domain-containing enzymes. Like PRC2, Suv39h1 and Setdb1 also work in conjunction with readers to enable the formation of gene-repressive H3K9me3 domains, albeit through distinct mechanisms from each other²¹⁻²³. In this work, we focus our attention on the poorly understood third class of heterochromatin, H3K9me2.

H3K9me2 is catalyzed by the highly homologous SET domain-containing enzymes G9a and G9a-like protein (GLP)²⁴. G9a and GLP, which can form homodimers, are each independently capable of producing H3K9me1 and H3K9me2 *in vitro*. Despite this, a heterodimer of G9a-GLP is understood to be primarily responsible for H3K9me2 *in vivo*²⁵. G9a-GLP each contain an H3K9me1/2-reader, the Ankyrin repeat (ANK) domain, which contains an aromatic cage responsible for H3K9me1/2-binding²⁶. It would be straightforward to assume that like other repressive methyl-lysine reader-writers, G9a-GLP spread heterochromatin in a nucleosome-to-nucleosome manner which depends on H3K9me1/2 reading by the ANKs. However, H3K9me2 spreading in a nucleosome-to-nucleosome manner by G9a-GLP might not be the most appropriate model. Following differentiation of mouse stem cells, others found that small, pre-existing H3K9me2 domains expanded, in a cell-type specific manner, to encompass large genomic regions in a process that visually mirrors classical spreading²⁷, though this interpretation was later challenged²⁷⁻³². Additionally, newer methods have found conflicting evidence as to whether H3K9me2 spreads in these contexts at all^{31,33}. *In vitro*, ablation of G9a or GLP's product recognition does indeed seem to impede lateral spreading, but the *in vivo* effects of these perturbations do not mirror the spreading mutations and their phenotypes observed in homologous

systems^{34,35}. In summary, it is unclear if and in what contexts H3K9me2 spreads, the mechanism by which it spreads, or whether reading by the ANK domains is responsible for that spreading. Thus, the central question that remains is: how does G9a-GLP-catalyzed H3K9me2 spread?

To address this question, we pursued two complementary goals, centered on the molecular and cellular mechanism of H3K9me2 spreading. The first goal was to understand whether reading by the ANKs affected writing by G9a-GLP *in vitro*. Reader-writer feedback is the central mechanism by which heterochromatin is formed. Understanding whether this mode of regulation is present in a reduced biochemical system would allow us to speculate on *in vivo* spreading mechanisms, but also directly complement the results obtained from our second goal. Our second goal was to resolve whether H3K9me2 spreading occurs in a developmentally relevant context and use this understanding to inform whether H3K9me2 spreads by an internucleosomal propagation mechanism. In the first series of results, we show that readers do indeed shape H3K9me2 formation by writers and speculate as to how. In the second, we confirm that H3K9me2 spreading occurs *in vivo*, and specifically ask how changing the spreading environment affects spreading itself. These works establish that, while H3K9me2 does indeed spread by a product recognition mechanism, it likely does so in modes distinct from other lysine methyltransferases.

2: Product recognition by G9a-GLP is required for nucleosome dimethylation

Abstract

Heterochromatin formation is directed in part by dimethylation of Histone 3 lysine 9 (H3K9me₂), catalyzed by the G9a-GLP heterodimer.²⁵ Like other heterochromatic lysine methyltransferases, G9a-GLP possesses product-recognition, or 'reading' domains²⁶ Here, we explore how reading by G9a-GLP's ANK domains affects nucleosome methylation *in vitro*. We unveil that while the ANK domains are not required for monomethylation, they are indispensable for nucleosome dimethylation. Additionally, we discern the balance in their contribution to nucleosome dimethylation by swapping ANK domains between G9a and GLP, highlighting a heterotypic arrangement essential for efficient dimethylation. Our findings also illuminate the pivotal role of methyl-binding by the ANKs in facilitating dimethylation. We then demonstrate that nucleosome methylation specifies ANK engagement with the nucleosome, preferentially with G9a, and that abrogating these contacts removes the capacity of the heterodimer to produce nucleosome dimethylation. These insights underscore the complexity of the methylation landscape and open doors to understanding the interplay of molecular recognition and enzymatic activity in chromatin modification. In this work, we establish a clear link between methyl recognition via the ANK domains and dimethylation specifically. This study deepens our understanding of chromatin dynamics and opens new avenues for therapeutic interventions targeting diseases associated with dysregulated cell identity.

2.1 Introduction

Cell identity is hypothesized to be controlled in part through the formation of gene-repressive H3K9me2 heterochromatin domains. H3K9me2 is predominantly catalyzed *in vivo* by the euchromatic histone methyltransferases 2 and 1, or G9a and G9a-like protein (GLP), respectively^{25,36}. Both G9a and GLP contain an H3K9me1/2 catalysis domain (SET), and an H3K9me1/2 recognition domain, the Ankyrin repeat domain (ANK)²⁶. The N-termini of G9a and GLP contain methylation sites and protein-protein interaction domains that may facilitate G9a-GLP localization on chromatin^{34,36,37}.

The SET of G9a and GLP, like Suv39h1/h2, catalyzes H3K9 methylation, predominantly H3K9me1/2, but can perform H3K9me3 and H3K27me1/2. Like PRC2, Setdb1, and Suv39h1/h2, G9a and GLP also work in conjunction with domains that recognize the primary products of their respective enzyme's catalysis. In each of these enzymes, methyl recognition promotes heterochromatin formation through distinct mechanisms^{18,19,38}. Thus, the coupling of recognition with catalysis, read with write, is central to heterochromatin formation, and consequently, cell identity.

G9a and GLP are unique among heterochromatic lysine methyltransferases in that G9a and GLP act as an obligate heterodimer in cells^{25,36}. G9a and GLP both form stable homodimers or heterodimers primarily through SET/SET contacts. Cellular knock-out of either G9a or GLP results in a near loss of H3K9me2, and a less dramatic loss of H3K9me1. Thus, the G9a-GLP heterodimer is the dominant cellular H3K9 dimethylase. We previously demonstrated that both reading and writing are significantly enhanced in the heterodimer compared to homodimers of G9a or GLP. The heterodimer has higher recognition of H3K9me2, and a 10-fold increased turnover rate for

nucleosomal substrates under multiple turnover conditions, which is not evident on histone tail peptide substrates³⁹. Given the coupling of read and write in homologous systems, we wondered whether the heterodimer's increased ability to read affected its' ability to write, or vice versa. We hypothesized that like other chromatin-modifying enzymes, catalysis might be regulated by methyl-reading in the heterodimer. Here, we bacterially express wild-type and mutant G9a-GLP heterodimers and query the effect altered engagement with a nucleosome substrate has on catalysis. We demonstrate a requirement of product recognition specifically for catalysis of the biologically relevant H3K9me2 modification. This work is among the first we are aware of which performs mutational analysis on the heterodimer, the dominant cellular form, with its native substrate, the nucleosome.

2.2 Results

The ANK domains are required for nucleosome dimethylation by G9a-GLP

We previously demonstrated enhanced nucleosome methylation by the G9a-GLP heterodimer compared to the G9a homodimer *in vitro*, an activity coupled to enhanced methyl-peptide recognition. In other heterochromatic lysine methyltransferases, product recognition stimulates methylation; we wondered whether this level of regulation was present within the G9a-GLP heterodimer.

To query G9a-GLP activity on its native substrate, we assay nucleosome methylation over two hours following bacterial purification of truncated (**Fig. 2.1A**) G9a-GLP ANK-SET heterodimers at 1:1 stoichiometry. Given that we sought to modulate product recognition directly, we chose a subsaturating regime of nucleosomes with respect to the Michaelis-Menten curve we previously reported, to sensitize the system to binding differences by the heterodimer (**Fig. 2.1B**).

To understand how reading by the ANKs affects writing by the SETs, we first removed the ANK domains from our ANK-SET heterodimer, producing a SET-only heterodimer. We performed methylation reactions in parallel with WT ANK-SET heterodimer and found that loss of both G9a and GLP's ANK domains does not affect nucleosome monomethylation but results in a near complete loss of dimethylation (**Fig. 2.1C-D**). As there is no precedent for a methyl-state-specific requirement for product recognition, we wondered whether the loss of the ANK domains altered the intrinsic activity of the SETs.

To test this possibility, we assayed methyltransferase activity on H3 tail peptide substrates. To gain methyl-state specificity, we monitored methylation activity via Matrix-

Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF), under saturating conditions. From these peptide activity assays, it is clear the SET-only heterodimer can produce dimethylation on an H3(1-20) tail peptide (**Fig. 2.1E**) but is incapable of doing so on a nucleosome. This demonstrates an absolute requirement of ANK domains for dimethylation, but not monomethylation, on a nucleosome.

Methyl recognition via the ANK domains promotes nucleosome dimethylation

The central ascribed biochemical activity of the G9a and GLP ANK domains *in vitro* is the recognition of methyl tail peptide²⁶. In G9a and GLP homodimers, H3K9me1/2 recognition promotes internucleosome methylation *in vitro* and dimethylation *in vivo*, while loss of G9a ANK's methyl-binding activity is dispensable for nucleosome dimethylation *in vivo*³⁴. However, the role of methyl binding by the ANKs in 1) mononucleosome methylation, 2) mono- vs dimethylation, or 3) methylation in the heterodimeric context, remains unexplored. Given that the ANK domains are required to convert H3K9me1 to me2, we hypothesized that methyl-tail recognition, the dominant *in vitro* activity of the ANKs, is responsible for this conversion.

To test the hypothesis that methyl recognition is required for nucleosome dimethylation, we employed Tryptophan to Alanine mutations in the aromatic cages of the ANK domains (G9a 3A or GLP 3A) that abrogate methyl reading by the ANKs. We installed 3A mutations alone or in combination within an ANK-SET heterodimer. First, we sought to dissect whether the ANKs individually or together contributed to the binding of methylated nucleosomes. Given the poor binding of G9a-GLP to nucleosomes, we used specially modified H3K9me2 nucleosomes in which one tail is modified with the enzyme trapping Norleucine amino acid (Nle), which we refer to as an

asymmetric nucleosome. Via Electrophoretic Mobility Shift Assay (EMSA), we observed complex formation following incubation of 6 μ M ANK-SET heterodimer with 500 nM nucleosome (**Fig. 2.2A-B**). This finding encouraged us to query 3A mutant binding on the asymmetric nucleosome. We found that binding of the GLP 3A heterodimer to the asymmetrically modified nucleosome is mildly reduced compared to WT, while both the single G9a 3A and double G9a 3A/GLP 3A point mutant exhibit a near complete loss of binding (**Fig. 2.2A-B**). We additionally explored a range of enzyme concentrations and confirmed these observations, with the GLP 3A single point mutant displaying further reduced binding compared to WT (data not shown). These results indicate that the aromatic cages of both G9a and GLP's ANK domains contribute to the binding of a methylated nucleosome within the heterodimeric context.

To test the hypothesis that methyl recognition is required for nucleosome dimethylation, we performed parallel methylation reactions with the 3A mutants and WT. Loss of both aromatic cages within the ANK-SET heterodimer resulted in an 80% loss of nucleosome dimethylation and a small amount of excess monomethyl accumulation (**Fig. 2.2C-D**). This indicates an absolute requirement of methyl-binding by the ANKs for efficient dimethylation. Given the asymmetric contribution of G9a and GLP's ANK domains to nucleosome recognition, we wondered whether there was a similar asymmetry in the contribution from methyl-reading by G9a and GLP's ANKs. Our experiments indicate that loss of methyl-reading by G9a or GLP's ANK is dispensable for nucleosome mono- and dimethylation. Interestingly, the combined effect of losing methyl-reading in both ANKs was non-additive, but synergistic with respect to dimethylation.

A *cis* composition of heterotypic ANKs is required for nucleosome dimethylation

The major reported role for the ANK domains *in vitro* is methyl tail reading. Were the sole role of the ANK domains methyl tail reading, we would expect the double 3A ANK-SET heterodimer to display a loss of methylation as in the SET-only heterodimer. However, the double 3A ANK-SET heterodimer retains ~20% dimethylation as compared to the WT ANK-SET heterodimer (**Fig 2.2C-D**). The retention of modest dimethylation activity in the double 3A point mutant suggests that non-aromatic cage residues on the ANK domain might contribute to nucleosome dimethylation through an unknown mechanism. Further, if the sole role of the ANK domains is methyl recognition, then the ANKs should be exchangeable between G9a and GLP given the similarity in the two ANK domains structurally, in sequence, and with respect to the aromatic cages.

To test whether the ANK domains were indeed equivalent, we tested all combinations of G9a and GLP's ANK domains within the ANK-SET heterodimer. To design the replacement of one ANK with another, we integrated crystallographic structures of isolated SET and ANK domains of G9a and GLP, the amino acid sequence homology between the two, as well as our previous Crosslinking Mass Spectrometry (CLMS) data of intramolecular G9a-GLP contacts. Integrating this information, we excised GLP's ANK 1-8, leaving intact the N-terminus, ANK-SET linker, and SET domains, and replaced this region with the homologous residues of G9a's ANK, which includes regions previously annotated as ANK-SET linker (**Fig. 2.3A**). We replace either G9a's ANK with GLP's, here termed G9aC-GLP, or GLP's ANK with G9a's, designated G9a-GLPC. Replacement of GLP's ANK domain with G9a's did not appreciably affect nucleosome monomethylation but did result in a 20% decrease in dimethylation (**Fig.**

2.3B-C). Thus, G9a's ANK domain is partially able to complement a loss of GLP's ANK. In contrast, replacing G9a's ANK with GLP's results in the loss of more than 70% dimethylation. A decrease in the dimethylation rate is again accompanied by maintenance of the monomethylation rate (**Fig. 2.3B-C**). This indicates that GLP's ANK is less able to complement the loss of G9a's. Given the contribution of both ANKs to dimethylation, we next asked whether the ANKs functioned *in cis* or whether they could serve their dimethylation-enhancement function *in trans*. A swap of ANK domains within the heterodimer, termed G9aC-GLPC, phenocopied the full ANK deletion seen in the SET-only heterodimers (**Fig. 2.3D-E**). That is, nucleosome monomethylation was unaffected while dimethylation was severely reduced. We conclude that a *cis* arrangement of heterotypic ANKs is required for nucleosome dimethylation.

Replacement of domains could have altered structures that affect activity. To understand the effects of replacing ANK domains we used AlphaFold2-multimer to model the structures of WT and chimeric enzymes. Modeling confirmed the WT ANK-SET heterodimer's structure, aligning well with existing crystal structures. Despite AlphaFold2's limitations in domain orientation prediction, our previous CLMS data supported the accuracy of this model. Chimeric models showed minor deviations without significant structural changes from the WT, whereas the G9a homodimer displayed a different domain arrangement, consistent with our previous CLMS data. These findings suggest that despite structural similarities, altering the connection between ANK domains and SETs disrupts nucleosome dimethylation, highlighting the importance of their heterotypic composition. These results suggest that both ANKs differentially contribute to nucleosome dimethylation, with G9a's ANK in particular being

nearly sufficient. We conclude that a heterotypic composition of ANK domains is necessary for nucleosome dimethylation and posit that this is without major structural rearrangement as compared to the ANK-SET heterodimer.

Nucleosome methylation induces ANK-nucleosome contacts

A *cis* arrangement of heterotypic ANKs is required to accomplish the me1 to me2 transition on nucleosomes (**Fig. 2.3**). That the ANKs are not required for monomethylation of an unmodified nucleosome suggests the ANKs do not engage with an unmodified nucleosome. In contrast, that the ANKs are absolutely required for dimethylation suggests they may be involved in recognizing or otherwise engaging with a methylated nucleosome, in contrast to an unmethylated nucleosome.

To test this hypothesis, we performed CLMS of enzyme bound to unmodified or H3K9me2-modified nucleosomes. As we were unable to obtain stable complex formation as assessed by EMSA (data not shown), and the K_d for unmodified nucleosomes is relatively weak, we utilize SET domain-trapping modifications, converting the catalytic lysine target to an aliphatic residue, either ethyl cysteine (Ecx) or norleucine (Nle) (**Fig. 2.4A**). To compare the impact of pre-existing H3K9 methylation on Enzyme-substrate complex formation, we contrast the association of G9a-GLP with a symmetric H3K9Ecx nucleosome, where both H3 tails contain a catalytic trapping H3K9Ecx, with asymmetric nucleosomes where one tail is H3K9me2 and the other a trapping H3K9Nle. First, we compare the overall binding of G9a-GLP to these nucleosomes and find that G9a-GLP has a clear preference for the asymmetric nucleosome, even though the H3K9Ecx symmetric nucleosomes contain two trapping residues (**Fig. 2.4B**). Given that the binding of G9a to the nucleosomes is dominated by

the catalytic SET, a base-level expectation would be a stronger association with two trapping tails. However, the asymmetric nucleosomes formed a much tighter complex with G9a-GLP. This result may indicate that this nucleosome represents a favorable intermediate enzyme-substrate complex. We thus analyzed complex formation by CLMS, i.e. first crosslinking with DSSO, protease digestion, and Mass-Spectrometry.

We first compared captured intranucleosome crosslinks with known contacts derived from structural studies and found satisfactory agreement. When analyzing Ecx nucleosome-G9a-GLP contacts, we observed that contacts occurred exclusively through the H3 tail and the SET domains of G9a and GLP (**Fig. 2.4C, E**). That is, we did not detect ANK-Ecx nucleosome contacts, nor contacts to the nucleosome core particle. These data suggest that SET binding to the H3 tail is the primary mode of engagement with an otherwise unmodified nucleosome. In stark contrast, the asymmetric H3K9me2/H3K9Nle nucleosome revealed extensive crosslinks with the ANK domains of G9a and GLP, and, particularly along the H3 tail, with more G9a crosslinks captured than GLP crosslinks (**Fig. 2.4D, F**). Of note, we also detected interactions with the N-termini of G9a and GLP and the C-terminus of H2A, which is structurally positioned near the base of the H3 tails. These results indicate that nucleosome methylation induces ANK-nucleosome contacts, as well as induces more extensive engagement with G9a-GLP. We speculate that these contacts are necessary for nucleosome dimethylation.

G9a ANK8 is essential for nucleosome dimethylation

The above data suggest that the specific contacts observed between the ANKs of the WT ANK-SET heterodimer and the methylated nucleosome are required for effective H3K9me2 production. Like our chimera data, our CLMS data also suggest the

possibility that G9a and GLP's ANKs do not equally contribute to dimethylation. CLMS data from the ANK-SET heterodimer alone revealed a crosslink between a previously unannotated region of G9a, homologous to GLP's ANK8, and ANK 2 of GLP. This same region of G9a ANK8 makes contact with the H3 tail of the methylated nucleosome but does not contact the H3K9Ecx nucleosome. The captured peptide is adjacent to the most poorly conserved stretch of residues between G9a and GLP in the ANK-SET construct: AWDLTPER in G9a. This divergence is represented in mouse G9a-GLP. We hypothesized that this poorly conserved sequence might be important not just for intramolecular contacts within the heterodimer but for nucleosome engagement and catalysis as well. To test this hypothesis, we replaced the divergent sequence, AWDLTPER on G9a with a GGGSGGG linker, within the ANK-SET heterodimer, and performed methyltransferase assays.

Preliminarily, the replacement of AWDLTPER on G9a within the ANK-SET context resulted in a complete failure to produce nucleosome dimethylation (data not shown). We emphasize that while we do not yet have a defined monomethyl rate as compared to the WT ANK-SET heterodimer, dimethylation is severely impacted. In the entire study, we have replaced entire domains or removed their key function, tail recognition. None of these perturbations have had as severe an effect on dimethylation as the replacement of these previously undescribed 8 residues within G9a ANK8. We speculate that binding to the methylated nucleosome may be mediated in part through these residues, and that this binding is necessary for dimethylation.

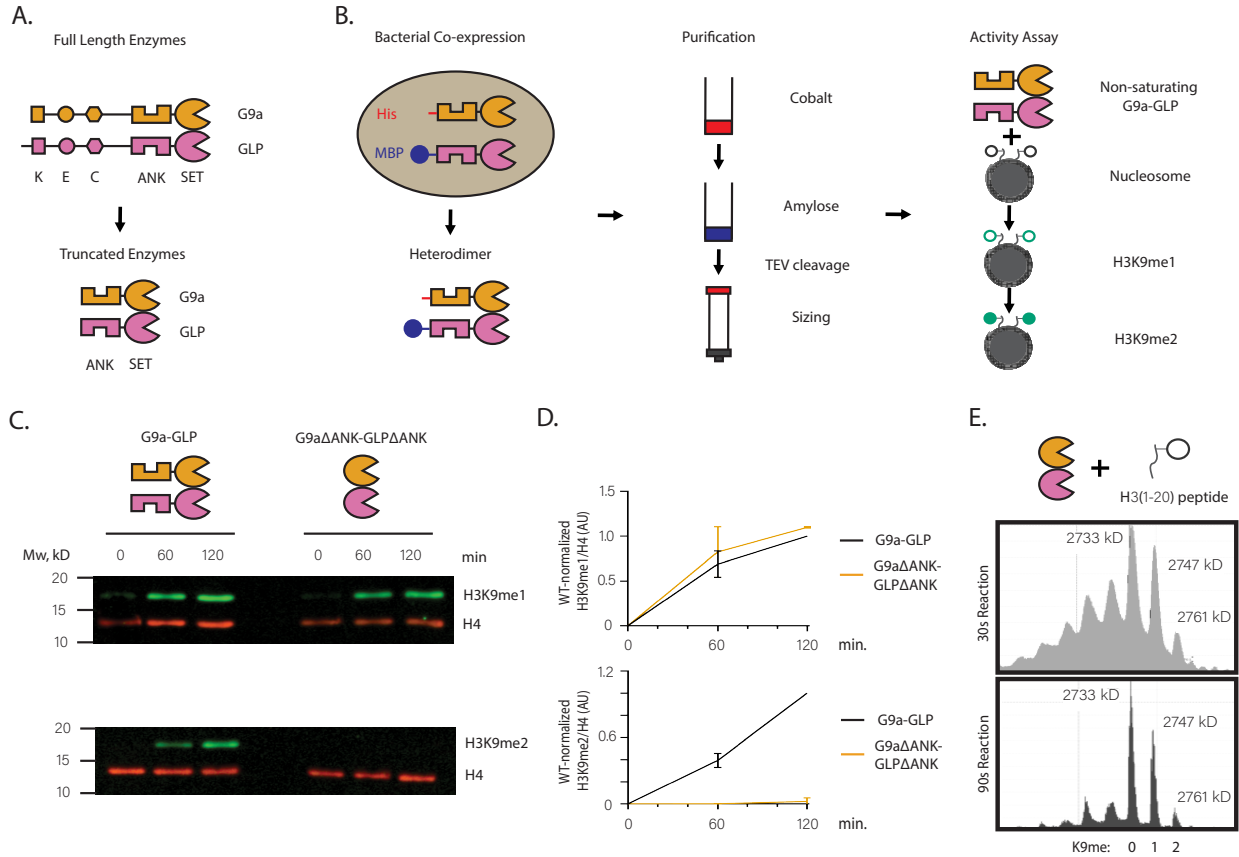


Figure 2.1: The ANK domains of G9a-GLP are required for nucleosome dimethylation

(A) Domain architecture of G9a-GLP. *Top*, full-length enzymes. The N-terminus of G9a and GLP features an automethylation residue (K), an acidic patch (E), and a cysteine-rich region (C). The C-termini contain ankyrin repeats (ANK) that bind H3K9me and a catalytic SET domain (SET) which is the dimerization interface. *Bottom*, truncated ANK-SET construct used in this study. (B) Purification of heterodimer. Constructs are expressed from a single plasmid. Sequential purification is followed by TEV cleavage and sizing and produces heterodimers at 1:1 stoichiometry. 0.8 μ M SETs are incubated with 1 μ M nucleosome for two hours and quenched by Laemli. (C) Western blot time course of H3K9me1 or H3K9me2 (green) production on mononucleosomes comparing WT ANK-SET heterodimer and SET-only heterodimer. H4 is blotted as an internal control (red). (D) Quantification of Western Blot time course. H4-normalized product at each time-point is normalized to WT at 2 hours. (E) Time course MALDI-TOF of H3(1-20) peptide reacted by SET-only heterodimer. Error bars denote standard error from 2 independent duplicate experiments.

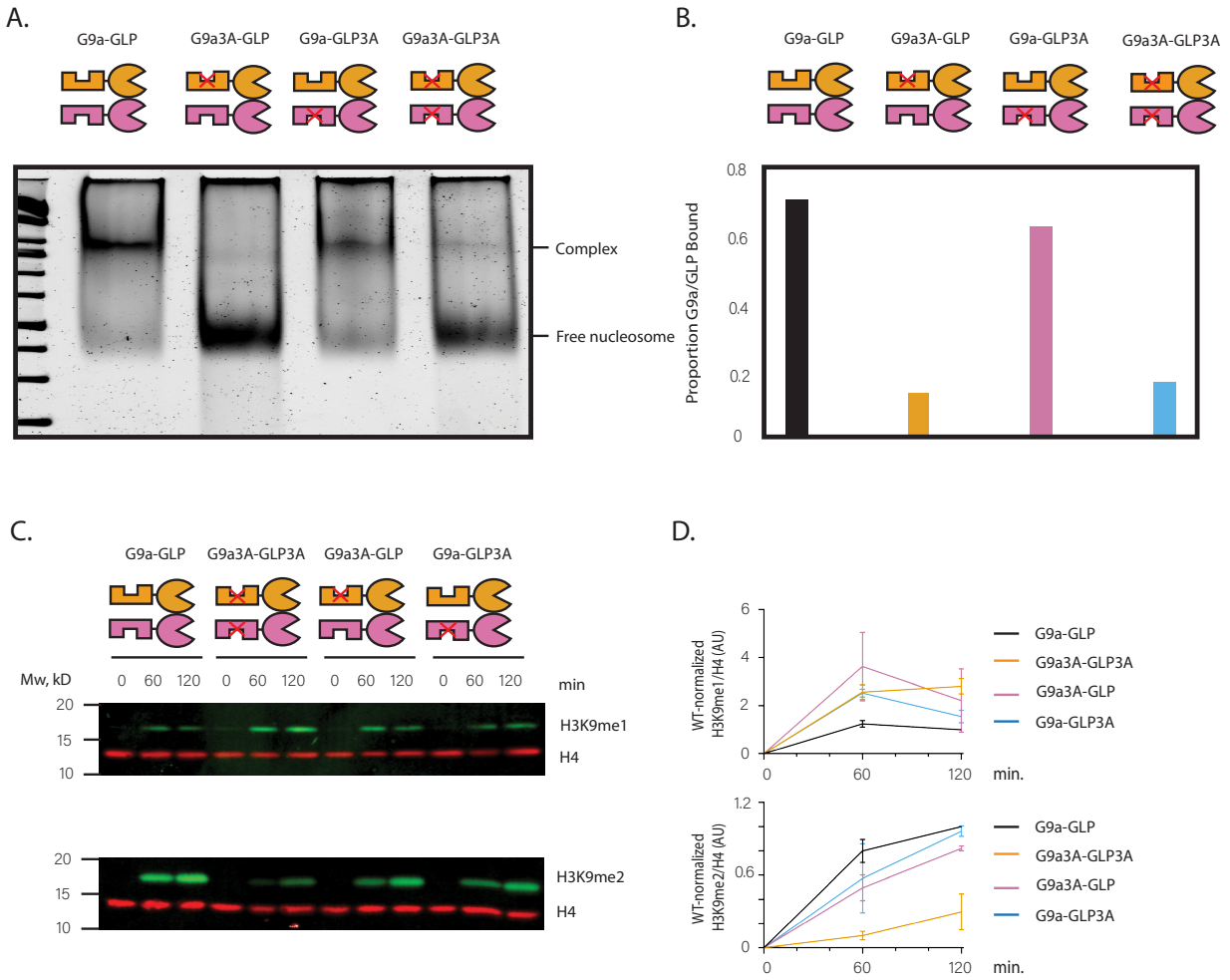


Figure 2.2: Methyl recognition via the ANKs promotes nucleosome dimethylation

(A) Nucleosome EMSA measuring binding of 6 μ M 3A mutants to 200 nM mononucleosome and (B) quantification of the proportion of G9a-GLP bound to the nucleosome. (C) Western Blot time course of mononucleosome methylation by 3A mutants and (D) quantification. Error bars denote standard error from 2 independent duplicate experiments.

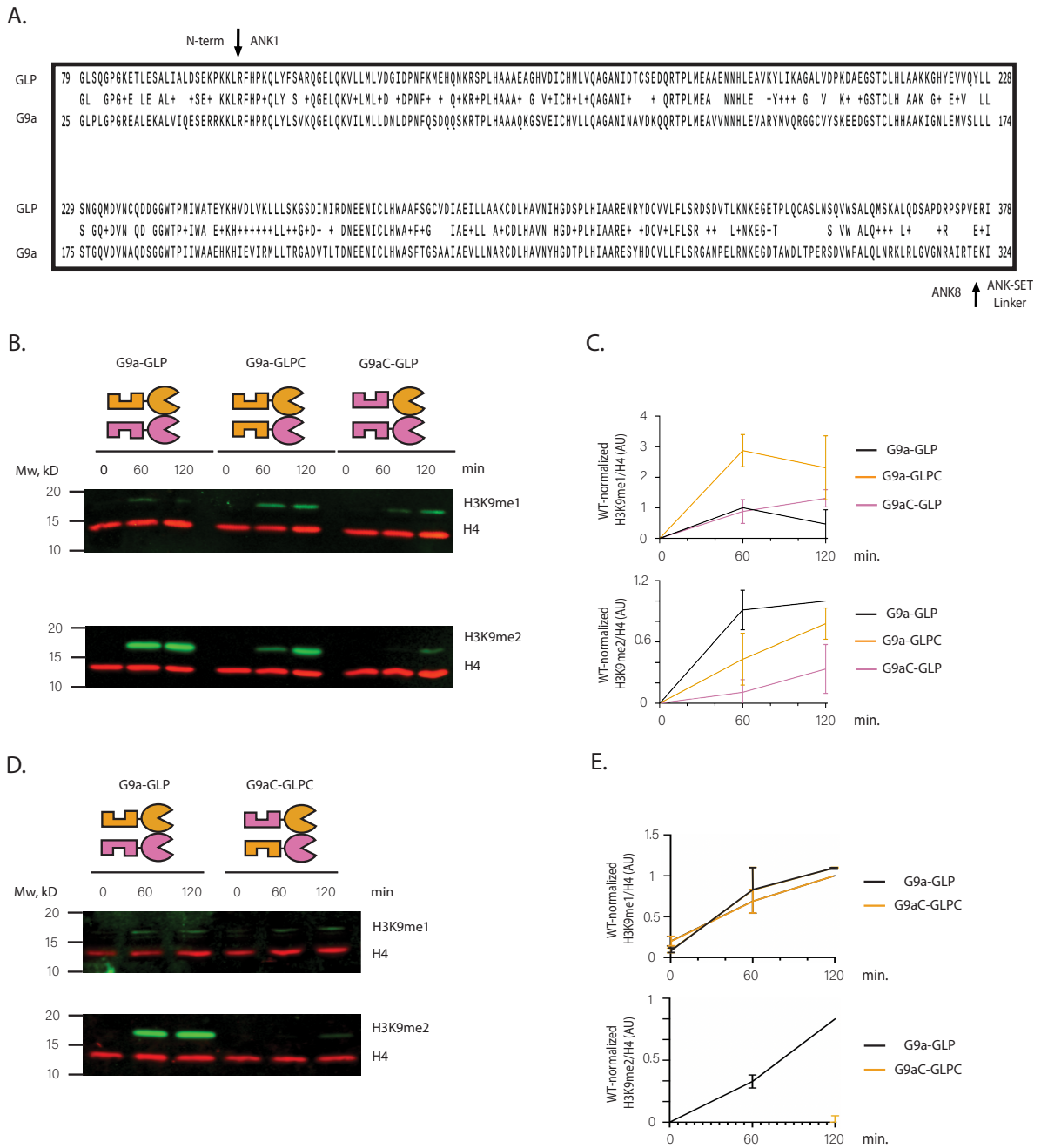


Figure 2.3: A cis composition of heterotypic ANKs is required for nucleosome dimethylation

(A) G9a and GLP amino acid sequence alignment with arrows indicating at which positions each respective ANK was (*Figure caption continued on the next page*)

(*Figure caption continued from the previous page*) excised for replacement with corresponding homologous ANK region. (B, D) Western Blot time course of mononucleosome methylation by (B) G9a-GLPC and G9aC-GLP or (D) G9aC-GLPC mutants and corresponding (C, E) quantifications. Error bars denote standard error from 2 independent duplicate experiments.

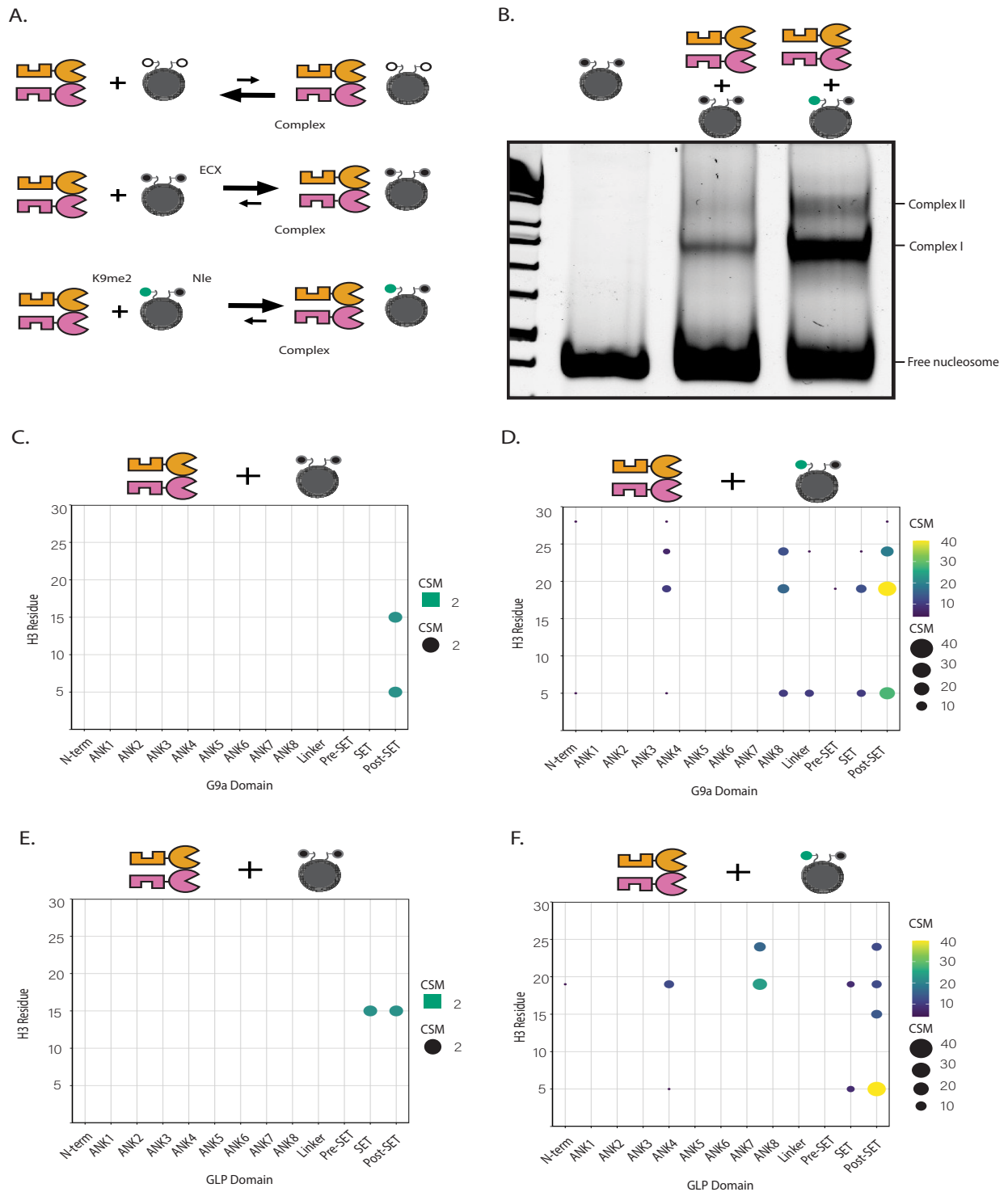


Figure 2.4: Nucleosome methylation induces ANK-nucleosome contacts

(A) WT ANK-SET heterodimer binding to trapping nucleosomes promotes binding compared to unmodified nucleosomes (*Figure caption continued on the next page*)

(Figure caption continued from the previous page) (B) Nucleosome EMSA showing binding of WT ANK-SET heterodimer to either H3K9Ecx or H3K9me2/H3K9Nle (asymmetric) mononucleosome. Complexed products were used for CLMS. (C-F) Cross-linking Mass Spectrometry for WT-ANK set heterodimer to trapping nucleosomes showing G9a/Ecx nucleosome (C), G9a/Asymmetric nucleosome (D), GLP/Ecx nucleosome (E) and GLP/Asymmetric nucleosome contacts (D). Dot size and color indicate the number of cross-linked spectral matches (CSM) summarized for domain pairs.

2.3 Discussion

Unique among repressive lysine methyltransferases, G9a and GLP require heterodimerization for effective H3K9me2 production in cells^{25,36}. We previously demonstrated that the G9a-GLP heterodimer, which contains both product-recognizing ANKs and catalytic SETs, is enhanced in both product recognition and catalysis compared to homodimers of either protein³⁹. In this study, we sought to understand if, like all repressive lysine methyltransferases, product recognition and catalysis are coupled in the heterodimer.

We show that removal or replacement of the *cis* heterotypic ANK unit on an otherwise WT ANK-SET heterodimer compromised dimethyl production without affecting monomethyl production on the nucleosome. Additionally, abrogation of the aromatic cages or ANK8 of G9a, which we found to be involved in forming contacts with a methylated nucleosome, similarly resulted in compromised dimethyl production. Thus, the primary finding of this work is that product recognition via a *cis* arrangement of heterotypic ANKs is required for nucleosome dimethylation but not monomethylation.

Put simply, we've demonstrated that product recognition enhances catalysis by the G9a-GLP heterodimer. Coupling catalytic enhancement to product recognition is a fundamental feature of repressive lysine methyltransferases^{18,21,38}. The most well-characterized mechanism for such regulation is PRC2. H3K27me3-binding to the EED reader's aromatic cage *in cis* or *trans* leads to an allosteric rearrangement within the writer that stabilizes the active site, promoting a general methylation enhancement^{18,19,40}. Like PRC2, Suv39h1-catalyzed H3K9me3 is enabled by a reader, the chromodomain (CD). Unlike PRC2, the CD is part of the same molecule as the

writer. Suv39h1, inactive in its free form, samples chromatin through its CD. CD recognition of H3K9me3 *in cis* or *trans* allosterically activates a latent chromatin binding motif to anchor the enzyme to regions of pre-existing H3K9me3. This second step in turn stimulates H3K9 methylation generally ^{20–22}. In contrast, for the homolog of Suv39h1, Clr4, product-recognition-driven catalytic enhancement is hypothesized to be controlled by a product guidance mechanism. CD binding to the methylated substrate *in cis* guides the Clr4 active site to the proper orientation with respect to the substrate, resulting, again, in a general catalytic enhancement ³⁸. We contrast the general catalytic enhancement seen in PRC2, Suv39h1, and Clr4, with the specific enhancement in dimethylation and not monomethylation. We propose the following mechanism for this specific enhancement.

G9a's SET domain is required to bind unmodified nucleosomes in the WT ANK-SET heterodimer³⁹, suggesting that G9a's SET domain makes first contact with the H3 tail of the nucleosome. On the Ecx-nucleosome, our CLMS shows contacts between both G9a and GLP's SET domains with the H3 tail. Thus, concomitant or following the first methylation event by G9a's SET, GLP may perform the second monomethylation on the second tail. A second possibility is that the G9a SET unbinds from the monomethylated product, binds the unmethylated second tail, and performs a second monomethylation. To continue, our methyltransferase activity assays, as well as failure to capture ANK-nucleosome contacts on an Ecx-modified nucleosome, indicate monomethylation does not involve the ANKs or ANK-nucleosome engagement.

One outstanding question in the field has been whether G9a-GLP accomplishes the H3K9me1 to me2 transition on nucleosomes without substrate disassociation. Were

the SETs capable of processive methylation, we would not expect a loss of the ANKs to affect catalysis. We therefore speculate that following nucleosome monomethylation, the heterodimer disassociates from the monomethylated substrate, in contrast to its proposed processivity on H3 tail substrates. Next, given the presence of two active aromatic cages, we hypothesize that the ANKs bind H3K9me1 nucleosome tail. Because non-aromatic cage-H3 tail contacts are only captured in the presence of methylated substrate, we believe aromatic cage engagement with methyl-peptide precedes and is required for the broader ANK engagement we capture with the tail. If these contacts occurred absent aromatic cage binding to methylated tail, we would have expected to capture them in our CLMS with Ecx-modified nucleosome. Why then, does G9a-GLP require the ANKs to dimethylate a monomethylated nucleosome but not monomethylate an unmodified nucleosome? We believe the strict requirement of ANK-methyl tail contacts is to facilitate proper substrate engagement. We speculate that the methylated nucleosome is a poor substrate for the SET domains. This would explain the specific enhancement seen in the dimethylation step. If recognition is necessary for only the terminal methylation step, we predict a K_M difference between unmodified and modified substrates.

While we believe overcoming poor substrate engagement explains the requirement of the ANKs for nucleosome dimethylation, we acknowledge one non-mutually exclusive alternative: product recognition or production induces an ANK-regulated conformational change within the enzyme that enables efficient dimethylation. From our own CLMS data, all the intra-heterodimer crosslinks we captured on the Ecx nucleosome were represented on the asymmetrically modified nucleosome. However,

we captured several novel/strengthened intra-heterodimer contacts specifically with the asymmetrically modified nucleosome. These data provide compelling, albeit circumstantial evidence that a conformational change within the heterodimer does indeed occur when encountering a methylated mononucleosome. Whether this conformational change occurs or is relevant to the observed biology is unresolved.

The cellular consequences of this hypothesized regulation remain open to speculation. For all repressive lysine methyltransferases, product recognition recruits enzymes to regions of preexisting modification, enabling their maintenance. Additionally, because catalytic stimulation occurs via modification on neighboring nucleosomes, this allows neighboring, unmodified nucleosomes to become methylated, providing a mechanism for nucleosome-to-nucleosome heterochromatin spreading over short and long distances. We emphasize that the regulation uncovered here within the G9a-GLP heterodimer is restricted to the mononucleosome. Like PRC2, Suv39h1, and Clr4, we expect product recognition to guide the heterodimer to regions of pre-existing H3K9me1/2. However, how the growth and expansion of domains occurs following nucleation, or *de novo* remains undescribed. Following recruitment to sites of pre-existing modification, it is possible that the N-terminus of G9a-GLP, which is truncated in our study, binds other DNA-interacting proteins that may guide it along a domain. Indeed, a host of cellular G9a-GLP interactors including Wiz, Dnmt1, and Cyclin D1, may function to spread G9a-GLP along chromatin or recruit it to *de novo* domains^{37,41,42}. The influence of these and other factors on H3K9me2 spreading remains to be explored.

Additionally, our data reconcile prior observations about the formation of H3K9me1 and me2 by different complements of G9a or GLP enzymes *in vivo* with their ability to perform H3K9me1 and me2 *in vitro*. In cells, loss of G9a or GLP results in a near total loss of H3K9me2 without affecting H3K9me1²⁴. *In vitro*, an ANK-SET G9a homodimer is unimpaired for H3K9me1 production but produces much less H3K9me2 compared to an ANK-SET heterodimer. Together, these findings suggest that G9a or GLP homodimers might be sufficient to generate H3K9me1 when either partner is knocked out. In line with this, were both G9a and/or GLP homodimers active, and responsible, for H3K9me1 in cells, knockout of both would result in a loss of H3K9me1, which is indeed the case. Our data, in combination with the cellular knock-out phenotypes, suggest that the specific function of the WT ANK-SET heterodimer is H3K9me2 production on chromatin.

In cells, and in contrast to H3K9me1, H3K9me2 is dominantly associated with the repression of repetitive and transposable elements and euchromatic cell-type inappropriate genes^{33,43-48}. The stringent control of H3K9me2 production on mononucleosomes uncovered in this study might hint at the regulatory mechanisms underlying the control of H3K9me2's spatiotemporal dynamics during cell state transitions.

2.4 Methods

Purification of G9a-GLP Heterodimers

To isolate the ANK-SET G9a-GLP heterodimer, we coexpressed N-terminally tagged His-G9a and MBP-GLP from a single plasmid (QB3 Berkeley Macrolab expression vectors) in *E. coli* DE3 Rosetta cells and performed sequential cobalt- and amylose-charged resin affinity chromatography purification. Cells co-expressing His and MBP constructs were lysed on ice via sonication in lysis buffer (100mM Tris pH 8, 300mM NaCl, 10% glycerol (v/v), 0.1% Tween-20, with freshly added 1mM β -mercaptoethanol (BME), 1mM PMSF, 5mM Benzamidine, 200 μ M Leupeptin, Aprotinin, Pepstatin, Phenantroline). The clarified lysate was then bound to cobalt-charged resin (Takara) for 1hr and washed twice with lysis buffer. His tagged proteins were eluted with lysis buffer containing 400mM imidazole and bound immediately to amylose resin (NEB) for 1hr. MBP-tagged proteins were eluted with lysis buffer + 20 mM maltose. Affinity tags were then removed by incubation with 12mg TEV protease for 1hr at 25°C. TEV protease was absorbed into cobalt resin and the cleaved heterodimer was further purified by size exclusion chromatography (Superdex 200 Increase 10/300 column) and buffer exchanged into storage buffer (100mM Tris pH 8, 100mM KCl, 10% glycerol, 1mM MgCl₂, 20mM ZnSO₄, 10mM BME). All protein constructs were quantified using an SDS page with BSA standards and a Sypro Red stain.

Preparation of mononucleosome substrates

Histone H3.1 peptides (amino acids 1-34) were synthesized using Fmoc-based solid-phase peptide synthesis, incorporating either Fmoc-K(me₂)-OH or Fmoc-Nle at the H3K9 position. After synthesis, the peptides were cleaved from the solid phase,

purified using reverse-phase HPLC (RP-HPLC) with a gradient of acetonitrile in water, and their purity was confirmed via electrospray ionization mass spectrometry (ESI-MS). Quantification was done by analytical RP-HPLC. Concurrently, histone octamer components were co-expressed in LOBSTR Rosetta E. coli, harboring plasmids for histones H2A, H2B, H4, and a truncated version of H3. Post-induction and harvesting, the cells underwent lysis, and histones were purified via affinity chromatography on NiNTA agarose, followed by cleavage with TEV protease and further purification through size-exclusion chromatography.

Separately, Widom 601 DNA sequences of 147 and 185 base pairs were generated in E. coli DH5alpha, extracted through a series of lysis, precipitation, and purification steps, including RNase treatment to remove RNA, and then cut with EcoRV for insert release. The DNA was purified over anion-exchange columns and prepared for nucleosome reconstitution.

For nucleosome assembly, purified histone octamers and Widom 601 DNA were mixed in a reconstitution buffer and subjected to dialysis under a linear salt gradient to facilitate nucleosome formation. These crude nucleosomes were then incubated with synthetic H3.1 K9Nle peptides and a specific sortase mutant for selective modification, followed by purification via weak anion exchange chromatography. The process was repeated with H3.1 K9me2 peptides for further modification. Intermediate and final products were identified and analyzed using western blotting for H3 C-terminus and LC-MS, confirming the incorporation of the desired histone modifications. This detailed methodological approach allowed for the precise synthesis and assembly of modified nucleosomes for further biochemical and structural analyses.

Nucleosome methylation assays

Substrate nucleosomes were mixed in a solution containing 500 mM S-Adenosyl Methionine (SAM, Perkin Elmer) cofactor, and reactions were initiated upon the addition of the enzyme. Reactions were run in 100mM Tris pH 8, 100mM KCl, 1mM MgCl₂, 20mM ZnSO₄, 10mM BME, and quenched with laemmli buffer. Reactions were read out via Western Blot. Proteins were separated on an 18% SDS PAGE gel, transferred to PVDF membrane in Tris-Glycine with %20 methanol, blocked, and stained with anti-H3K9me₂ (Abcam, ab1220) or anti-H3K9me₁ (Sigma-Aldrich AB_2793303) and anti-H4 as loading control (Active Motif AB_2636967).

Nucleosome binding assays

To assess the binding affinity between G9a/GLP (wild type or mutants) and nucleosome core particles (NCPs), a specific binding assay protocol was followed. The reaction mixture was prepared with 100 nM NCP, 7 μM G9a/GLP (either wild type or mutants), and 250 μM S-adenosylmethionine (SAM). This mixture was incubated at 25°C for 60 minutes in a buffer solution consisting of 20 mM HEPES-NaOH (pH 7.5), 40 mM NaCl, 1 mM MgCl₂, 1.1 mM dithiothreitol, 0.03% NP-40, and 0.5% glycerol to facilitate the interaction between the proteins and NCPs.

Following incubation, the crosslinking reaction within the mixture was halted by adding 1 M Tris-HCl (pH 7.5) to achieve a final concentration of 50 mM. The mixture was then cooled on ice for 10 minutes to stabilize the reactions. To analyze the binding interactions, samples were subjected to 5% non-denaturing polyacrylamide gel electrophoresis, employing a buffer of 0.53 TBE (45 mM Tris-borate and 1 mM EDTA), which allows for the resolution of protein-DNA complexes without denaturing them.

After electrophoresis, the gel was stained using SYBR Gold, a nucleic acid stain, to visualize the nucleosomes and thus determine the binding efficiency of G9a/GLP to the NCPs. This methodological approach enables the quantification and analysis of the affinity and specificity of G9a/GLP (wild type or mutants) towards NCPs under the defined experimental conditions.

Peptide methylation assays

Substrate peptides were mixed in a solution containing 500uM S-adenosyl methionine (SAM, PerkinElmer) cofactor, and reactions were initiated upon addition of 0.4 to 0.8 μ M enzyme. Reactions were run in 100 mM Tris pH 8, 100 mM KCl, 1 mM $MgCl_2$, 20 μ M $ZnSO_4$, 10 mM BME and quenched with TFA.

Crosslinking Mass Spectrometry

CLMS was performed as described with minor modifications^{49,50}.

3: H3K9me2 propagation requires association with the nuclear periphery

Abstract

Heterochromatic loci marked by H3K9me2 are associated with the nuclear periphery in metazoans, but the consequences of this association on H3K9me2 maintenance and propagation are unknown. Both the nuclear lamins and the nuclear membrane protein lamin B receptor (LBR) have been proposed to maintain the association of H3K9me2 with the nuclear periphery in mammals. Yet, embryonic stem cells (mESCs) lacking all lamins retain both LBR and H3K9me2 at the nuclear periphery. We show that while H3K9me2 detaches from the nuclear periphery in lamin + LBR quadruple knockout (QKO) mESCs, the genomic location, and levels of H3K9me2 remain constant. In contrast, while QKO mESCs sustain naïve pluripotency, they fail to expand H3K9me2 across the genome during EpiLC differentiation. These results demonstrate that H3K9me2 spreads during EpiLC differentiation and establish that the association of H3K9me2 with the nuclear periphery regulates this expansion.

3.1 Introduction

Cell identity is hypothesized to be controlled in part through the formation of gene-repressive H3K9me2 heterochromatin domains. H3K9me2 is predominantly catalyzed *in vivo* by the euchromatic histone methyltransferases G9a and G9a-like protein (GLP). In developing embryos loss of H3K9me2, G9a, or GLP, is associated with failed developmental competency and death. During critical cell state transitions including epiblast formation, hematopoiesis, and neurogenesis, H3K9me2 patterns are hypothesized to change to reinforce the transition and the underlying transcriptomic changes in a process termed heterochromatin spreading^{3,32,48,51–53}. The extent, contexts, and mechanisms of H3K9me2 spreading are currently unknown.

H3K9me2 is uniquely enriched at the nuclear periphery in many eukaryotes^{54–58}. The proteins that localize H3K9me2-marked loci to the nuclear periphery vary across eukaryotes. For instance, the inner nuclear membrane (INM) protein Amo1 tethers H3K9me2 to the nuclear periphery in *S. pombe*⁵⁴. In *C elegans*, the INM protein CEC-4 tethers H3K9me2-marked chromatin in embryos but is dispensable for this process in differentiated tissues, perhaps due to compensation by other proteins^{55,59}. H3K9me2 is highly enriched in lamina-associated domains (LADs) of chromatin^{56,60–63}, raising the possibility that the lamins and/or the lamin binding receptor (LBR) control the peripheral positioning of heterochromatin bearing this modification. Ablating the lamins and LBR appears to affect the spatial enrichment, but not the total abundance, of H3K9me2 (unpublished data). This observation suggests that the deposition of H3K9me2 onto loci occurs independently of tethering to the nuclear periphery. While we know that dimethylation deposition repositions loci to the periphery^{43,64,65}, it is unknown how the

periphery itself regulates dimethylation deposition, especially during cell state transitions.

To further investigate the functional significance of the nuclear periphery in regulating H3K9me2 dynamics, we utilized mouse embryonic stem cells (mESCs) lacking lamins and LBR. This model allowed us to dissect the contribution of nuclear periphery association to the propagation of H3K9me2 during differentiation. Our findings demonstrate that H3K9me2 spreads across the genome during differentiation into epiblast-like cells (EpiLCs). However, this spreading is compromised when H3K9me2-marked heterochromatin is disassociated from the nuclear periphery. This suggests that the nuclear periphery acts as a facilitator for the rapid expansion of H3K9me2-marked heterochromatin, highlighting its importance in cell state transitions. These insights underscore the complexity of chromatin organization and its regulation, emphasizing the nuclear periphery's role in facilitating dynamic changes in chromatin structure that are essential for cellular differentiation and development.

3.2 Results

H3K9me2 maintenance does not depend on its association with the nuclear periphery

Since mESCs lacking the lamins and LBR (QKO mESCs), but not either (TKO or LBRKO), have internalized their H3K9me2-marked heterochromatin, we asked whether this change in cellular localization affects the genomic abundance and distribution of H3K9me2. To evaluate the genomic position and abundance of H3K9me2, we used a monoclonal antibody with validated selectivity for this mark in genome-binding assays^{56,57} to perform CUT&RUN with spike-in control⁶⁶. Alignments show excellent agreement between replicates (**Fig. 3.1A**) with correlations between genotypes exceeding 0.8 in all cases. Visual inspection of H3K9me2 tracks also revealed remarkably similar distributions of the modification (**Fig. 3.1B**). We then applied a three-state hidden Markov model (HMM)⁶⁷ to identify domains of absent, intermediate (class 1), or high (class 2) H3K9me2 density (**Fig. 3.1B-C**). We noted a modestly higher density of H3K9me2 within both class 1 and class 2 domains in QKO mESCs compared to other genotypes (**Fig. 3.1C-D**). The association of genes with the nuclear periphery is correlated with repression of transcription⁶⁸⁻⁷². Consistently, we find a strong correlation between H3K9me2 modification and low transcription; genes found within H3K9me2-marked domains are less expressed than genes found outside these regions, and the strength of repression correlates with the density of H3K9me2 (**Fig. 3.1E**).

H3K9me2 domains appear similar across the genomes of WT, LBRKO, lamin TKO, and QKO mESCs (**Fig. 3.1A-B**). All genotypes possess similar numbers of class 1 and class 2 domains that cover approximately 60% of the genome in kilobase- to

megabase-long tracts (**Fig. 3.2A-C**; ~59%, ~65%, ~61%, and ~57% of the genome covered in WT, LBRKO, TKO, and QKO mESCs, respectively). Approximately 66% of H3K9me2 domains are shared across all four genotypes (**Fig. 3.2E**), and over 70% are present in both WT and QKO mESCs. Taken together, these data indicate that ablating association with the nuclear periphery does not affect the genomic position of H3K9me2 in mESCs.

H3K9me2 propagation depends on its association with the nuclear periphery

Previous reports have questioned whether and how H3K9me2 expands across the genome during differentiation^{28,33,73,74}. To test the influence of the nuclear periphery on this proposed H3K9me2 propagation, we asked whether peripheral association is required for the transition from the “naïve” pluripotent state of the inner cell mass of an embryo (approximated in culture by mESCs grown in 2i + LIF conditions) to the “primed” pluripotent state of the epiblast (approximated in culture by treatment with FGF and Activin A growth factors for 24-48 hours)^{75,76} (**Fig. 3.3A**). Importantly, this developmental transition is hypothesized to be dependent on a G9a-GLP-dependent expansion of H3K9me2 across the genome^{77,78}.

Microscopy analyses indicate that, like mESCs, H3K9me2 is displaced from the nuclear periphery in QKO EpiLCs (data not shown). Additionally, H3K9me2 abundance appears to increase during the transition from naïve to primed pluripotency in all genotypes, in line with some reports. We note that QKO mESCs, in which H3K9me2 is disassociated from the nuclear periphery, were severely impaired in their ability to generate EpiLCs (data not shown). These data therefore encouraged us to query

whether the change genomic abundance or distribution is dependent on the association with the nuclear periphery.

To quantify H3K9me2 abundance across the genome in EpiLCs, we again applied spike-in-controlled H3K9me2 CUT&RUN. We noted from correlation analysis (**Fig. 3.3B**) that although the technical quality of the experiment was high, there was less correlation between genotypes than in mESCs. Visual inspection of H3K9me2 tracks was also suggestive of increased signal in the QKOs, and an altered H3K9me2 distribution compared to other genotypes (**Fig. 3.3C**). We then identified domains of absent, intermediate (class 1), or high (class 2) H3K9me2 intensity (**Fig. 3.3C-E**). H3K9me2 domains that are shared in naïve and primed pluripotency, which we refer to as “constitutive H3K9me2,” are less transcribed than *de novo* H3K9me2 domains (**Fig. 3.3F**). We additionally noted that despite having higher overall signal in both class 1 and class 2 domains, the proportion of the genome covered in the QKO EpiLCs was less than the other genotypes (**Fig. 3.3G**), these results encouraged us to query the distribution of H3K9me2 domains in more detail.

Previous reports have questioned whether H3K9me2 expands across the genome during differentiation^{73,74}, our spike-in-controlled analysis demonstrated expansion of H3K9me2 across the genome in WT EpiLCs *versus* mESCs that was apparent by visual inspection, signal quantification in domains, and by quantifying total genome coverage (~66% of genome within H3K9me2 domains in EpiLCs vs. ~59% in ESCs) (**Fig. 3.4A**). We additionally quantified the median contiguous length of H3K9me2 domains in WT ESCs vs. EpiLCs (**Fig. 3.4B**; 130 kb in WT ESCs vs 190 kb in WT EpiLCs), and by tracking the net flow of genes into H3K9me2 domains in WT

EpiLCs *versus* ESCs (**Fig. 3.4C**). While H3K9me2 expands across the genome as LBRKO and lamin TKO mESCs transition to EpiLCs (**Fig. 3.4A-E**), this expansion is disrupted in survived QKO EpiLCs (**Fig. 3.4B, F**; median contiguous domain length of 120 kb in QKO ESCs vs 110 kb in QKO EpiLCs). Instead, H3K9me2 domains are smaller in QKO EpiLCs compared to other genotypes, fewer genes join H3K9me2 domains as QKO ESCs differentiate into EpiLCs (**Fig. 3.4F**). Additionally, QKO EpiLCs have a higher number of class 1 / intermediate density H3K9me2 domains than other genotypes. Interestingly, however, H3K9me2 accumulates to significantly higher levels on the genome of QKO EpiLCs compared to other genotypes (**Fig. 3.3**). Taken together, these analyses indicate that H3K9me2 is focally deposited at higher local densities within smaller, more fragmented H3K9me2 domains in the absence of peripheral association. Thus, while the total dose of H3K9me2 still increases as QKO ESCs differentiate into EpiLCs, the modification covers less of the genome. Because of this altered deposition of H3K9me2, fewer domains are preserved in QKO EpiLCs than in other genotypes. Approximately 55% of H3K9me2 domains are shared across all four genotypes (**Fig. 3.4G-H**); only 59% of domains are shared between WT and QKO EpiLCs, while 76% of domains are shared between WT and LBR KO EpiLCs. These data indicate that H3K9me2 propagation occurs during EpiLC differentiation and is dependent on an association with the nuclear periphery.

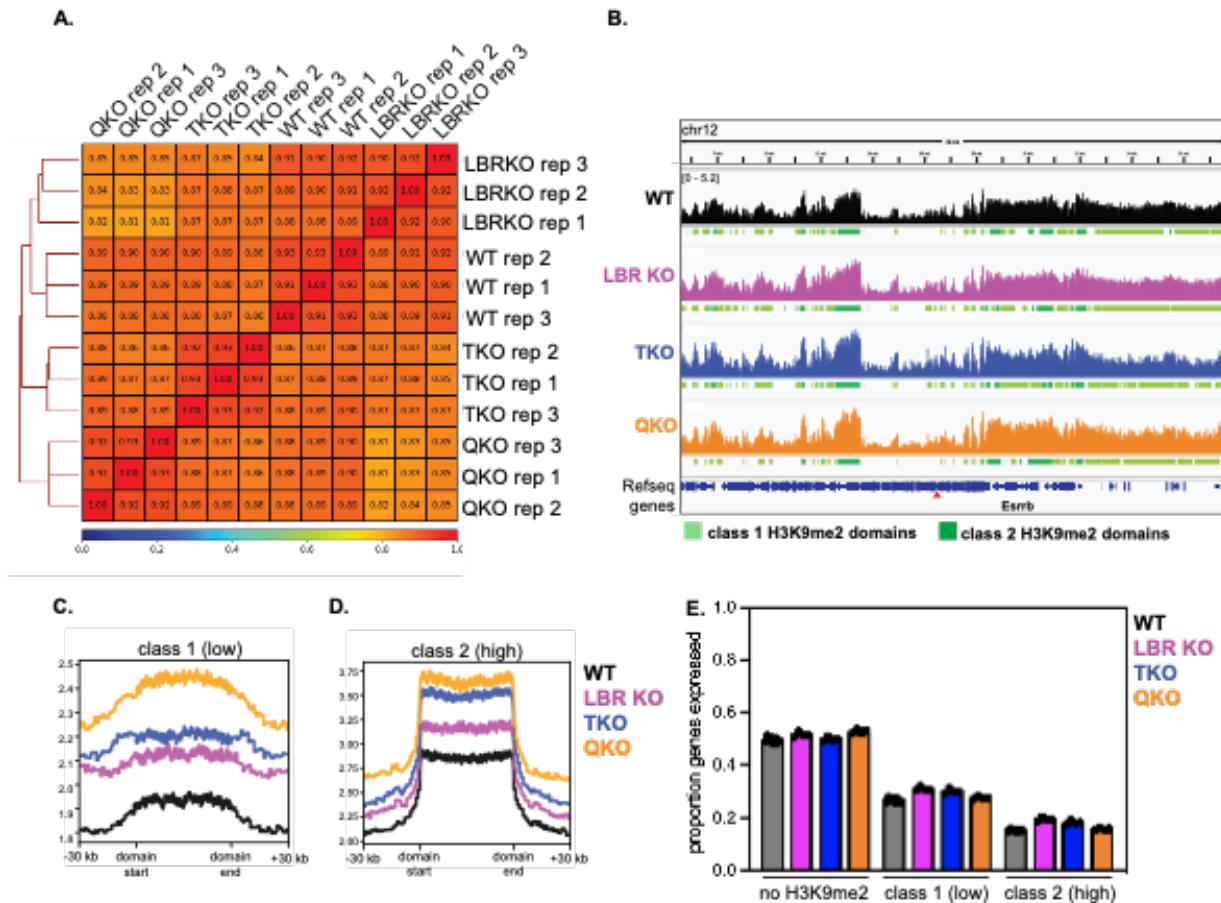


Figure 3.1. H3K9me2 CUT&RUN in mESCs

(A) Dendrogram and heatmap of individual H3K9me2 CUT&RUN replicates (3 per condition) showing similarity of replicates for each genotype. (B) Representative genome tracks and domain calls for H3K9me2 in WT, LBR KO, TKO, and QKO ESCs on a 20 Mb section of chromosome 12, including the *Esrrb* gene. The y-axis range indicated at the top left is the same for all tracks shown. Low K9me2 density “class 1” domains are marked in light green and high K9me2 density “class 2” domains are marked in dark green. Averaged density of H3K9me2 in all class 1 (C) and all class 2 (D) domains across genotypes. (E) The proportion of genes expressed (minimum of 5 TPMs) outside of H3K9me2 domains, within H3K9me2 class 1 domains, or H3K9me2 class 2 domains across genotypes.

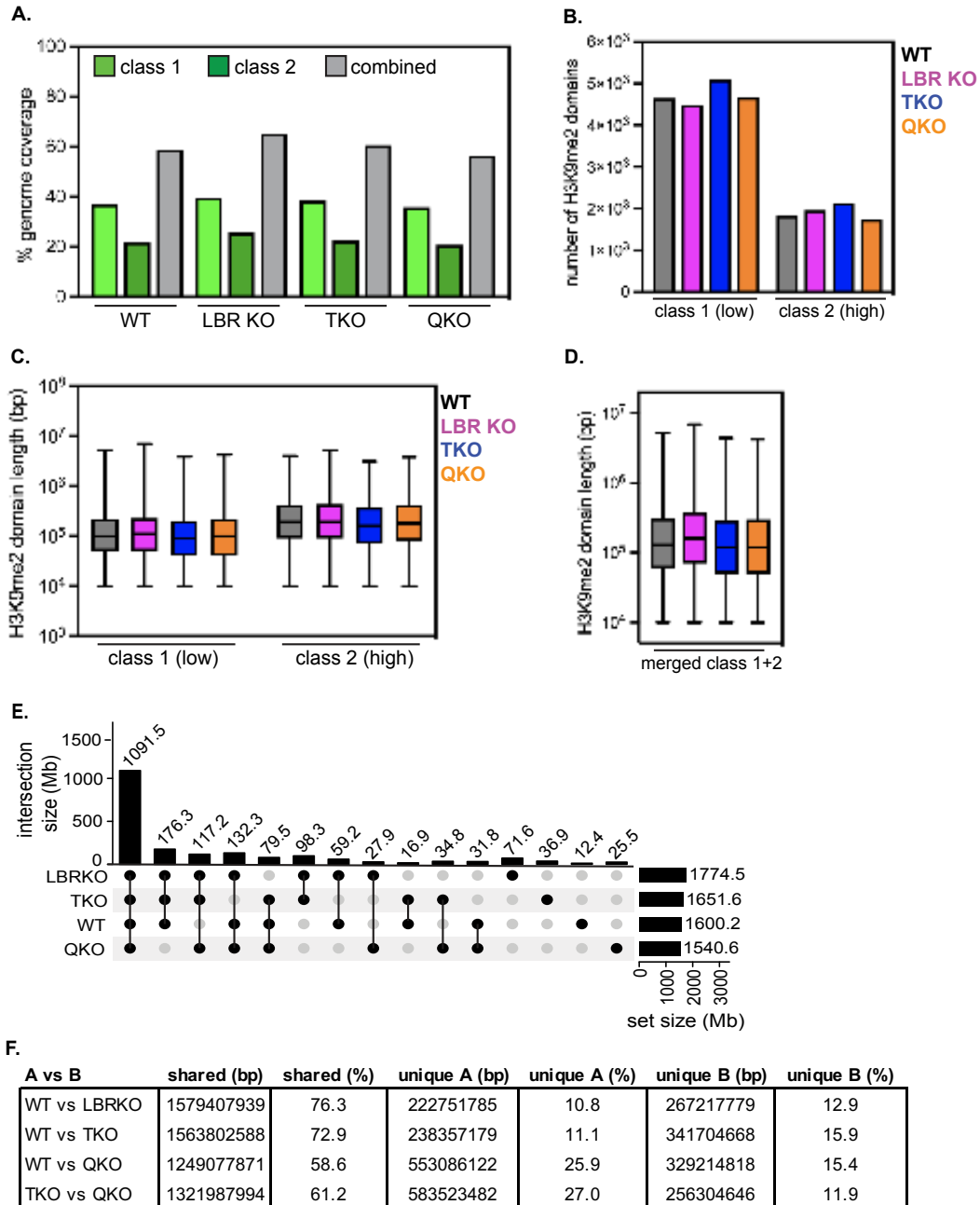


Figure 3.2. H3K9me2 maintenance does not depend on its association with the nuclear periphery

(B) Total genome coverage statistics for class 1 H3K9me2 domains, class 2 H3K9me2 domains, and merged H3K9me2 domains in each genotype of mESCs. (B) Number of HMM class 1 (low K9me2 density) and (Figure caption continued on the next page)

(Figure caption continued from the previous page) class 2 (high K9me2 density) domains called across genotypes. (C) Size of contiguous HMM class 1 (low K9me2 density) and class 2 (high K9me2 density) domains called across genotypes. (D) Size of contiguous HMM domains called across genotypes, adjacent class 1 and class 2 domains have been merged. (E) UpSet plot showing intersections of H3K9me2 domains between genotypes. (F) Summary of overlapping and unique H3K9me2 domains in pairwise comparisons between genotypes.

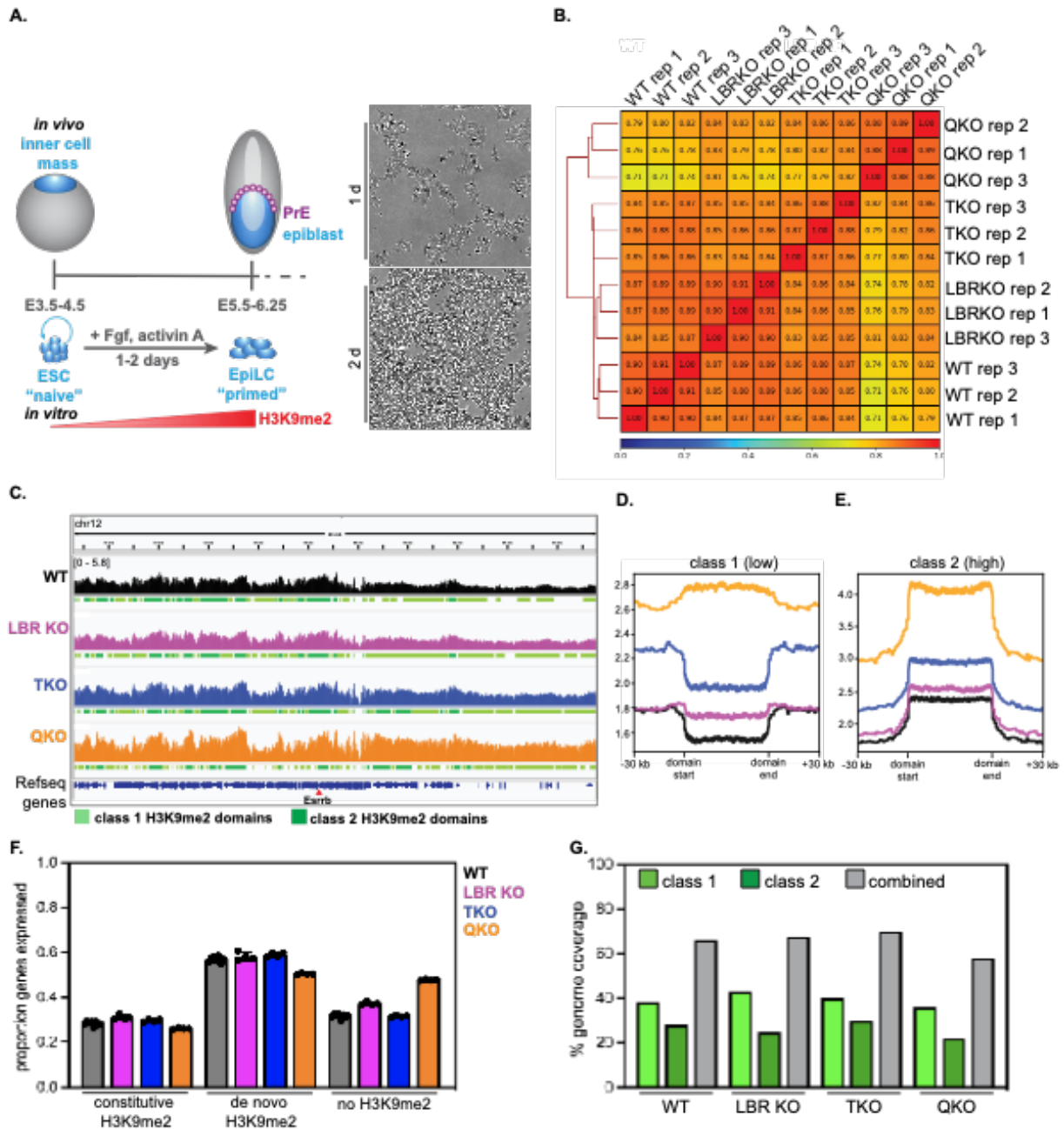


Figure 3.3. Peripheral association of H3K9me2 limits its production in EpiLCs

(A) Diagram of naïve (ICM / mESC) to primed (epiblast / EpiLC) developmental transition *in vivo* and its approximation *in vitro*. (B) Representative brightfield microscopy images of WT, LBR KO, lamin TKO, and lamin + LBR QKO mESCs after 1 day (top panel) or 2 days (bottom panel) of culture in EpiLC differentiation conditions. (B) Dendrogram and heatmap of individual (*Figure caption continued on the next page*)

(Figure caption continued from the previous page) H3K9me2 CUT&RUN replicates (3 per condition) showing similarity of replicates for each genotype. (C) Representative genome tracks and domain calls for H3K9me2 in WT, LBR KO, TKO, and QKO ESCs on a 20 Mb section of chromosome 12, including the *Esrrb* gene. The y-axis range indicated at the top left is the same for all tracks shown. Low K9me2 density “class 1” domains are marked in light green and high K9me2 density “class 2” domains are marked in dark green. Averaged density of H3K9me2 in all class 1 (D) and all class 2 (E) domains across genotypes. (F) The proportion of genes expressed (minimum of 5 TPMs) outside of H3K9me2 domains, within H3K9me2 class 1 domains, or H3K9me2 class 2 domains across genotypes. (G) Total genome coverage statistics for class 1 H3K9me2 domains, class 2 H3K9me2 domains, and merged H3K9me2 domains in each genotype of mESCs.

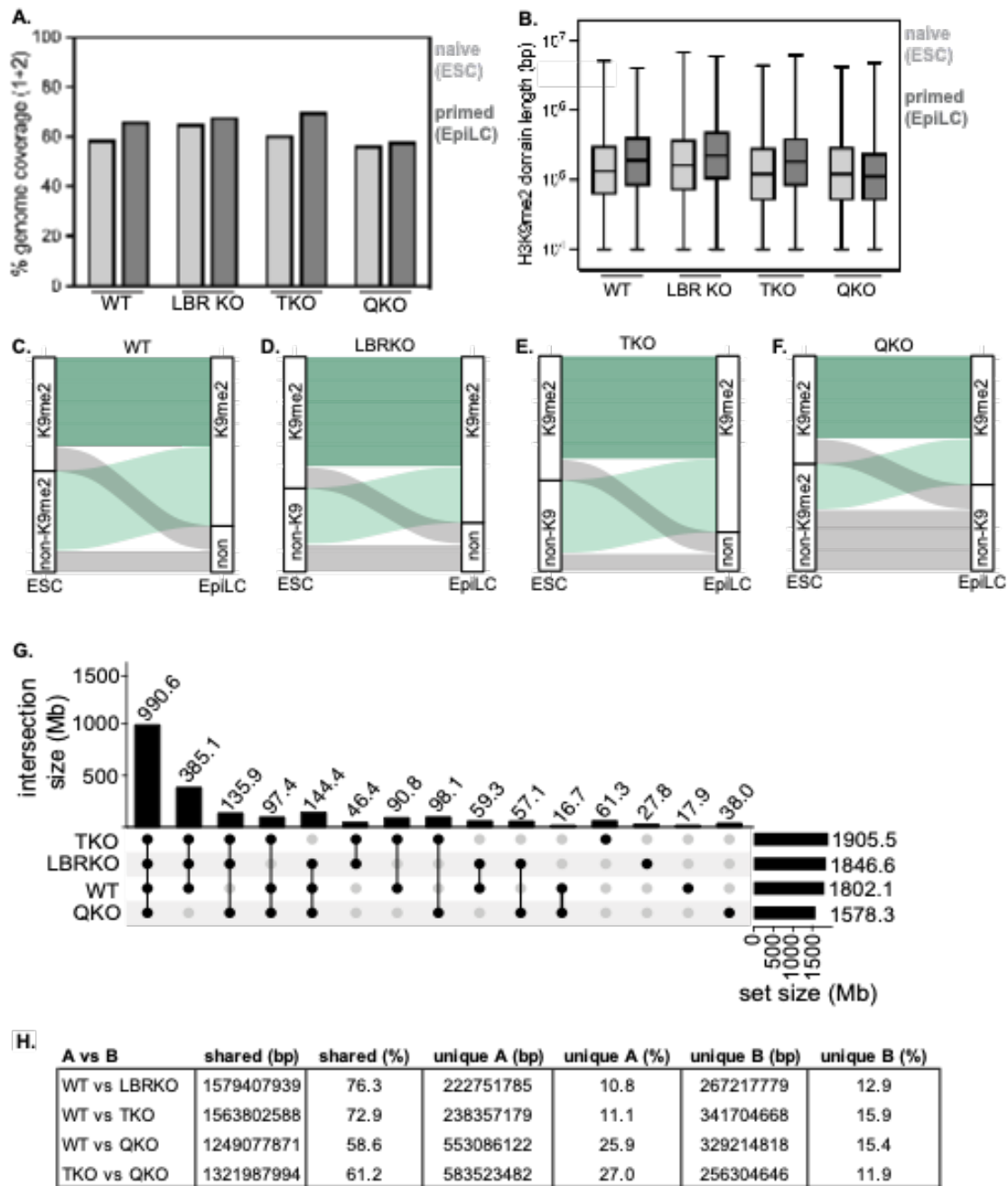


Figure 3.4. H3K9me2 propagation depends on its association with the nuclear periphery

(A) Comparison of total genome coverage in mESCs *versus* EpiLCs for each genotype. (B) Size of contiguous HMM class 1 (*Figure caption continued on the next page*)

(Figure caption continued from the previous page) (low K9me2 density) and class 2 (high K9me2 density) domains in ESCs versus EpiLCs of WT, LBR KO, lamin TKO, and lamin + LBR QKO cells. Alluvial plots showing the movement of genes into and out of H3K9me2 domains as ESCs differentiate into EpiLCs for WT (C), LBR KO (D), lamin TKO (E), and lamin + LBR QKO (F) cells. Genes found in H3K9me2 domains in both ESCs and EpiLCs are referred to as “constitutive” (dark green) while genes that move into H3K9me2 domains in EpiLCs are referred to as “de novo” (light green). (G) UpSet plot showing intersections of H3K9me2 domains between genotypes. (H) Summary of overlapping and unique H3K9me2 domains in pairwise comparisons between genotypes.

3.3 Discussion

Our data indicate that the lamins and LBR together exert broad control on heterochromatin organization in pluripotent cells. H3K9me2-marked loci are also displaced from the nuclear periphery and into nucleoplasmic foci in cells lacking these proteins. This phenotype is reminiscent of the intranuclear coalescence of heterochromatin that occurs in terminally differentiated photoreceptor neurons that lack lamin A/C and LBR^{79,80}, although less dramatic. Heterochromatin may not completely coalesce into a single intranuclear focus within one interphase in proliferative mESCs, while chromatin inversion takes place over several weeks after cell cycle exit in neurons⁷⁹. Alternatively, additional chromatin changes, such as relabeling of H3K9me2 loci with H3K9me3, may promote heterochromatin coalescence and inversion in neurons⁸⁰.

These data indicate that the lamins can sustain chromatin organization in the absence of LBR and vice versa. While LBR is much more highly expressed than lamin A/C in pluripotent cells, low levels of lamin A/C have been recently shown to influence gene expression in naïve pluripotency⁸¹. LBR alone can drive peripheral heterochromatin positioning when ectopically expressed in various cell types, while ectopically expressed Lamin A/C cannot⁷⁹. This implies that Lamin A/C's regulation of H3K9me2's peripheral association is mediated by additional factors with variable expression levels across tissues, while LBR can either directly tether heterochromatin or work through ubiquitously expressed intermediary protein(s). Various lamin-bound proteins such as LAP2b, HDAC3, PRR14, and others are candidates that could promote

lamin-mediated heterochromatin tethering^{56,71,82,83}, while LBR binds to the H3K9me2/3-binding protein HP1⁸⁴ and can interact with H3 tails^{85,86}.

During the transition from naïve to primed pluripotency, *de novo* H3K9me2 modification is found on loci that remain actively transcribed and harbor markers of ongoing transcription, such as H3K27ac⁷⁸. This transitory state may be a prerequisite for transcriptional inactivation of these genes to orchestrate differentiation^{75,78}. We discover that association with the nuclear periphery is required for the differentiation-coupled expansion of H3K9me2 across the genome, and find that in the absence of this association, *de novo* H3K9me2 does not expand as broadly, but H3K9me2 instead accumulates to higher local densities within more fragmented domains. These data raise the possibility that H3K9me2 tethering influences the deposition of H3K9me2 by G9a-GLP or enhances the stability of H3K9me2 on modified loci. Some evidence exists for the latter possibility, as H3K9me2 appears to persist for a longer time on LADs than on non-LAD loci when G9a-GLP is pharmacologically inhibited⁸⁷. Overall, this work establishes a model system whereby H3K9me2 expansion occurs and demonstrates that association with the periphery, mediated by the lamins and LBR, is necessary for its propagation.

3.4 Methods

mESC culture, EpiLC differentiation, and culture on Laminin

Cells were cultured in 5% CO₂ and 37°C using the Thermo Scientific Steri-Cycle i160 CO₂ Incubator. mESCs were grown in serum- and feeder-free 2i + LIF medium, which consists of N2B27 basal medium supplemented with the following compounds at a final concentration of 3 μM CHIR-99021 (Selleckchem S1263), 1 μM PD0325901 (Selleckchem S1036), 10³ U/mL LIF (MilliporeSigma™ ESG1107), 55 μM β-mercaptoethanol (Gibco™ 21985023), and 1X PenStrep (GenClone 25-512).

N2B27 medium was made by mixing equal parts of DMEM: F12 (GenClone 25-503) and Neurobasal medium (Gibco™ 21103049) then adding 1X N-2 (Gibco™ 17502001), 1X B-27 (Gibco™ 17504001) and 2 mM Glutamax (Gibco™ 35050061).

Media was sterile filtered using 0.2μm pore size filter bottles (VWR 97066-202). Dishes were coated with 0.1% gelatin (MilliporeSigma ES-006-B) for 30 minutes at 37°C before seeding cells. Serum-free 2i + LIF media was replenished every day and cells were passaged every other day, seeding 4x10⁵ cells per 6-well (approximately 4.21 x 10⁴ cells/cm²).

Differentiation of mESCs into epiblast-like cells was adapted from a published protocol. Briefly, 1 mL of a 1:200 solution of either Geltrex (Gibco™ A1413202) or Cultrex (Bio-Techne 3445-005-01) in N2B27 was used to coat a 6-well at 37°C. After one hour, that solution was aspirated and 2 x 10⁵ mESCs were seeded (approximately 2.11 x 10⁴ cells/cm²) with N2B27 media containing a final concentration of 20 ng/mL FGF (Peprotech 100-18B) and 12 ng/mL Activin A (Peprotech 120-14E). Fresh EpiLC media was added the next day.

CUT&RUN and library preparation

CUT&RUN was performed as described in Skene et al. 2018 (Nature Protocols). Live cells were harvested with accutase and washed once with PBS before starting the procedure. Three replicates containing 200,000 mESCs or EpiLCs each were used per condition plus 2,000 HEK293T cells as a spike-in control. H3K9me2 antibody (Abcam 1220) was used at 1:100 dilution in antibody buffer. Primary antibody incubation was performed at 4°C overnight. Rabbit anti-mouse secondary antibody was used at 1:100 dilution and incubated at 4°C for 1 hour. pA-MNase (batch #6 143µg/mL) gifted from the Hennikoff lab was used for the cleavage reaction.

DNA was isolated by phenol-chloroform extraction. Purified DNA was resuspended with 40µL of 1mM Trish-HCl pH 8.0 and 0.1mM EDTA. DNA concentration was measured using the Qubit dsDNA HS assay kit (Invitrogen Q32851) and Qubit 4 fluorometer. DNA quality was assayed using the Agilent Bioanalyzer 2100 with the High Sensitivity DNA kit. Libraries were prepared using the NEBNext Ultra II DNA Library Prep with Sample Purification Beads (NEB E7103S) following the manufacturer's protocol with 5ng of starting DNA input. Each library was uniquely indexed with NEBNext Multiplex Oligos for Illumina (NEB E6440S), then pooled together in equimolar amounts. Library size was analyzed with the Agilent Bioanalyzer 2100 using the High sensitivity DNA kit. Paired-end sequencing was performed on the pooled library using the Illumina NextSeq2000 platform (read length of 35bp and read depth of approximately 16 million reads per library).

CUT&RUN analysis

Raw sequencing reads were mapped to the “soft masked” mm39 mouse genome (<https://hgdownload.soe.ucsc.edu/goldenPath/mm39/bigZips/mm39.fa.gz>) using Bowtie2 version 2.3.5.1: `-X 2000 -N 1 --local --dovetail`. SAMtools version 1.10 was used to keep properly paired, primary alignments and filter out unassembled contigs, read duplicates, and reads mapped to mitochondria. For the spike-in control scale factor calculations, raw sequencing reads were mapped to the “soft masked” hg38 human genome (<https://hgdownload2.soe.ucsc.edu/goldenPath/hg38/bigZips/hg38.fa.gz>). As was done for mm39, Bowtie2 and SAMtools were used for alignment and post-alignment processing, respectively. The SAMtools function `flagstat` was used to find the number of properly paired reads in hg38 alignments. These were used to calculate a scale factor that was defined by dividing an arbitrary constant number (30,000) by the number of properly paired reads (Zheng Y et al (2020). Protocol.io). The scale factors for each library were used to generate bigWig files of 1kb and 10kb bin, spike-in normalized RPKM signal coverage tracks with the deepTools2 function `bamCoverage`. Hidden Markov models were implemented on each 10kb bin bigwig file with the pomegranate package in Python to call four regions of different H3K9me2 signal 1. background signal 2. low signal 3. high signal and 4. very high artifact signal representative of “blacklist” regions (Jacob Schreiber. 2017. Pomegranate: fast and flexible probabilistic modeling in python. J. Mach. Learn. Res. 18, 1 (January 2017), 5992–5997.). Replicate BED files containing low (class 1) and high (class 2) H3K9me2 domains were merged using the bedtools function `multiinter`. For each condition, replicate 1kb bigWig files were averaged using the deepTools2 function `bigwigAverage`.

Concluding Thoughts

In this work, we sought to understand the mechanism by which G9a-GLP-catalyzed H3K9me2 spreads. In homologous systems, product recognition drives heterochromatin formation. Following ablation of product recognition, heterochromatin domains collapse, leaving nucleation sequences as the only remaining heterochromatin. Via various techniques that temporally track the formation of heterochromatin following its ablation including ChIP-seq and the Heterochromatin Spreading Sensor (HSS), we understand that wild-type methyltransferases linearly propagate heterochromatin outward from these small nucleation sites. The discovery of small H3K9me2 domains in mESCs that appeared to expand in a cell-type specific manner, as well as the discovery of G9a-GLP's product recognition domains, led to the model that H3K9me2 formation, too, proceeds by a linear mechanism. To test this model, our original intention was to ablate product recognition in G9a-GLP and track heterochromatin formation in ESCs via an analogous HSS, thereby recreating classic experiments in a novel model system.

We quickly realized that H3K9me2 was not itself instructive of repression, which the HSS relies on as its output, in the same manner as H3K9me3. Additionally, we were unable to demonstrate product recognition-driven catalytic stimulation *in vitro*. Further, the extent and contexts in which H3K9me2 spread *in vivo* were and continue to be challenged. Thus, we utilized CUT&RUN to 1) determine whether spreading itself occurs in a disputed context, thereby 2) establishing a model system in which to study spreading following the ablation of product recognition of G9a-GLP, and 3) identify regions that gained dimethylation that were also dependent on that methylation for repression, enabling the development of a mammalian HSS.

Towards these ends, we established via CUT&RUN and other unpublished data that EpiLC differentiation is accompanied by an expansion of H3K9me2. That is, H3K9me2 does indeed spread. However, we challenge the notion that H3K9me2 exists in sparse, small domains in mESCs before differentiation. In contrast, we find that ~60% of the genome comprises H3K9me2 domains ranging in the 10s of kilobases to 10s of megabases in size. While these domains do indeed grow on average in size during differentiation, we note a dramatic redistribution of a large proportion of these domains and the establishment of domains *de novo* independent of pre-existing H3K9me2. With the development of novel analytic tools to analyze spreading by other methyltransferases, future directions include performing CUT&RUN at higher time resolution in EpiLC differentiation and other contexts to identify mechanisms by which H3K9me2 spreads. Additionally, this study does not describe regions of H3K9me2 that are regulated by G9a-GLP. Thus, additional follow-ups include utilizing G9a-GLP inhibitors or knockout cell lines in combination with CUT&RUN to identify G9a-GLP-regulated regions.

Initially, to identify spreading mutants that are relevant to the investigation in the above context, we orthogonally began querying G9a-GLP mutants *in vitro*. While we first sought to use mononucleosome methylation as a starting point toward eventual dinucleosome substrates, our discovery that dimethylation on the mononucleosome is specifically regulated led us down an unexpected path. Mechanistically, we demonstrate an absolute requirement for a *cis* complement of product recognition domains in H3K9me2 production but do not identify the molecular mechanism by which this regulation occurs. Thus, gaining structural insight into the modes of wild-type and

mutant G9a-GLP engagement with differentially methylated nucleosomes is an area for fruitful exploration. The mode of H3K9me2 regulation we've uncovered *in vitro* stands in stark contrast to other repressive lysine methyltransferases, where product recognition stimulates catalysis generally and is involved in linear heterochromatin spreading. Because our study is limited to mononucleosome substrates, future directions *in vitro* include assaying these mutants on dinucleosomes or polysome substrates.

While we expect product recognition by G9a-GLP to aid in its recruitment to regions of pre-existing H3K9me2, this study provides no insight into how the formation of a domain occurs *de novo* or following hypothesized nucleation. As described, elucidating these mechanisms involves ablating heterochromatin genetically or chemically, and tracking its formation temporally following removal of chemical inhibition or re-introduction of removed proteins. In these studies, we have established the essential ingredients to undertake a full *in vivo* exploration of G9a-GLP-catalyzed H3K9me2 formation. We have demonstrated that H3K9me2 does indeed spread during EpiLC differentiation, and uncovered and described how product recognition by G9a-GLP affects methylation on the repeat unit of chromatin, the mononucleosome. Given the dearth of evidence for linear H3K9me2 spreading by G9a-GLP, we strongly believe that the mechanism for propagation *in vivo* is unlike any that is currently described. A combination of CUT&RUN with mutants described in this text would provide an excellent system in which to study how product recognition affects G9a-GLP-catalyzed H3K9me2 domain during cell state transitions.

References

1. Hathaway, N.A., Bell, O., Hodges, C., Miller, E.L., Neel, D.S., and Crabtree, G.R. (2012). Dynamics and Memory of Heterochromatin in Living Cells. *Cell* 149. 10.1016/j.cell.2012.03.052.
2. Hall, I.M., Shankaranarayana, G.D., Noma, K., Ayoub, N., Cohen, A., and Grewal, S.I.S. (2002). Establishment and Maintenance of a Heterochromatin Domain. *Science* 297, 2232–2237. 10.1126/science.1076466.
3. Ugarte, F., Sousae, R., Cinquin, B., Martin, E.W., Krietsch, J., Sanchez, G., Inman, M., Tsang, H., Warr, M., Passequé, E., et al. (2015). Progressive Chromatin Condensation and H3K9 Methylation Regulate the Differentiation of Embryonic and Hematopoietic Stem Cells. *Stem Cell Rep* 5, 728–740. 10.1016/j.stemcr.2015.09.009.
4. Allshire, R.C., and Madhani, H.D. (2017). Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Bio* 19, 229. 10.1038/nrm.2017.119.
5. Verschure, P.J., Kraan, I. van der, Leeuw, W. de, Vlag, J. van der, Carpenter, A.E., Belmont, A.S., and Driel, R. van (2005). In Vivo HP1 Targeting Causes Large-Scale Chromatin Condensation and Enhanced Histone Lysine Methylation†. *Mol Cell Biol* 25, 4552–4564. 10.1128/mcb.25.11.4552-4564.2005.
6. Lachner, M., O’Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116. 10.1038/35065132.

7. Wang, X., and Moazed, D. (2017). DNA sequence-dependent epigenetic inheritance of gene silencing and histone H3K9 methylation. *Science* 356, 88–91. 10.1126/science.aaj2114.
8. Yu, R., Wang, X., and Moazed, D. (2018). Epigenetic inheritance mediated by coupling of RNAi and histone H3K9 methylation. *Nature* 558, 615–619. 10.1038/s41586-018-0239-3.
9. Rangunathan, K., Jih, G., and Moazed, D. (2015). Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* 348, 1258699. 10.1126/science.1258699.
10. Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell* 128, 693–705. 10.1016/j.cell.2007.02.005.
11. Xie, W.J., and Zhang, B. (2019). Learning the Formation Mechanism of Domain-Level Chromatin States with Epigenomics Data. *Biophys J* 116, 2047–2056. 10.1016/j.bpj.2019.04.006.
12. Obersriebnig, M.J., Pallesen, E.M.H., Sneppen, K., Trusina, A., and Thon, G. (2016). Nucleation and spreading of a heterochromatic domain in fission yeast. *Nat Commun* 7, 11518. 10.1038/ncomms11518.
13. Pan, M.-R., Hsu, M.-C., Chen, L.-T., and Hung, W.-C. (2018). Orchestration of H3K27 methylation: mechanisms and therapeutic implication. *Cell Mol Life Sci* 75, 209–223. 10.1007/s00018-017-2596-8.

14. Haddad, N., Jost, D., and Vaillant, C. (2017). Perspectives: using polymer modeling to understand the formation and function of nuclear compartments. *Chromosome Res* 25, 35–50. 10.1007/s10577-016-9548-2.
15. Osborne, E.A., Dudoit, S., and Rine, J. (2009). The establishment of gene silencing at single-cell resolution. *Nat Genet* 41, ng.402. 10.1038/ng.402.
16. Jones, A., and Wang, H. (2010). Polycomb repressive complex 2 in embryonic stem cells: an overview. *Protein Cell* 1, 1056–1062. 10.1007/s13238-010-0142-7.
17. Oksuz, O., Narendra, V., Lee, C.-H., Descostes, N., LeRoy, G., Raviram, R., Blumenberg, L., Karch, K., Rocha, P.P., Garcia, B.A., et al. (2018). Capturing the Onset of PRC2-Mediated Repressive Domain Formation. *Mol Cell* 70, 1149-1162.e5. 10.1016/j.molcel.2018.05.023.
18. Lee, C.-H., Yu, J.-R., Kumar, S., Jin, Y., LeRoy, G., Bhanu, N., Kaneko, S., Garcia, B.A., Hamilton, A.D., and Reinberg, D. (2018). Allosteric Activation Dictates PRC2 Activity Independent of Its Recruitment to Chromatin. *Mol Cell* 70, 422-434.e6. 10.1016/j.molcel.2018.03.020.
19. Margueron, R., Justin, N., Ohno, K., Sharpe, M.L., Son, J., III, W.J.D., Voigt, P., Martin, S.R., Taylor, W.R., Marco, V.D., et al. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* 461, 762. 10.1038/nature08398.

20. Weirich, S., Khella, M.S., and Jeltsch, A. (2021). Structure, Activity and Function of the Suv39h1 and Suv39h2 Protein Lysine Methyltransferases. *Life* 11, 703. 10.3390/life11070703.
21. Müller, M.M., Fierz, B., Bittova, L., Liszczak, G., and Muir, T.W. (2016). A two-state activation mechanism controls the histone methyltransferase Suv39h1. *Nat. Chem. Biol.* 12, 188–193. 10.1038/nchembio.2008.
22. Wang, T., Xu, C., Liu, Y., Fan, K., Li, Z., Sun, X., Ouyang, H., Zhang, X., Zhang, J., Li, Y., et al. (2012). Crystal Structure of the Human SUV39H1 Chromodomain and Its Recognition of Histone H3K9me_{2/3}. *PLoS ONE* 7, e52977. 10.1371/journal.pone.0052977.
23. Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y., and Allis, C.D. (2003). Histone Methyltransferases Direct Different Degrees of Methylation to Define Distinct Chromatin Domains. *Mol Cell* 12, 1591–1598. 10.1016/s1097-2765(03)00479-9.
24. Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., et al. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Gene Dev* 16, 1779–1791. 10.1101/gad.989402.
25. Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T., and Shinkai, Y. (2005). Histone methyltransferases G9a

and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Gene Dev* 19, 815–826. 10.1101/gad.1284005.

26. Collins, R.E., Northrop, J.P., Horton, J.R., Lee, D.Y., Zhang, X., Stallcup, M.R., and Cheng, X. (2008). The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nat Struct Mol Biology* 15, nsmb.1384. 10.1038/nsmb.1384.

27. Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet* 41, ng.297. 10.1038/ng.297.

28. Lienert, F., Mohn, F., Tiwari, V.K., Baubec, T., Roloff, T.C., Gaidatzis, D., Stadler, M.B., and Schübeler, D. (2011). Genomic Prevalence of Heterochromatic H3K9me2 and Transcription Do Not Discriminate Pluripotent from Terminally Differentiated Cells. *Plos Genet* 7, e1002090. 10.1371/journal.pgen.1002090.

29. Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2010). Reply to “Reassessing the abundance of H3K9me2 chromatin domains in embryonic stem cells.” *Nat Genet* 42, 5. 10.1038/ng0110-5.

30. Fillion, G.J., and Steensel, B. van (2010). Reassessing the abundance of H3K9me2 chromatin domains in embryonic stem cells. *Nat Genet* 42, 4. 10.1038/ng0110-4.

31. Yang, P., Humphrey, S.J., Cinghu, S., Pathania, R., Oldfield, A.J., Kumar, D., Perera, D., Yang, J.Y.H., James, D.E., Mann, M., et al. (2019). Multi-omic Profiling

Reveals Dynamics of the Phased Progression of Pluripotency. *Cell Syst* 8, 427-445.e10. 10.1016/j.cels.2019.03.012.

32. Kurimoto, K., Yabuta, Y., Hayashi, K., Ohta, H., Kiyonari, H., Mitani, T., Moritoki, Y., Kohri, K., Kimura, H., Yamamoto, T., et al. (2015). Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells. *Cell Stem Cell* 16, 517–532. 10.1016/j.stem.2015.03.002.

33. Zyllicz, J.J., Dietmann, S., Günesdogan, U., Hackett, J.A., Cougot, D., Lee, C., and Surani, M.A. (2015). Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. *Elife* 4, e09571. 10.7554/elife.09571.

34. Bittencourt, D., Lee, B.H., Gao, L., Gerke, D.S., and Stallcup, M.R. (2014). Role of distinct surfaces of the G9a ankyrin repeat domain in histone and DNA methylation during embryonic stem cell self-renewal and differentiation. *Epigenet Chromatin* 7, 27. 10.1186/1756-8935-7-27.

35. Liu, N., Zhang, Z., Wu, H., Jiang, Y., Meng, L., Xiong, J., Zhao, Z., Zhou, X., Li, J., Li, H., et al. (2015). Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability. *Gene Dev* 29, 379–393. 10.1101/gad.254425.114.

36. Shinkai, Y., and Tachibana, M. (2011). H3K9 methyltransferase G9a and the related molecule GLP. *Gene Dev* 25, 781–788. 10.1101/gad.2027411.

37. Bian, C., Chen, Q., and Yu, X. (2015). The zinc finger proteins ZNF644 and WIZ regulate the G9a/GLP complex for gene repression. *Elife* 4, e05606. 10.7554/elife.05606.
38. Al-Sady, B., Madhani, H.D., and Narlikar, G.J. (2013). Division of Labor between the Chromodomains of HP1 and Suv39 Methylase Enables Coordination of Heterochromatin Spread. *Mol Cell* 51, 80–91. 10.1016/j.molcel.2013.06.013.
39. Sanchez, N.A., Kallweit, L.M., Trnka, M.J., Clemmer, C.L., and Al-Sady, B. (2021). Heterodimerization of H3K9 histone methyltransferases G9a and GLP activates methyl reading and writing capabilities. *J. Biol. Chem.* 297, 101276. 10.1016/j.jbc.2021.101276.
40. Kalantry, S., and Magnuson, T. (2006). The Polycomb Group Protein EED Is Dispensable for the Initiation of Random X-Chromosome Inactivation. *Plos Genet* 2, e66. 10.1371/journal.pgen.0020066.
41. Li, Z., Jiao, X., Sante, G.D., Ertel, A., Casimiro, M.C., Wang, M., Katiyar, S., Ju, X., Klopfenstein, D.V., Tozeren, A., et al. (2019). Cyclin D1 integrates G9a-mediated histone methylation. *Oncogene* 38, 4232–4249. 10.1038/s41388-019-0723-8.
42. Fukuda, M., Sakaue-Sawano, A., Shimura, C., Tachibana, M., Miyawaki, A., and Shinkai, Y. (2019). G9a-dependent histone methylation can be induced in G1 phase of cell cycle. *Sci Rep-uk* 9, 956. 10.1038/s41598-018-37507-5.
43. Poleshko, A., Smith, C.L., Nguyen, S.C., Sivaramakrishnan, P., Wong, K.G., Murray, J.I., Lakadamyali, M., Joyce, E.F., Jain, R., and Epstein, J.A. (2019). H3K9me2

orchestrates inheritance of spatial positioning of peripheral heterochromatin through mitosis. *Elife* 8, e49278. 10.7554/elife.49278.

44. Rao, R.A., Ketkar, A.A., Kedia, N., Krishnamoorthy, V.K., Lakshmanan, V., Kumar, P., Mohanty, A., Kumar, S., Raja, S.O., Gulyani, A., et al. (2019). Euchromatic histone methyltransferases regulate peripheral heterochromatin tethering via histone and non-histone protein methylations. *Biorxiv*, 240952. 10.1101/240952.

45. Zylitz, J.J., Borensztein, M., Wong, F.C., Huang, Y., Lee, C., Dietmann, S., and Surani, M.A. (2018). G9a regulates temporal preimplantation developmental program and lineage segregation in blastocyst. *Elife* 7, e33361. 10.7554/elife.33361.

46. Yokochi, T., Poduch, K., Ryba, T., Lu, J., Hiratani, I., Tachibana, M., Shinkai, Y., and Gilbert, D.M. (2009). G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proc National Acad Sci* 106, 19363–19368. 10.1073/pnas.0906142106.

47. Zhang, T., Termanis, A., Özkan, B., Bao, X.X., Culley, J., Alves, F. de L., Rappsilber, J., Ramsahoye, B., and Stancheva, I. (2016). G9a/GLP Complex Maintains Imprinted DNA Methylation in Embryonic Stem Cells. *Cell Reports* 15, 77–85. 10.1016/j.celrep.2016.03.007.

48. Chen, X., Skutt-Kakaria, K., Davison, J., Ou, Y.-L., Choi, E., Malik, P., Loeb, K., Wood, B., Georges, G., Torok-Storb, B., et al. (2012). G9a/GLP-dependent histone H3K9me2 patterning during human hematopoietic stem cell lineage commitment. *Gene Dev* 26, 2499–2511. 10.1101/gad.200329.112.

49. McGilvray, P.T., Anghel, S.A., Sundaram, A., Zhong, F., Trnka, M.J., Fuller, J.R., Hu, H., Burlingame, A.L., and Keenan, R.J. (2020). An ER translocon for multi-pass membrane protein biogenesis. *eLife* 9, e56889. [10.7554/elife.56889](https://doi.org/10.7554/elife.56889).
50. Trnka, M.J., Baker, P.R., Robinson, P.J.J., Burlingame, A.L., and Chalkley, R.J. (2014). Matching Cross-linked Peptide Spectra: Only as Good as the Worst Identification*. *Mol. Cell. Proteom.* 13, 420–434. [10.1074/mcp.m113.034009](https://doi.org/10.1074/mcp.m113.034009).
51. Consortium, R.E., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317. [10.1038/nature14248](https://doi.org/10.1038/nature14248).
52. Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of Enhancer Patterns in Transition from Naive to Primed Pluripotency. *Cell Stem Cell* 14, 838–853. [10.1016/j.stem.2014.04.003](https://doi.org/10.1016/j.stem.2014.04.003).
53. Argelaguet, R., Mohammed, H., Clark, S., Stapel, C., Krueger, C., Kapourani, C.A., Xiang, Y., Hanna, C., Smallwood, S., Soria, X.I., et al. (2019). Single cell multi-omics profiling reveals a hierarchical epigenetic landscape during mammalian germ layer specification. *Biorxiv*, 519207. [10.1101/519207](https://doi.org/10.1101/519207).
54. Holla, S., Dhakshnamoorthy, J., Folco, H.D., Balachandran, V., Xiao, H., Sun, L., Wheeler, D., Zofall, M., and Grewal, S.I.S. (2019). Positioning Heterochromatin at the Nuclear Periphery Suppresses Histone Turnover to Promote Epigenetic Inheritance. *Cell*, 1–31.

55. Gonzalez-Sandoval, A., Towbin, B.D., Kalck, V., Cebianca, D.S., Gaidatzis, D., Hauer, M.H., Geng, L., Wang, L., Yang, T., Wang, X., et al. (2015). Perinuclear Anchoring of H3K9-Methylated Chromatin Stabilizes Induced Cell Fate in *C. elegans* Embryos. *Cell* 163, 1333–1347.
56. Poleshko, A., Shah, P.P., Gupta, M., Babu, A., Morley, M.P., Manderfield, L.J., Ifkovits, J.L., Calderon, D., Aghajanian, H., Sierra-Pagán, J.E., et al. (2017). Genome-Nuclear Lamina Interactions Regulate Cardiac Stem Cell Lineage Restriction. *Cell* 171, 573-580.e14.
57. Poleshko, A., Smith, C.L., Nguyen, S.C., Sivaramakrishnan, P., Wong, K.G., Murray, J.I., Lakadamyali, M., Joyce, E.F., Jain, R., and Epstein, J.A. (2019). H3K9me2 orchestrates inheritance of spatial positioning of peripheral heterochromatin through mitosis. *eLife* 8, 61.
58. See, K., Kiseleva, A.A., Smith, C.L., Liu, F., Li, J., Poleshko, A., and Epstein, J.A. (2020). Histone methyltransferase activity programs nuclear peripheral genome positioning. *Dev Biol* 466, 90–98. 10.1016/j.ydbio.2020.07.010.
59. Cebianca, D.S., Muñoz-Jiménez, C., Kalck, V., Gaidatzis, D., Padeken, J., Seeber, A., Askjaer, P., and Gasser, S.M. (2019). Active chromatin marks drive spatial sequestration of heterochromatin in *C. elegans* nuclei. Nature Publishing Group, 1–20.
60. Wu, R., Terry, A.V., Singh, P.B., and Gilbert, D.M. (2005). Differential Subnuclear Localization and Replication Timing of Histone H3 Lysine 9 Methylation States. *Mol. Biol. Cell* 16, 2872–2881. 10.1091/mbc.e04-11-0997.

61. Yokochi, T., Poduch, K., Ryba, T., Lu, J., Hiratani, I., Tachibana, M., Shinkai, Y., and Gilbert, D.M. (2009). G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proc. Natl. Acad. Sci.* *106*, 19363–19368. [10.1073/pnas.0906142106](https://doi.org/10.1073/pnas.0906142106).
62. Kind, J., Pagie, L., Ortazokoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-Cell Dynamics of Genome-Nuclear Lamina Interactions. *Cell* *153*, 178–192. [10.1016/j.cell.2013.02.028](https://doi.org/10.1016/j.cell.2013.02.028).
63. Shah, P.P., Keough, K.C., Gjoni, K., Santini, G.T., Abdill, R.J., Wickramasinghe, N.M., Dundes, C.E., Karnay, A., Chen, A., Salomon, R.E.A., et al. (2023). An atlas of lamina-associated chromatin across twelve human cell types reveals an intermediate chromatin subtype. *Genome Biol* *24*, 16. [10.1186/s13059-023-02849-5](https://doi.org/10.1186/s13059-023-02849-5).
64. Harr, J.C., Luperchio, T.R., Wong, X., Cohen, E., Wheelan, S.J., and Reddy, K.L. (2015). Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *J Cell Biology* *208*, 33–52. [10.1083/jcb.201405110](https://doi.org/10.1083/jcb.201405110).
65. Poleshko, A., Shah, P.P., Gupta, M., Babu, A., Morley, M.P., Manderfield, L.J., Ifkovits, J.L., Calderon, D., Aghajanian, H., Sierra-Pagán, J.E., et al. (2017). Genome-Nuclear Lamina Interactions Regulate Cardiac Stem Cell Lineage Restriction. *Cell* *171*, 573-587.e14. [10.1016/j.cell.2017.09.018](https://doi.org/10.1016/j.cell.2017.09.018).
66. Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* *6*, e21856. [10.7554/elife.21856](https://doi.org/10.7554/elife.21856).

67. Spracklin, G., and Pradhan, S. (2020). Protect-seq: genome-wide profiling of nuclease inaccessible domains reveals physical properties of chromatin. *Nucleic Acids Res.* 48, e16–e16. 10.1093/nar/gkz1150.
68. Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., Klein, A. de, Wessels, L., Laat, W. de, et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453, 948–951.
69. Finlan, L.E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J.R., and Bickmore, W.A. (2008). Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells. *PLoS Genet.* 4, e1000039. 10.1371/journal.pgen.1000039.
70. Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 452, 243–247.
71. Zullo, J.M., Demarco, I.A., Piqué-Regi, R., Gaffney, D.J., Epstein, C.B., Spooner, C.J., Luperchio, T.R., Bernstein, B.E., Pritchard, J.K., Reddy, K.L., et al. (2012). DNA Sequence-Dependent Compartmentalization and Silencing of Chromatin at the Nuclear Lamina. *Cell* 149, 1474–1487.
72. Leemans, C., Zwalm, M.C.H. van der, Brueckner, L., Comoglio, F., Schaik, T. van, Pagie, L., Arensbergen, J. van, and Steensel, B. van (2019). Promoter-Intrinsic and

Local Chromatin Features Determine Gene Repression in LADs. *Cell* 177, 852-864.e14.
10.1016/j.cell.2019.03.009.

73. Filion, G.J., and Steensel, B. van (2010). Reassessing the abundance of H3K9me2 chromatin domains in embryonic stem cells. *Nat Genet* 42, 4–4. 10.1038/ng01110-4.

74. Kurimoto, K., Yabuta, Y., Hayashi, K., Ohta, H., Kiyonari, H., Mitani, T., Moritoki, Y., Kohri, K., Kimura, H., Yamamoto, T., et al. (2015). Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells. *Cell Stem Cell* 16, 517–532. 10.1016/j.stem.2015.03.002.

75. Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of Enhancer Patterns in Transition from Naive to Primed Pluripotency. *Cell Stem Cell* 14, 838–853.
10.1016/j.stem.2014.04.003.

76. Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells. *Cell* 146, 519–532. 10.1016/j.cell.2011.06.052.

77. Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nature Genetics* 41, 246–250.

78. Zyllicz, J.J., Dietmann, S., Günesdogan, U., Hackett, J.A., Cougot, D., Lee, C., and Surani, M.A. (2015). Chromatin dynamics and the role of G9a in gene regulation and

enhancer silencing during early mouse development. *eLife* 4, e09571.
10.7554/elife.09571.

79. Solovei, I., Wang, A.S., Thanisch, K., Schmidt, C.S., Krebs, S., Zwerger, M., Cohen, T.V., Devys, D., Foisner, R., Peichl, L., et al. (2013). LBR and Lamin A/C Sequentially Tether Peripheral Heterochromatin and Inversely Regulate Differentiation. *Cell* 152, 584–598.

80. Smith, C.L., Lan, Y., Jain, R., Epstein, J.A., and Poleshko, A. (2021). Global chromatin relabeling accompanies spatial inversion of chromatin in rod photoreceptors. *Sci. Adv.* 7, eabj3035. 10.1126/sciadv.abj3035.

81. Wang, Y., Elsherbiny, A., Kessler, L., Cordero, J., Shi, H., Serke, H., Lityagina, O., Trogisch, F.A., Mohammadi, M.M., El-Battrawy, I., et al. (2022). Lamin A/C-dependent chromatin architecture safeguards naïve pluripotency to prevent aberrant cardiovascular cell fate and function. *Nat. Commun.* 13, 6663. 10.1038/s41467-022-34366-7.

82. Poleshko, A., Mansfield, K.M., Burlingame, C.C., Andrade, M.D., Shah, N.R., and Katz, R.A. (2013). The Human Protein PRR14 Tethers Heterochromatin to the Nuclear Lamina during Interphase and Mitotic Exit. *CellReports* 5, 292–301.

83. Robson, M.I., Heras, J.I. de las, Czapiewski, R., Thành, P.L., Booth, D.G., Kelly, D.A., Webb, S., Kerr, A.R.W., and Schirmer, E.C. (2016). Tissue-Specific Gene Repositioning by Muscle Nuclear Membrane Proteins Enhances Repression of Critical Developmental Genes during Myogenesis. *Mol. Cell* 62, 834–847.
10.1016/j.molcel.2016.04.035.

84. Ye, Q., Callebaut, I., Pezhman, A., Courvalin, J.C., and Worman, H.J. (1997). Domain-specific interactions of human HP1-type chromodomain proteins and inner nuclear membrane protein LBR. *Journal of Biological Chemistry* 272, 14983–14989.
85. Liokatis, S., Edlich, C., Soupsana, K., Giannios, I., Panagiotidou, P., Tripsianes, K., Sattler, M., Georgatos, S.D., and Politou, A.S. (2012). Solution Structure and Molecular Interactions of Lamin B Receptor Tudor Domain*. *J Biol Chem* 287, 1032–1042. 10.1074/jbc.m111.281303.
86. Hirano, Y., Hizume, K., Kimura, H., Takeyasu, K., Haraguchi, T., and Hiraoka, Y. (2012). Lamin B Receptor Recognizes Specific Modifications of Histone H4 in Heterochromatin Formation*. *J. Biol. Chem.* 287, 42654–42663. 10.1074/jbc.m112.397950.
87. Yan, Z., Ji, L., Huo, X., Wang, Q., Zhang, Y., and Wen, B. (2020). G9a/GLP-sensitivity of H3K9me2 Demarcates Two Types of Genomic Compartments. *Genom., Proteom. Bioinform.* 18, 359–370. 10.1016/j.gpb.2020.08.001.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

Eric Simental

7631DAAC6443419...

Author Signature

3/20/2024

Date