Title

Enhancer Cis Elements Instruct Enhancer-Promoter Interaction, Enhancer-Promoter Insulation and Compartmental Segregation

Permalink

https://escholarship.org/uc/item/94r5f4s2

Author

Lu, Hanbin

Publication Date

2022

UNIVERSITY OF CALIFORNIA SAN DIEGO

Enhancer Cis Elements Instruct Enhancer-Promoter Interaction, Enhancer-Promoter Insulation and Compartmental Segregation

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Biology

by

Hanbin Lu

Committee in charge:

Professor Cornelis Murre, Chair
Professor Ken W. Y. Cho
Professor Xiang-Dong Fu
Professor Christopher K. Glass
Professor Clodagh O'Shea

2022

The Dissertation of Hanbin Lu is approved, and it is acceptable in
quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

In loving memory of my father and my cat Chloe

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Specially thanks to my wife Paula Pham and my family. They love me and support my career. My wife also significantly contributed to the projects presented in this thesis.

I want to thank my thesis advisor Dr. Cornelis Murre for mentoring, encouraging, and supporting me during the time pursuing my PhD degree. He taught me about scientific research and critical thinking. Kees supervised and oversaw all the work presented here.

I want to thank my colleague Dr. Yina Zhu who trained me wet lab skills. She had shared with me the invaluable experiences and insights that will benefit all times in my career.

I want to thank Dr. Ken Cho for his support and collaborations. He supervised the project described in chapter 5.

I want to thank all my colleagues from the Murre lab, especially Matthew Denholtz; Jeff Jiajing Zhou from the Cho lab; and Han Han, Wenqi Wang from the Wang lab.

I also want to thank all my friends for supporting and cheering me during the rough times in my life.

Chapter 2 and 3 are currently prepared for submission for publication authored by Hanbin Lu, Paula Pham, Han Han, Jeff Jiajing Zhou, Yina Zhu, Wenqi Wang, Ken W. Y. Cho and Cornelis Murre.

Chapter 4.1 reprinted the paper in *Genes & Development*, 2020. “Upon microbial challenge, human neutrophils undergo rapid changes in nuclear architecture and chromatin folding to orchestrate an immediate inflammatory gene program”. Matthew Denholtz, Yina Zhu, Zhaoren He, Hanbin Lu, Takeshi Isoda, Simon Döhrmann, Victor Nizet and Cornelis Murre.

Chapter 4.2 reprinted another paper in *Genes & Development*, 2020. “Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization”. Yina Zhu, Matthew Denholtz, Hanbin Lu and Cornelis Murre.

Chapter 5 is currently prepared for submission for publication authored by Paula Pham, Hanbin Lu, Han Han, Jeff Jiajing Zhou, Wenqi Wang, Ken W. Y. Cho and Cornelis Murre.

VITA

EDUCATION

2022 Doctor of Philosophy, Biology, University of California San Diego

2013 Bachelor of Science, Biology, Sun Yat-Sen University

PUBLICATIONS

9. Pham, P.D.*, **Lu, H.***, Han, H., Zhou, J.J., Wang W., Murre, C., & Cho, K.W.Y. A multifaceted network of transcription factors enforces lineage plasticity in emerging extraembryonic endoderm cells. (Manuscript In Preparation).

8. **Lu, H.***, Pham, P.D.*, Han, H., Zhou, J.J., Zhu, Y., Wang W., Cho, K.W.Y., & Murre, C. Enhancer Cis Elements Instructs Promoter-Enhancer Interaction, Promoter-Enhancer Insulation and Compartmental Segregation. (Manuscript In Preparation).

7. Zhu, Y., Denholtz, M., **Lu, H**., & Murre, C. (2021). Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization. Genes & Development, 35(65–81).

6. Denholtz, M., Zhu, Y., He, Z., **Lu, H**., Isoda, T., Döhrmann, S., Nizet, V., & Murre, C. (2020). Upon microbial challenge, human neutrophils undergo rapid changes in nuclear architecture and chromatin folding to orchestrate an immediate inflammatory gene program. Genes & Development, 34(3-4).

5. Barajas-Mora, E. M., Kleiman, E., Xu, J., Carrico, N. C., **Lu, H**., Oltz, E. M., ..., & Feeney, A. J. (2019). A B-cell-specific enhancer orchestrates nuclear architecture to generate a diverse antigen receptor repertoire. Molecular Cell, 73(1), 48-60.

4. Rahnamoun, H., Lee, J., Sun, Z., **Lu, H**., Ramsey, K. M., Komives, E. A., & Lauberth, S. M. (2018). RNAs interact with BRD4 to promote enhanced chromatin engagement and transcription activation. Nature Structural & Molecular Biology, 25(8), 687-697.

3. Rahnamoun, H., Hong, J., Sun, Z., Lee, J., **Lu, H**., & Lauberth, S. M. (2018). Mutant p53 regulates enhancer-associated H3K4 monomethylation through interactions with the methyltransferase MLL4. Journal of Biological Chemistry, 293(34), 13234-13246.2. Su, S.,

2. Shao, X., Zhu, C., Xu, J., **Lu, H**. ,…, Huang, X. (2018) Transcriptome-wide analysis reveals the origin of peloria in Chinese Cymbidium (Cymbidium sinense). Plant Cell Physiol., 59, 2064– 2074.

1. Rahnamoun, **H., Lu**, H., Duttke, S. H., Benner, C., Glass, C. K., & Lauberth, S. M. (2017). Mutant p53 shapes the enhancer landscape of cancer cells in response to chronic immune signaling. Nature Communications, 8(1), 1-14.

*** co-first authors**

ABSTRACT OF THE DISSERTATION

Enhancer Cis Elements Instruct Enhancer-Promoter Interaction, Enhancer-Promoter Insulation and Compartmental Segregation

by

Hanbin Lu

Doctor of Philosophy in Biology

University of California San Diego, 2022

Professor Cornelis Murre, Chair

Enhancers harbor instructions encoded for the interactions between cis-elements and transcription factors to orchestrate lineage specific gene programs. Here we developed a modified method for chromosome conformation capture (3C), named MID Hi-C, to reveal how in mouse embryonic stem cells differential cooperation of enhancers and the chromatin remodeler BAF, as instructed by the underlying transcription factor motifs, modulate enhancer-promoter communication. We show that BAF-dependent enhancers permit genomic interactions beyond enhancer boundaries. BAF-dependent enhancers do not dictate genomic interactions within enhancer-promoter loop domains but rather act to instruct remote enhancer-promoter communication. In contrast, BAF-independent enhancers interact with promoter regions within tightly insulated enhancer-promoter loop domains that are marked by promoter and enhancer boundary elements. In addition, enhancer activeness modulated by BAF enforces compartment segregation. Based on these observations, we propose that enhancer cis elements instruct with great precision BAF-induced enhancer-promoter communication and compartmental segregation. We also characterize the gene programs and enhancer landscapes that instructed the cell fate bifurcation in the extraembryonic endoderm lineage orchestrated by lineage specific transcription factors. The simplicity of this developmental system will be valuable for dissecting how enhancer cis elements exploit the 3D genome principles to inscribe proper cell differentiation.

CHAPTER 1

Introduction to 3D Genome Organization

1.1 High-throughput Chromosome Conformation Capture (3C) Technology

The early studies in chromosome structure (Münkel and Langowski, 1998; Paulson and Laemmli, 1977; Rattner and Lin, 1985; Sedat and Manuelidis, 1978) and the three-dimensional (3D) structure of gene loci, such as at the globin genes (Baù et al., 2011; Tolhuis et al., 2002; Zhou et al., 2006) and the immunoglobulin heavy-chain locus (*Igh*) (Guo et al., 2011; Jhunjhunwala et al., 2008), generated a broad picture that the genome is folded in sophisticated 3D configurations. Since then, the maturing of chromosome conformation capture (3C) technology, such as 3C (Dekker et al., 2002), 4C (Simonis et al., 2006), as well as 5C (Dostie et al., 2006) assays became the state-of-the-art tool in mapping chromatin interactions. Despite the increasing throughputs of 3C method variants, only a minute fraction of chromatin interactions in the nucleus were sampled in a single run. Therefore, the enormous complexity of the 3D genome was still beyond reach. It was not until the arrival of the era of next-generation sequencing and the innovation of high-throughput chromosome conformation capture (Hi-C) assay (Lieberman-Aiden et al., 2009), as well as its mature version, in situ Hi-C (Rao et al., 2014), that allowed one to map genome-wide chromatin interactions in an unbiased manner (‘all-to-all’).

Hi-C assay is experimentally simple and inherently robust. In brief, crosslinked chromatin (nuclei) is subjected to fragmentation by the restriction enzyme (RE) followed by the proximal ligation to capture the physical proximity information from chromatin interactions (Lajoie et al., 2015; Lieberman-Aiden et al., 2009; Rao et al., 2014). Numerous endeavors, including the 4D nucleome project (https://www.4dnucleome.org/; Dekker et al., 2017), were directed to utilize Hi-C to comprehensively map the 3D genome in various biological contexts (Bonev et al., 2017; Cuartero et al., 2018; Denholtz et al., 2020; Gibcus et al., 2018; Isoda et al., 2017; Kieffer-Kwon et al., 2013, 2017; Lin et al., 2012; Link et al., 2018; Niu et al., 2021; Yang et al., 2020; Zheng and Xie, 2019).

Systematic survey of chromatin interactions by Hi-C assay revealed elaborate architectures by chromatin folding. We now know that chromatin is folded into compartments, self-interacting topological associated domains (TAD), CTCF loops, and transcriptional-element-engaged interactions (Dixon et al., 2012; Lieberman-Aiden et al., 2009; Mifsud et al., 2015; Nora et al., 2012; Phillips-Cremins et al., 2013; Rao et al., 2014; Sexton et al., 2012). These features are prominent in the Hi-C contact maps at low (100 ~ 500 KB), median (25 ~ 100 KB), median high (2 ~ 25 KB), and super high (~ 1 KB) resolutions respectively (see more details in Chapter 1.2; Jerkovic´ and Cavalli, 2021). Although there are growing interests to investigate chromatin interactions among cis-regulatory elements, these missions were proven to be challenging because the chromatin interaction space at the required super high resolution is immense, and thus demands a very deep sequencing (> 2 billion reads; Bonev et al., 2017; Lajoie et al., 2015; Rao et al., 2014). To improve cost efficiency, enrichment strategies, such as target capture via immunoprecipitation or DNA/RNA probes as implemented by HiChIP (Mumbach et al., 2016) or (NG) Capture-C (Davies et al., 2016; Mifsud et al., 2015), were applied to the Hi-C library to prioritize sequencing coverage at genomic regions of interest.

However, the performances of the aforementioned strategies are discounted by the inherent inefficacy of Hi-C method in capturing high-resolution chromatin interactions (Hsieh et al., 2020, 2021; Krietenstein et al., 2020; Oksuz et al., 2021). Theoretically, the highest resolution of in situ Hi-C assay is 256 BP, determined by the expected cutting frequency of 4-bp restriction enzyme (RE) used in chromatin fragmentation. In practice, proteins coated on chromatin tamper with the efficiency of fragmentation, which in turn impairs the resolution. To overcome incomplete chromatin digestion, a new 3C-based method variant, called Micro-C, was developed (Hsieh et al., 2016, 2020; Krietenstein et al., 2020). By using Mnase as a surrogate for RE, Micro-C assay successfully maps chromatin interactions in mammalian genome down to 200 BP resolution, approximate to the size of a single nucleosome yielded from Mnase digestion. However, like other Mnase-based assays, Micro-C demands delicate experimental controls to constrain the hyper-activity of Mnase, since over-digestion damages DNA fragments while under-digestion raises noise (Hsieh et al., 2020). Due to this technical barrier, the adoption of Micro-C assay by the wider genomics field did not attain as much momentum as Hi-C assay used to gain.

In consequence, a pressing challenge is to develop a manageable method, like Hi-C, to scale up efforts in mapping chromatin interaction at super high resolutions. To address the challenge, we developed a new Hi-C approach, termed MID Hi-C, featuring high performance in chromosome conformation capture like Micro-C, as well as experimental simplicity and robustness like Hi-C (Chapter 2). Given sufficient sequencing depth, MID Hi-C can generate feature-rich contact maps down to ~500 BP resolutions.

1.2 Chromatin 3D Architectures

In the past decade, chromatin 3D architectures of different genomic scales have been discovered (Dixon et al., 2012; Lieberman-Aiden et al., 2009; Mifsud et al., 2015; Nora et al., 2012; Phillips-Cremins et al., 2013; Rao et al., 2014; Sexton et al., 2012). The mechanisms driving their formations are under active research. At the mega-base scale, large spans of chromatin are segregated into two mutually repulsive clusters, termed A and B compartments, appearing as binary checkerboard patterns on contact maps at 100 to 500 KB resolution (Lieberman-Aiden et al., 2009). Genomic regions in A compartment are generally positioned toward the nuclear interior, e.g., nuclear speckle associating domains (SPADs), and associated with euchromatin in transcriptionally active states that are enriched for active histone marks like H3K27ac (Bickmore and van Steensel, 2013; Lieberman-Aiden et al., 2009; Rao et al., 2014; Zhang et al., 2020). Conversely, genomic regions in B compartment are generally positioned in the nuclear periphery, e.g., nuclear lamina associating domains (LADs), and associated with heterochromatin in transcriptionally inert states that are enriched for repressive histone marks like H3K9me3 (Bickmore and van Steensel, 2013; Lieberman-Aiden et al., 2009; Rao et al., 2014; Zhang et al., 2020).

Recent studies have shown that heterochromatin protein 1a (HP1a), ‘reader’ of H3K9me3, undergoes ‘liquid-liquid demixing’ and forms droplets in solution under molecular crowding environments (Larson et al., 2017; Strom et al., 2017). Increasing salt concentration effectively dissolved the droplets. The self-aggregation forces arise from weak, but multivalent interactions among HP1a proteins, mediated by the intrinsically disordered regions (IDR) embedded in the protein domains of HP1a as revealed by biophysics analyses (Larson et al., 2017; Strom et al., 2017). The phenomenon is termed liquid-liquid phase separation (LLPS) which drives the self-assembly of diverse membrane-less organelles, such as stress granules in cytoplasm and nucleoli in the nucleus (Boeynaems et al., 2018). Since heterochromatin is heavily coated by HP1a, LLPS is proposed to drive the condensation of heterochromatin and physically sequester active chromatin (Larson et al., 2017; Strom et al., 2017). Likewise, many transcriptional activators enriched in euchromatin can mediate LLPS as well (Boija et al., 2018; Gibson et al., 2019; Lu et al., 2018; Sabari et al., 2018). Consequently, the segregation of A and B compartments manifests LLPS phenomena in chromatin.

At the sub-mega-base scale, local chromatin is organized by self-interacting TADs (Dixon et al., 2012; Lieberman-Aiden et al., 2009; Nora et al., 2012; Phillips-Cremins et al., 2013; Rao et al., 2014; Sexton et al., 2012). These chromatin domains can be maintained for many cell cycles as well as conserved across distinct cell types and even across related species. In many cases, TAD is encapsulated in a CTCF loop linking the boundaries occupied by converging CTCF sites. Therefore, they are referred to sometimes as loop domains or contact domains (Rao et al., 2014).

CTCF loops and TADs are proposed to be generated by dynamic loop extrusion (Fudenberg et al., 2016; Sanborn et al., 2015). Specifically, loop extrusion process is orchestrated by architecture proteins (complexes) comprising CTCF, cohesin, NIPBL, and WAPL. Underlying the loop extrusion model, cohesin ring, which is loaded onto chromatin by NIPBL, entraps a stretch of DNA that gives birth to the initial DNA loop. Then, cohesin extrudes and enlarges the loop by moving along the chromatin in opposite directions. Extruding cohesin can be halted by a pair of CTCF binding sites, which in turn forms CTCF loops. The loop remains closed until cohesin is released from chromatin by WAPL and recycled. Loop extrusion model accurately predicted the loop defects caused by acute depletion of individual architecture proteins. For example, depletion of CTCF resulted in reversible and CTCF-dose-dependent loss of CTCF loops as well as insulation at TAD (loop domain) boundaries (Nora et al., 2017). On the other hand, removing cohesin from chromatin either through depletion of cohesin or NIPBL abolished all CTCF loops across the genome. A good portion of the disrupted loops, some may even span mega-base sizes, recovered in an hour when cohesin proteins were restored (Rao et al., 2017; Schwarzer et al., 2017). In contrast, depletion of WAPL resulted in longer dwelling time of cohesin on chromatin, which gave rise to ectopically enlarged loops (Haarhuis et al., 2017).

Beside illuminating the mechanism of CTCF loop formation, these studies shed light on the functional relations between CTCF loop and chromatin compartmentalization (also gene regulation, see more detail in Chapter 1.3). Neither genome-wide removal nor elongation of CTCF loops in the architectural protein depletion studies produced wide-spread alterations in compartmental organizations (Haarhuis et al., 2017; Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017). Therefore, compartmentalization and CTCF loops are independent layers for organizing chromatin at the global scales. So do the underlying LLPS and loop extrusion. However, in the local chromatin organizations, the subtle changes in compartmental interaction patterns indicated that compartmentalization and loop extrusion may be counteracting. For example, NIPBL depletion unmasked small B-like regions within A compartment (Schwarzer et al., 2017). They were usually devoid of active histone marks as opposed by the surrounding A compartment regions. Similarly, cohesin depletion causes super-enhancer to colocalize (Rao et al., 2017). Both the appearance of B-like regions and colocalization of super-enhancer are correlated to enhancer compartmental segregation, suggesting that loop extrusion can act against finer compartmentalization.

Despite mounting evidence pointing to the existence of loop extrusion, how cohesin moves along the chromatin is still unclear. RNA polymerase is thought to be one of the motors since convergent transcriptions accumulate cohesin at the intersecting 3' end, forming so-call cohesin islands, when the primary cohesin-position regulators, CTCF and WAPL, are deleted (Busslinger et al., 2017). Recently, in vitro studies showed that the ATPase activity of cohesin can consumes ATP and translocates cohesin in steps (Kim et al., 2020; Pradhan et al., 2021; Ryu et al., 2021). Consistently, cohesin assembled with SMC3 ATPase mutations demonstrates 50% less efficiency in translocating to CTCF sites than cohesin with WT SMC3 in vivo (Vian et al., 2018). Furthermore, configuration of 'stripe' domains (Vian et al., 2018) as well as the imminent connections between IgH recombination and loop extrusion are compelling evidence that cohesin scans the entire loop domain (Ba et al., 2020; Dai et al., 2021; Hill et al., 2020; Jain et al., 2018; Zhang et al., 2019a, 2019b). However, direct observations of loop extrusion in vivo are still needed to integrate the multiline evidence to a unified model.

At even finer scales, long range interactions engaged by transcriptional elements, like enhancer-promoter (E-P) interactions, promoter-promoter (P-P) interactions, and gene loops, etc., are prevalent as revealed by promoter capture-C (Mifsud et al., 2015; Schoenfelder et al., 2015) and Micro-C (Hsieh et al., 2020, 2021). A recent study reported that P-P and E-P interactions can cross TAD boundaries, renewing our understanding about CTCF loop insulation (Hsieh et al., 2021). However, a full characterization of the fine-scale chromatin interactions remains to be worked out using more comprehensive ‘all-to-all’ 3C tools, such as the newly developed Micro-C and MID Hi-C.

Our understanding of the mechanisms driving the formation of E-P interactions is rudimental relative to that of compartmentalization and CTCF loops. Since both cohesin and LLPS-competent factors (e.g., BRD4 and mediator) are often found in enhancers and promoters, loop extrusion and LLPS are speculated to participate in establishing E-P interactions (Boija et al., 2018; Hnisz et al., 2017; Isoda et al., 2017; Khattabi et al., 2019; Linares-Saldana et al., 2021; Liu et al., 2021; Sabari et al., 2018). However, neither depleting LLPS-competent factors (Crump et al., 2021; Jaeger et al., 2020; Khattabi et al., 2019) nor architecture proteins (Hsieh et al., 2021; Thiecke et al., 2020) that mediate loop extrusion had abolished E-P interactions. Thus, the mechanism is awaiting to be unveiled.

1.3 Impact of Chromatin 3D Architectures in Gene Regulation

Whether and how chromatin structures influence gene expression is currently under intense debate. Numerous studies have shown that tissue-specific gene expressions are correlated with changes in chromatin conformations (Bonev et al., 2017; Stadhouders et al., 2018). At the level of compartmentalization, cell-type determination factors were repositioned from nuclear lamina to nuclear interior during lineage commitments, such as in the cases of Ebf1 and Bcl11b during B cell and T cell lineage choice (Isoda et al., 2017; Lin et al., 2012). TAD and CTCF loop facilitate proper communication between enhancer and promoter by restricting enhancer from contacting non-cognate promoters (Dixon et al., 2016; Laat and Duboule, 2013). Indeed, several reports shown that enhancer-hijacking events triggered by disruption in TAD boundaries contributes to developmental defects and cancer (Franke et al., 2016; Hnisz et al., 2016; Ibn-Salem et al., 2014; Lupiáñez et al., 2015). In line with these observations, one would expect that genome-wide removal of TAD will cause vast transcriptional changes. However, all architecture protein depletion experiments mentioned above have seen only minor disruptions on gene expressions (Haarhuis et al., 2017; Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017). Although the changes in local chromatin interaction of a handful altered genes when depleting RAD21 and WAPL suggested that some enhancers may need cohesin to establish interactions with target promoters (Liu et al., 2021; Rao et al., 2017), the majority of E-P interactions were not affected by removal of architecture proteins (Hsieh et al., 2021; Thiecke et al., 2020). These observations could explain why only minor transcriptional changes occur when TADs and CTCF loops are disrupted.

Nevertheless, the argument assumes that E-P interactions play a central role in regulating gene expression. In fact, forced enhancer-promoter interactions increased the transcriptional output because of elevated transcriptional bursting (Deng et al., 2012; Fukaya et al., 2016; Morgan et al., 2017). Two recent studies shown that the formation of E-P interacting hub preceded TAD formation and cell-type-specific gene expression during Drosophila embryo development (Espinola et al., 2021; Ing-Simmons et al., 2021), suggesting that E-P interactions per se are less developmentally controlled as we previously thought. Furthermore, other studies reported that some genes do not require enhancer proximity for their activations (Alexander et al., 2019; Benabdallah et al., 2019). More future work needs to be done to understand how E-P interaction regulates gene expression.

1.4 Layout of Contents

In the second chapter of this thesis work, we developed a new and powerful Hi-C approach, named MID Hi-C, to comprehensively map E-P interactions. We benchmarked the performance of MID Hi-C against that of Micro-C and in situ Hi-C. Then, we demonstrated that MID Hi-C can be easily repurposed to upgrade existing Hi-C based methods. As a proof of principle, we developed MID HiChIP. We observed dramatical enhancement as compared to HiChIP.

In the third chapter, we used MID Hi-C to investigate enhancer-dependent effects in altering the status of the active (A) compartment as well as E-P interactions when the BAF complex was perturbed.

In the fourth chapter, I included two published studies that investigated chromatin structural changes in the mouse and human neutrophils stimulated by calcium signaling or microbial challenges.

In the fifth chapter, we studied how lineage-determining transcription factors cooperate with cis-regulatory elements to instruct antagonistic cell fates in the extraembryonic lineage during early mouse embryo development.

The last chapter is the concluding remarks.

CHAPTER 2

Development of MID Hi-C Method

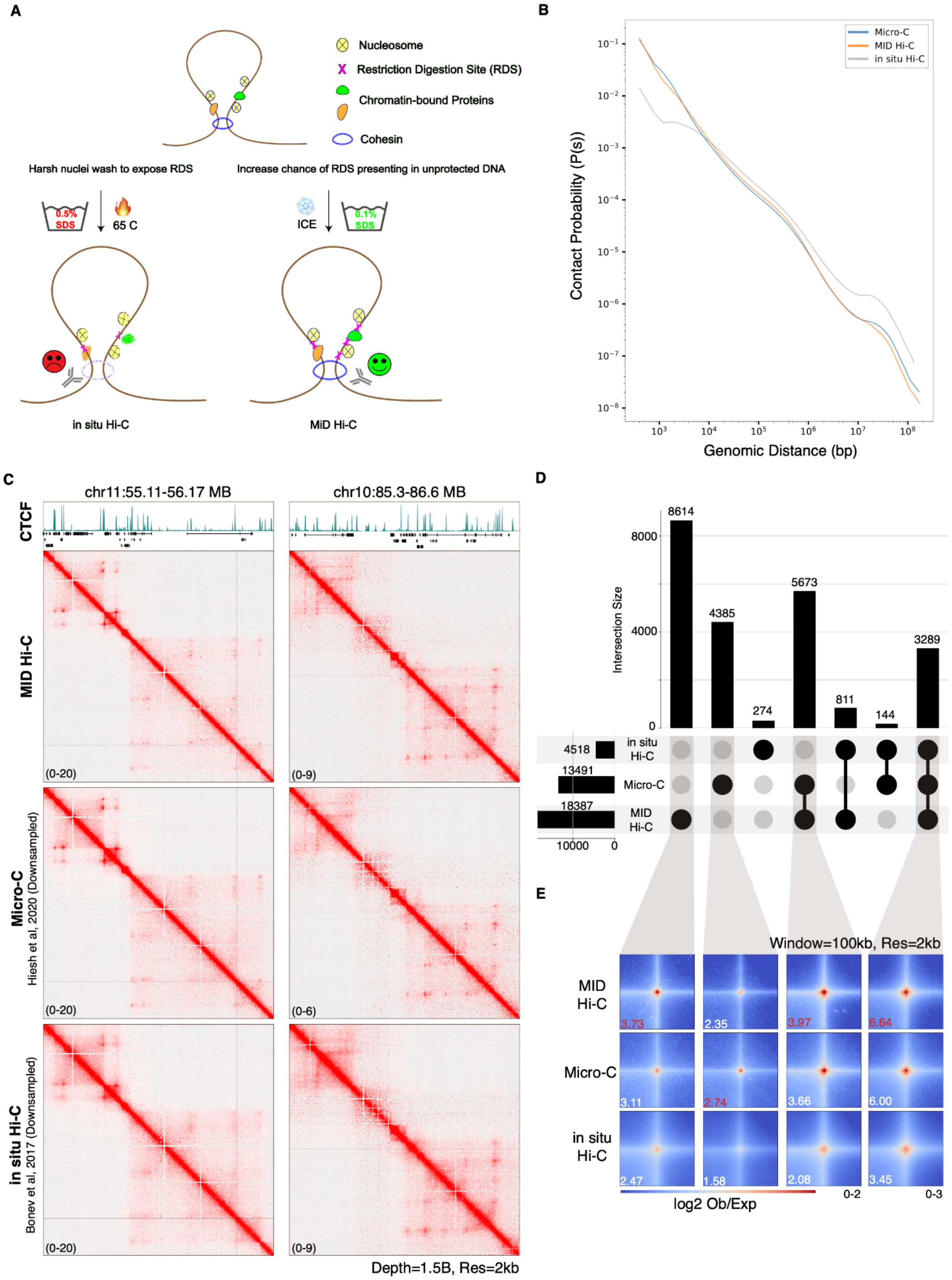
In this chapter, we presented a new Hi-C approach, termed MID Hi-C, that supersedes in situ Hi-C, the most widely used chromosome conformation capture (3C) tool nowadays, to map chromatin interactions in an unbiased manner. We first discussed the critical optimizations in MID Hi-C assay. We then benchmarked the performance of MID Hi-C in profiling CTCF loops and enhancer-promoter interactions against Micro-C, the state-of-the-art 3C tool that we are competing with to replace in situ Hi-C. Finally, we utilized MID Hi-C to upgrade the HiChIP protocol and developed MID HiChIP.

2.1 Rational of MID Hi-C

Recent studies developed optimized Hi-C protocols, named Micro-C and Hi-C3.0, that enable detection of DNA loops with relatively high signal-to-noise ratios (Oksuz et al., 2021). Micro-C revealed two major advantages over in situ Hi-C (Rao et al., 2014). First, Micro-C enhanced focal enrichment of looping signals. Second, Micro-C detects more loops in higher resolution when compared to in situ Hi-C, in particular for genomic regions encoding enhancers and promoters (Hsieh et al., 2016, 2020; Krietenstein et al., 2020). Another approach, named Hi-C 3.0, upgraded in situ Hi-C protocol by applying a second crosslinker and double restriction enzyme (RE) digestion. Hi-C 3.0 provides increased experimental simplicity relative to Micro-C. However, despite considerable increase in signal-to-noise ratios compared to in situ Hi-C, Hi-C 3.0 still significantly underperformed Micro-C in DNA loop detection (Oksuz et al., 2021). Hence, we tried to further improve Hi-C 3.0 by using more frequently digested restriction enzymes. Specifically, we used Mse1 and Dde1 to digest double-crosslinked chromatin (Figure 2.2A). We found that our updated approach still underperformed Micro-C in detecting DNA loops (data not shown).

To optimize the conditions further, we explored the possibility that the relatively harsh denaturing wash (0.5% SDS at 65 C for 10 min), when preparing nuclei for efficient chromatin fragmentation, could impair the preservation of chromatin interactions as a consequence of reverse-crosslinking. We reasoned that by applying Mse1 and Dde1 as restriction enzymes, the harsh denaturing nuclei wash might not be required for efficient chromatin fragmentation because Mse1 and Dde1 recognition sites are frequently present across the genome including genomic regions that are depleted for nucleosome and thus can be accessed by the enzymes (Figure 2.1A). Indeed, by using Mse1 and Dde1 as fragmentation enzymes, we were able to fragment chromatin to suitable sizes (< 1 kb) in a ‘nearly-denature-free’ condition (Figure 2.2B). Hence, we named this method as minimized interruption and double-digestion (MID) Hi-C.

Figure 2.1. Robust detection of CTCF loops by MID Hi-C. (A) Schematic showing chromatin can be fragmented without the denaturing nuclei wash if there are high chances of RE recognizable motifs presenting in unprotected regions. (B) Comparison of the contact scaling curves. (C) Comparison of the contact maps. (D) Intersection analysis of CTCF loops called from the datasets of the three methods using HiCCUPS. (E) APA for the loop subsets in (D) (see also in Figure 2.2B). The enrichment scores were calculated by dividing the intensity of the central pixel with the average intensity of the 3x3 pixels at bottom left. The highest enrichment scores for each loop subset (column-wise max) were highlighted in red.


A
Nucleosome
Restriction Digestion Site (RDS)
Chromatin-bound Proteins
Cohesin
Harsh nuclei wash to expose RDS
Increase chance of RDS presenting in unprotected DNA
0.5% SDS
65 C
ICE
0.1% SDS
in situ Hi-C
MiD Hi-C
B
Micro-C
MID Hi-C
in situ Hi-C
Contact Probability (P(s))
Genomic Distance (bp)
C
chr11:55.11-56.17 MB
chr10:85.3-86.6 MB
CTCF
MID Hi-C
Micro-C
Hiesh et al, 2020 (Downsampled)
in situ Hi-C
Bonev et al, 2017 (Downsampled)
(0-20)
(0-9)
(0-6)
Depth=1.5B, Res=2kb
D
Intersection Size
8614
4385
274
5673
811
144
3289
4518
13491
18387
in situ Hi-C
Micro-C
MID Hi-C
E
Window=100kb, Res=2kb
MID Hi-C
Micro-C
in situ Hi-C
3.73
2.35
3.97
6.64
3.11
2.74
3.66
6.00
2.47
1.58
2.08
3.45
log2 Ob/Exp
0-2
0-3

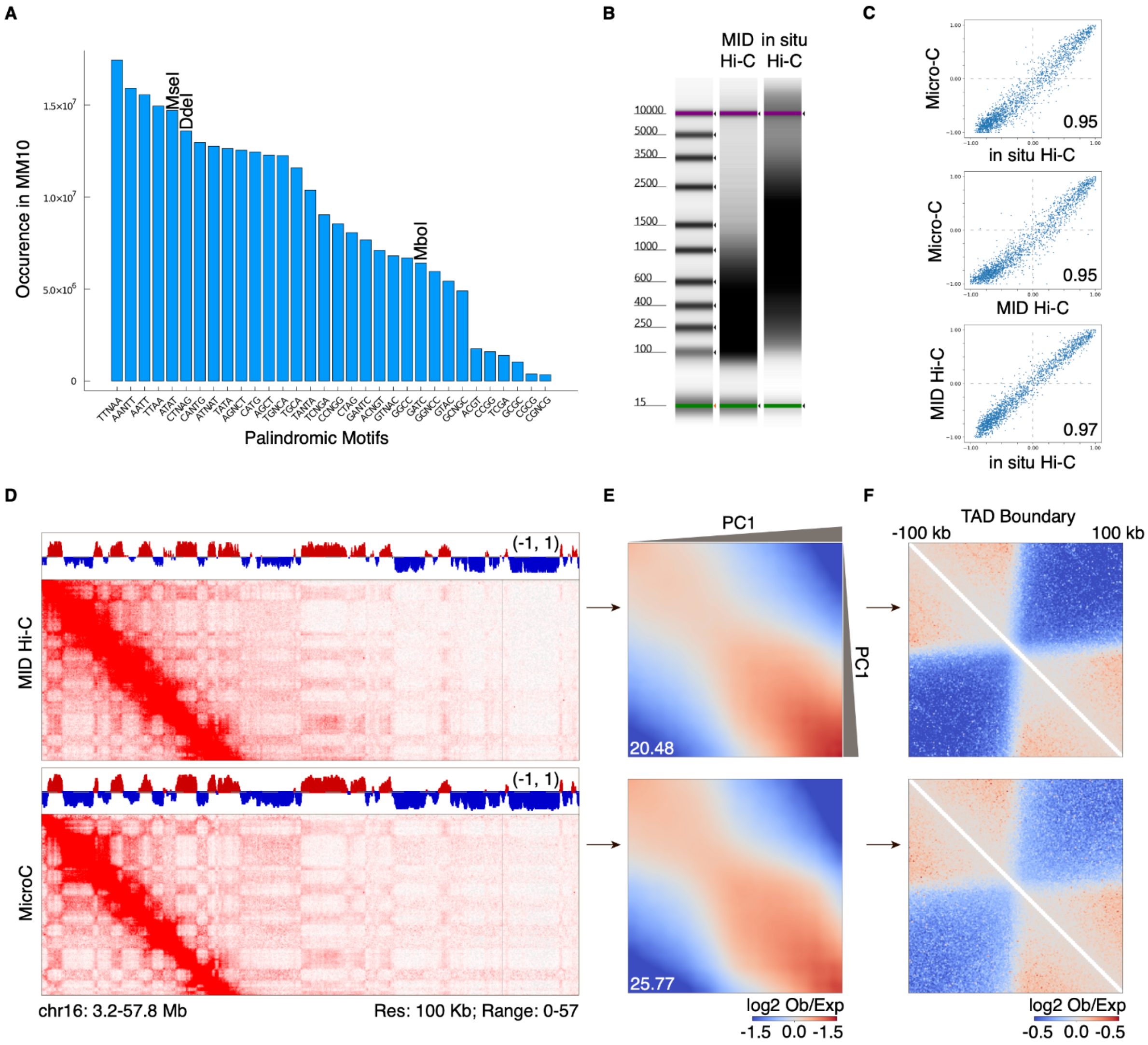
2.2 MID Hi-C Recapitulates Compartmentalization and TAD with Lower Noise

To determine the quality of MID Hi-C in mapping chromatin interactions, we generated a MID HI-C library (1.5 billion interaction reads) for mouse embryonic stem cells (mESCs). We next evaluated MID Hi-C performance by comparing it to deeply sequenced Micro-C (Hsieh et al., 2020) and in situ Hi-C (Bonev et al., 2017) datasets derived from mESCs. As a first approach we examined the scaling patterns (decay of contact frequencies as a function of genomic distance) for the three datasets. We found the scaling pattern of MID Hi-C resembled that of Micro-C (Figure 2.1B). Notably, the noise level associated with MID Hi-C, measured by the contact frequency at far cis (> 100 MB), was substantially lower when compared to in situ Hi-C (Figure 2.1B). In comparison to Micro-C, MID Hi-C noise levels were slightly lower (Figure 2.1B). Interaction frequencies involving paired genomic elements separated between 5 kb and 10 Mb, were virtually identical for Micro-C and MID Hi-C (Figure 2.1B).

To determine whether MID Hi-C recapitulates the features of compartmentalization and TADs, contact matrices of MID Hi-C were generated and compared to those derived for in situ HiC and Micro-C. We found that the assembly of compartments as revealed using the aforementioned three protocols were highly similar (Figure 2.2C, spearman correlation >0.945). The 'checkerboard' pattern as revealed in contact matrices using MID Hi-C was slightly weaker than that observed for Micro-C (Figure 2.2D). A similar trend was observed by analyzing compartmental strength using saddle plots (Figure 2.2E; Gibcus et al., 2018). We note that these differences between Micro-C and MID Hi-C plausibly relates to a steeper decline in contact coverages for paired genomic elements spanning vast genomic distances (>10 Mb) as observed for MID Hi-C (Figure 2.1B). To evaluate detection of TADs using MID Hi-C, we aggregated the contact maps of MID Hi-C and Micro-C surrounding the TAD boundaries detected from Micro-

C data using TopDom (Shin et al., 2016). We found an enhanced contrast between intra- vs. inter-TAD interaction strengths for MID Hi-C as compared to Micro-C, unveiling the boundary patterns associated with TADs (Figure 2.2F). Taken together, this analysis indicates that Micro-C and MID Hi-C protocols are equivalent in terms of identifying TADs and compartmentalization at a genome-wide level.

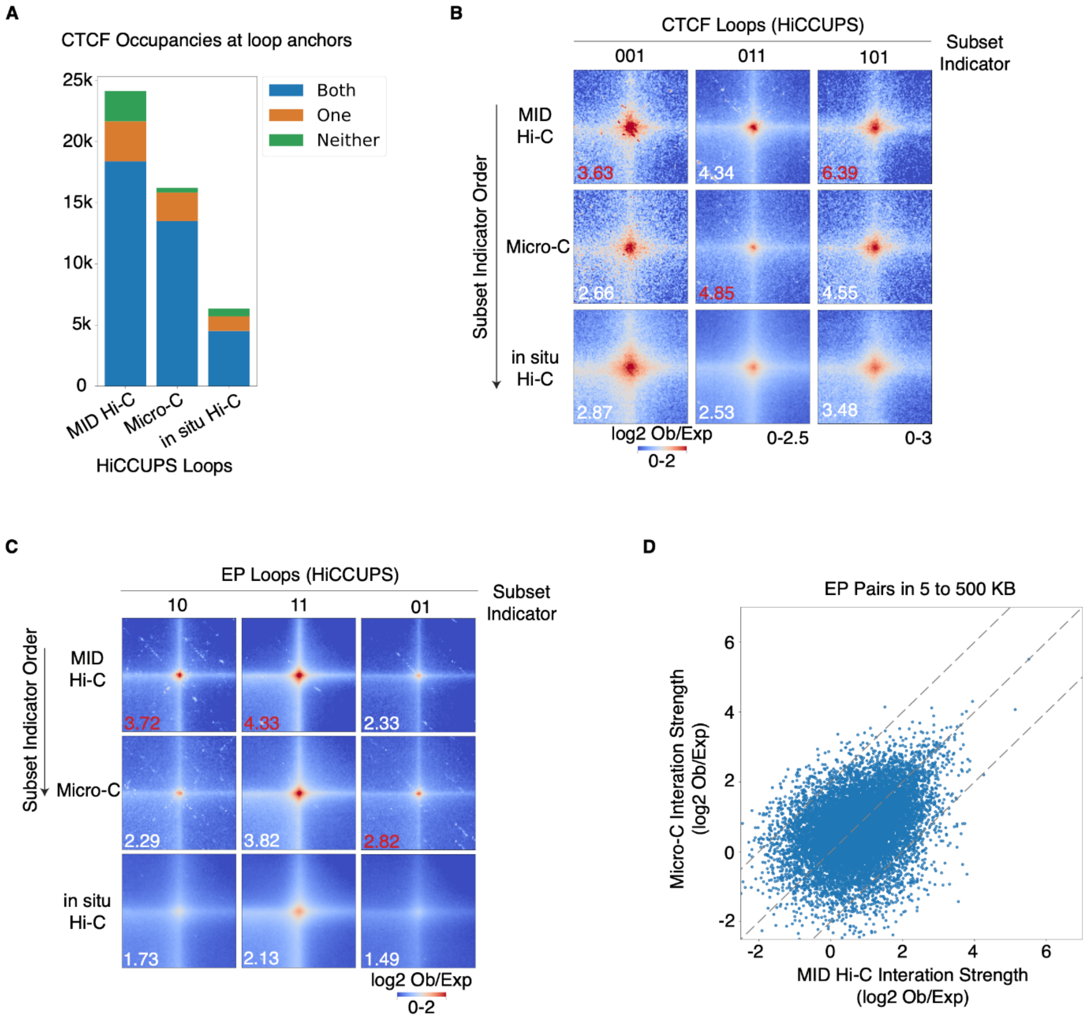
Figure 2.2. Comparable detection of compartments and TADs by MID Hi-C and Micro-C. (A) Numbers of palindromic (4 and 4 plus bp) motifs occur in the mouse genome. (B) Tapestation profiling of the fragment length distribution for the digested chromatin. (C) Scatter plot displaying the pairwise comparisons of compartment scores among MID Hi-C, in situ Hi-C and Micro-C. Bin size was 25 KB. (D) Contact maps showing ‘checkerboard’ pattern (compartmentalization) detected by MID Hi-C and Micro-C. (E) Saddle plots displaying the averaged contact frequencies between bins stratified by compartment scores. The summarized statistics (left corner) were calculated by first multiplying the strengths of A-A and B-B interactions of the top 20% A or B bins (intra-compartment, diagonal) and then dividing to the square of the strength of A-B interactions from these bins (inter-compartment, off-diagonal). (F) Aggregated contact maps at the 200KB neighborhoods centering on the TAD boundaries detected from Micro-C.


A
Occurence in MM10
1.5×10^7
1.0×10^7
5.0×10^6
0
MseI
DdeI
MboI
TTNAA
AANTT
AATT
TTAA
ATAT
CTNAG
CANTG
ATNAT
TATA
AGNCT
CATG
AGCT
TGNCA
TGCA
TANTA
TCNGA
CCNGG
CTAG
GANTC
ACNGT
GTNAC
GGCC
GATC
GGNCC
GTAC
GCNGC
ACGT
CCGG
TCGA
GCGC
CGCG
CGNCG
Palindromic Motifs
B
MID Hi-C
in situ Hi-C
10000
5000
3500
2500
1500
1000
600
400
250
100
15
C
Micro-C
in situ Hi-C
0.95
Micro-C
MID Hi-C
0.95
MID Hi-C
in situ Hi-C
0.97
1.00
0.00
-1.00
D
(-1, 1)
MID Hi-C
(-1, 1)
MicroC
chr16: 3.2-57.8 Mb
Res: 100 Kb; Range: 0-57
E
PC1
PC1
20.48
25.77
log2 Ob/Exp
-1.5 0.0 -1.5
F
TAD Boundary
-100 kb
100 kb
log2 Ob/Exp
-0.5 0.0 -0.5

2.3 Robust Detection of CTCF Loops using MID Hi-C

To determine how MID Hi-C compares to Micro-C and in situ Hi-C in terms of detecting CTCF-mediated loops, we matched the numbers of MID Hi-C, Micro-C and in situ Hi-C interaction reads by down-sampling the Micro-C and in situ Hi-C datasets. Initially, we focused on analyzing paired CTCF-bound sites appeared as sharp dots on contact maps. In line with the previous reports, signals associated with paired CTCF binding sites as revealed by in situ Hi-C was diffusive when compared to Micro-C (Figure 2.1C; Hsieh et al., 2020). On the other hand, MID Hi-C showed pronounced focal enrichment of paired CTCF binding sites (Figure 2.1C). To systematically assess the signal quality at paired CTCF binding sites, we used HiCCUPS (Rao et al., 2014) to call significant interactions from three datasets at 5 and 10 KB resolutions. We detected 24125, 16213, and 6342 loops by analyzing the equivalent number of interaction reads derived from MID Hi-C, Micro-C and in situ Hi-C, respectively. For all three protocols, >90% of the loops showed paired CTCF occupancy (Figure 2.3A). A large fraction of CTCF loops detected by the three different protocols overlapped (Figure 2.1D). Specifically, nearly all (91%) CTCF loops detected by in situ Hi-C were identified as well using MID Hi-C versus 75% identified using Micro-C. Notably, 8962 (66%) CTCF loops revealed by Micro-C were identified using MID Hi-C (49% CTCF-mediated looping as detected using MID Hi-C). MID Hi-C also allowed detection of 8614 additional CTCF loops, almost twice the number as detected uniquely using Micro-C (Figure 2.1D). To quantify and compare enrichment signal strength for these loop subsets for in situ Hi-C, Micro-C and MID Hi-C, we performed Aggregate Peak Analyses (APA, a.k.a pileup analysis; Rao et al., 2014). We found that, unlike in situ Hi-C, enrichment signal of MID Hi-C was comparable to that of Micro-C (Figure 2.1E and 2.3B). We detected a 1.9- and 1.75-times signal boost as revealed using MID Hi-C and Micro-C datasets respectively versus in situ Hi-C at the consensus CTCF loop subsets (Figure 2.1E). For Micro-C specific CTCF loops that were ineffectively captured by in situ Hi-C, a modest enrichment was seen from MID Hi-C (Figure 2.1E). Finally, MID Hi-C specific CTCF loops were also associated with strong signal strength as detected using Micro-C further validating the notion that these are bona-fide CTCF-mediated loops. Collectively, these observations indicate that MID Hi-C enables accurate and robust detection of CTCF-mediated looping.

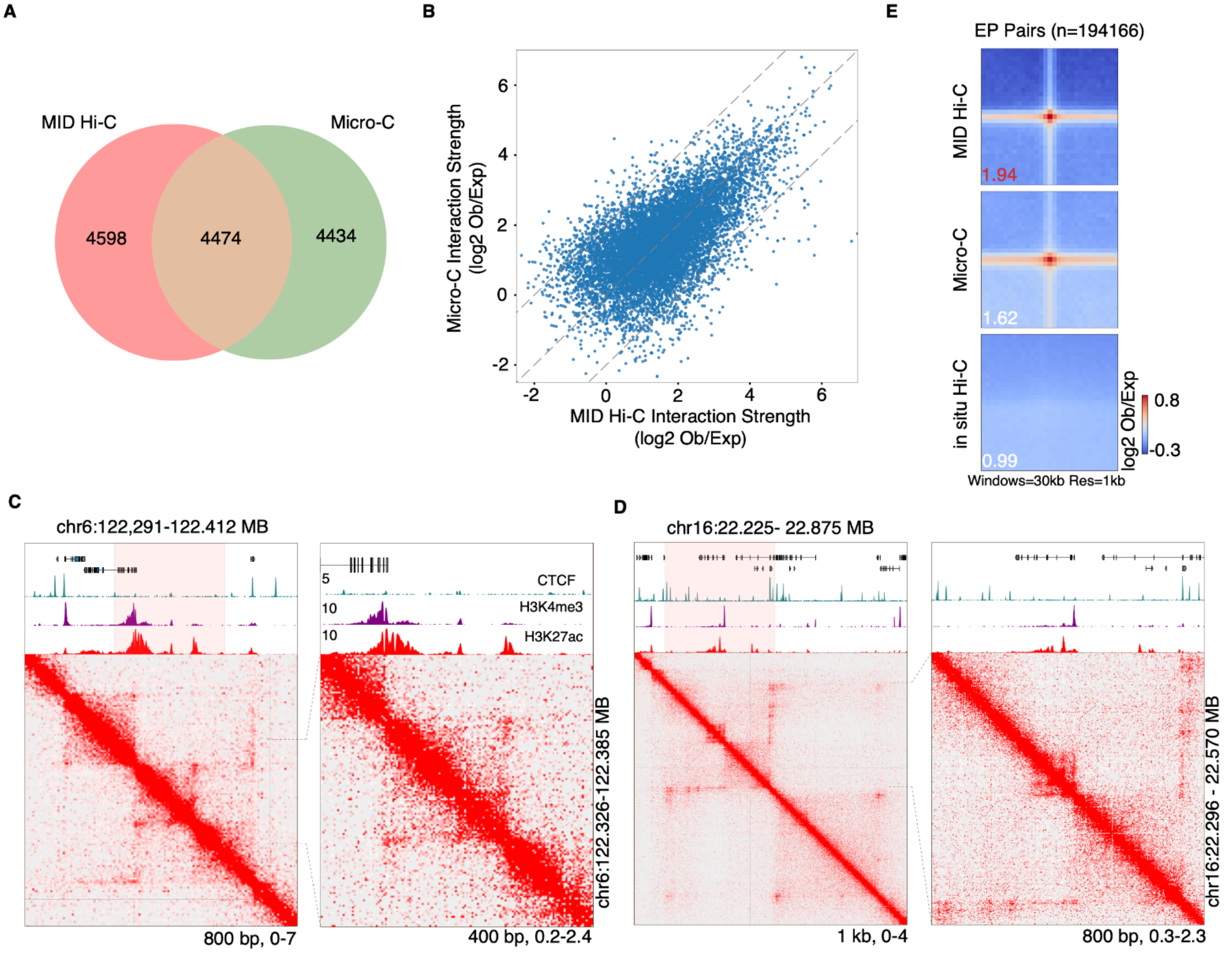
Figure 2.3. MID Hi-C captures strong signals at CTCF loops and E-P interactions. (A) Analysis of CTCF occupancies at the anchors of the HiCCUPS loops. (B) APA at the intersecting CTCF loop subsets. The three-digit code of each column is the indicator denoting whether the intersecting set contains loops from the datasets in the order of the rows (arrow on the left). For example, '101' stands for loop subset that is shared by MID Hi-C and in situ Hi-C but not by Micro-C. The calculation of the enrichment scores is the same as Figure 2.1E. (C) APA at the intersecting E-P loop subsets that are called by HiCCUPS. The coding of the indicator and calculation of the enrichment scores are the same to (B). (D) Scatter plot contrasting the interaction strengths of all possible 'CTCF-free' E-P pairs (separated by 5 to 500 KB) as measured by MID Hi-C (x axis) and Micro-C (y axis).


A
CTCF Occupancies at loop anchors
25k
20k
15k
10k
5k
0
Both
One
Neither
MID Hi-C
Micro-C
in situ Hi-C
HiCCUPS Loops
B
CTCF Loops (HiCCUPS)
001
011
101
Subset Indicator
Subset Indicator Order
MID Hi-C
Micro-C
in situ Hi-C
3.63
4.34
6.39
2.66
4.85
4.55
2.87
2.53
3.48
log2 Ob/Exp
0-2
0-2.5
0-3
C
EP Loops (HiCCUPS)
10
11
01
Subset Indicator
Subset Indicator Order
MID Hi-C
Micro-C
in situ Hi-C
3.72
4.33
2.33
2.29
3.82
2.82
1.73
2.13
1.49
log2 Ob/Exp
0-2
D
EP Pairs in 5 to 500 KB
Micro-C Interation Strength (log2 Ob/Exp)
MID Hi-C Interation Strength (log2 Ob/Exp)
-2
0
2
4
6

2.4 Robust Detection of Enhancer-Promoter Interactions using MID Hi-C

To evaluate the potential of MID Hi-C in detecting enhancer-promoter (E-P) interactions, we identified the E-P interactions from HiCCUPS loop sets by annotating anchors using ChromHMM (Ernst and Kellis, 2012, 2017). This analysis found 9032, 8908, and 1660 loops involving E-P interactions using MID Hi-C, Micro-C, and in situ Hi-C protocols, respectively. We focused our analysis on E-P loops detected by MID Hi-C and Micro-C since the majority of E-P interactions revealed by in situ Hi-C were also identified using MID Hi-C and Micro-C. We found that nearly 50% of E-P interactions overlapped for both loop-sets (Figure 2.4A). APA for the E-P loop subsets showed that looping signals of both MID Hi-C and Micro-C were highly enriched (Figure 2.3C). Likewise, we found that using MID Hi-C interaction strengths of combined E-P loops correlated well with that observed using Micro-C (Figure 2.4B). We next focused our analysis on E-P interactions that were not associated with CTCF occupancy. We note that in situ Hi-C merely detects E-P interactions devoid of CTCF occupancy (Hsieh et al., 2020; Krietenstein et al., 2020). In contrast, we found that MID- Hi-C readily detects E-P interactions not linked with CTCF binding, at high-resolution (<800 bp) (Figure 2.4C and 2.4D). To systematically confirm these observations, we aggregated contact maps for all possible pairs involving ‘CTCF-free’ enhancers and promoters that are separated by 5 to 500 kb. We found significant enrichment for both MID Hi-C and Micro-C but not in situ Hi-C (Figure 2.4E). Additionally, we found that the interaction strengths using MID Hi-C or Micro-C were quantitatively correlated (Figure 2.3D). Taken together, these data indicate that MID Hi-C readily allows detection of E-P interactions that are not associated with CTCF occupancy akin to that observed for Micro-C but not in situ Hi-C.

Figure 2.4. MID Hi-C can detects E-P interactions. (A) Venn diagram displaying the intersection of the E-P loops parsed from the HiCCUPS loops detected from MID Hi-C and Micro-C datasets. (B) Scatter plot contrasting the interaction strengths of the combined E-P HiCCUPS loops as measured by MID Hi-C (x axis) and Micro-C (y axis). (C and D) Two examples of MID Hi-C contact maps visualizing E-P interactions at high resolutions. (E) APA at all possible ‘CTCF-free’ E-P pairs that are separated by 5 to 500 KB.


A
MID Hi-C
Micro-C
4598
4474
4434
B
Micro-C Interaction Strength (log2 Ob/Exp)
MID Hi-C Interaction Strength (log2 Ob/Exp)
-2
0
2
4
6
E
EP Pairs (n=194166)
MID Hi-C
1.94
Micro-C
1.62
in situ Hi-C
0.99
log2 Ob/Exp
0.8
-0.3
Windows=30kb Res=1kb
C
chr6:122,291-122.412 MB
5
CTCF
10
H3K4me3
10
H3K27ac
chr6:122.326-122.385 MB
800 bp, 0-7
400 bp, 0.2-2.4
D
chr16:22.225- 22.875 MB
chr16:22.296 - 22.570 MB
1 kb, 0-4
800 bp, 0.3-2.3

2.5 Optimal HiChIP Derived from MID Hi-C

In principle the relatively mild denaturing steps used in MID Hi-C, allows for greater enrichment of target-bound chromatin contacts using chromatin immunoprecipitation sequencing (ChIP-seq). To explore this possibility, we included MID Hi-C in the HiChIP protocol (Mumbach et al., 2016) and named the derived method as MID HiChIP. We generated cohesin and H3K27ac MID HiChIP datasets in mESCs to compare them with the published HiChIP datasets (Mumbach et al., 2016). We began by comparing the efficiency of immunoprecipitation using MID Hi-C versus HiC conditions. Specifically, we overlaid datasets of ChIP-seq, MID HiChIP, and HiChIP at the cohesin (generated in lab) and H3K27ac (ENCODE, GSM1000126) peaks called from ChIP-seq profiles using MACS2 (Zhang et al., 2008). We found significant enrichment of MID HiChIP as compared to HiChIP (Figure 2.5A). Notably, the ChIP efficiency of MID HiChIP was comparable to that of using cohesin or H3K27Ac ChIP-seq alone (Figure 2.5A and 2.6A).

To determine whether MID HiChIP can enrich target-bound contacts and filter unbound contacts, we compared MID HiChIP versus HiChIP contact maps. We detected remarkably increased signal-to-noise ratios using MID HiChIP versus HiChIP, when enriched for either cohesin occupancy or deposition of H3K27Ac (Figure 2.5B). Enrichment was particularly striking using cohesin MID HiChIP revealing sharp interaction peaks of paired CTCF binding sites (Figure 2.5C and 2.6C). Specifically, we found that enrichments of interaction signals at paired CTCF binding sites were on average increased for a factor of 3.7-fold using cohesin MID HiCHIP when compared to cohesin HiChIP (Figure 2.5D and 2.6B). Likewise, signals associated with H3K27ac MID HiChIP at H3K27ac-anchored interactions were on average 2.3-fold stronger relative to H3K27ac HiChIP (Figure 2.5D and 2.6B). As expected, MID HiChIP signals associated with cohesin occupancy differed from that observed for H3K27ac MID HiChIP consisting with cohesin sites and H3K27Ac-marked enhancers being associated with distinct chromatin tethers (Figure 2.5B, 2.5D, and 2.6B). In contrast, the distinction between cohesin and H3K27ac HiChIP in contact maps was less pronounced (Figure 2.5B, 2.5D, and 2.6B). We found that fewer number of interaction reads were required by MID HiChIP (Figure 2.5B). As a proof of principal, we downsampled the cohesin MID HiChIP dataset to 20 million interaction reads. Even when used at low depth, we were able to visualize paired CTCF interactions in relatively high signal-to-noise ratios (Figure 2.6D), suggesting superior cost efficiency of MID HiChIP.

Finally, we compared the MID HiChIP and HiChIP protocols to identify significant paired cohesin-dependent loops using cohesin ChIP-Seq peaks as anchor input for loop calls by hichipper (Lareau and Aryee, 2018). We found a large percentage of significant paired-cohesin loops detected by MID HiChIP and HiChIP were shared (Figure 2.7B). To evaluate how well these two datasets in detecting cohesin-mediated loops, we next compared the recall rates for the highly confident paired-cohesin loops, combined loops of MID Hi-C and Micro-C called by HICCUPS (described in chapter 2.3), as a function of an increasing number of top ranked loops (decreasing stringency) that were rated using cohesin-dependent MID HiChIP or HiChIP. As expected, we found a substantially higher percentage of reference loops were re-identified using cohesin-dependent MID HiChIP as compared to HiChIP (Figure 2.5E). Taken together, these data demonstrate that MID Hi-C combined with chromatin immunoprecipitation enables significant enrichment for paired binding sites.

Figure 2.5. Enrichment of target-bound DNA contacts by MID HiChIP. (A) Tornado plot showing the enrichment of ChIP, MID HiChIP, and HiChIP at the cohesin or H3K27ac peaks by ChIP-seq. (B) Comparison of contact matrices of MID HiChIP and HiChIP. (C) Example of high-resolution contact maps displaying superior signal-to-noise ratio of cohesin MID HiChIP. (D) Violin plot showing the distributions of the interaction strengths of CTCF-anchored interactions (blue violin) or H3K27ac-anchored interactions (orange violin), measured by the indicated HiChIP datasets (y axis labels). The numbers on the arrows represent the fold differences of the means between the source violin versus the target violin. (E) Trend lines describing the recall rate (y axis) for the reference paired-cohesin loops, combined from HiCCUPS loop calls from MID Hi-C and Micro-C, as a function of the numbers (x axis) of top rank loops as rated by hichipper or OBJECT using ChIP-seq peaks as anchor inputs.

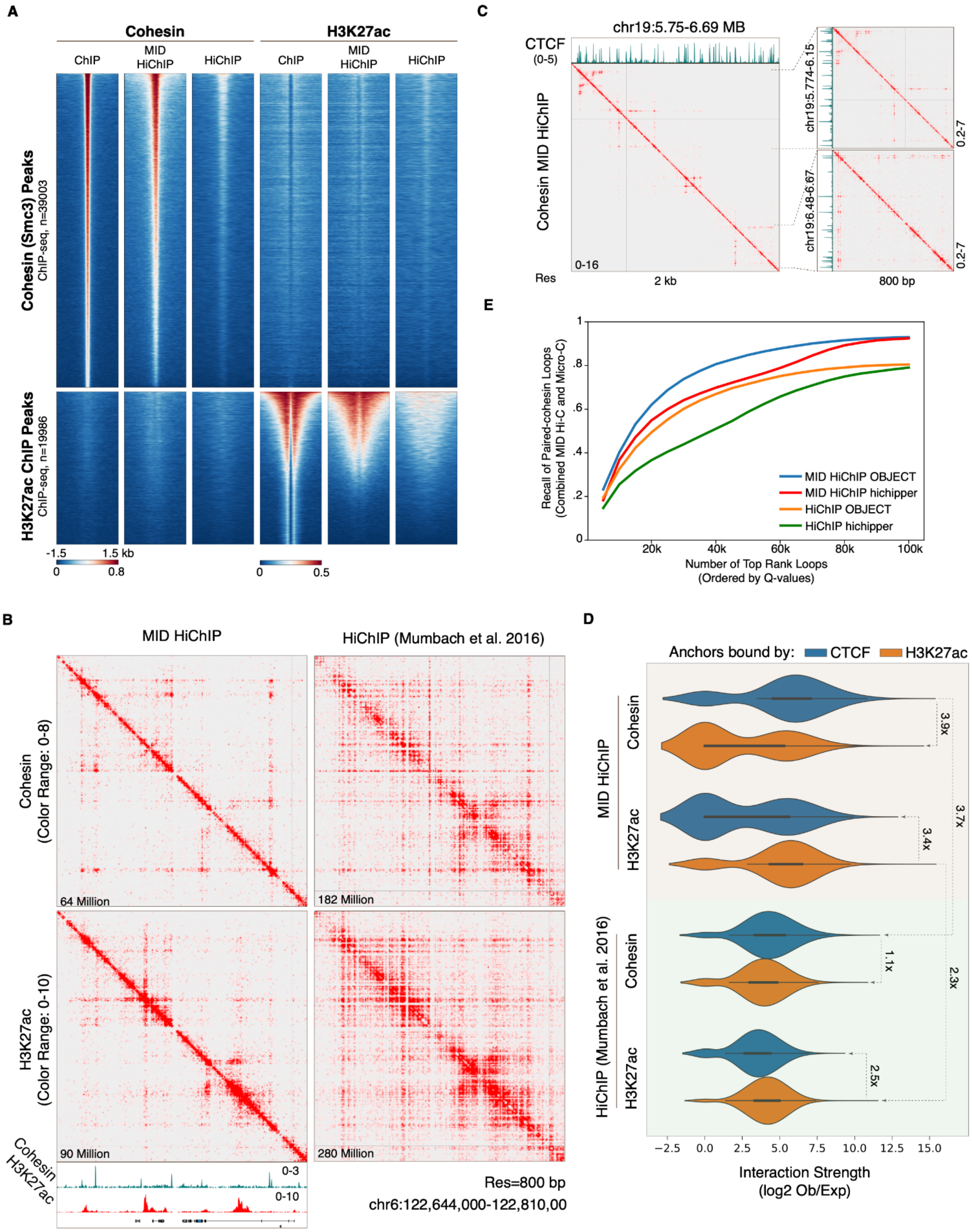
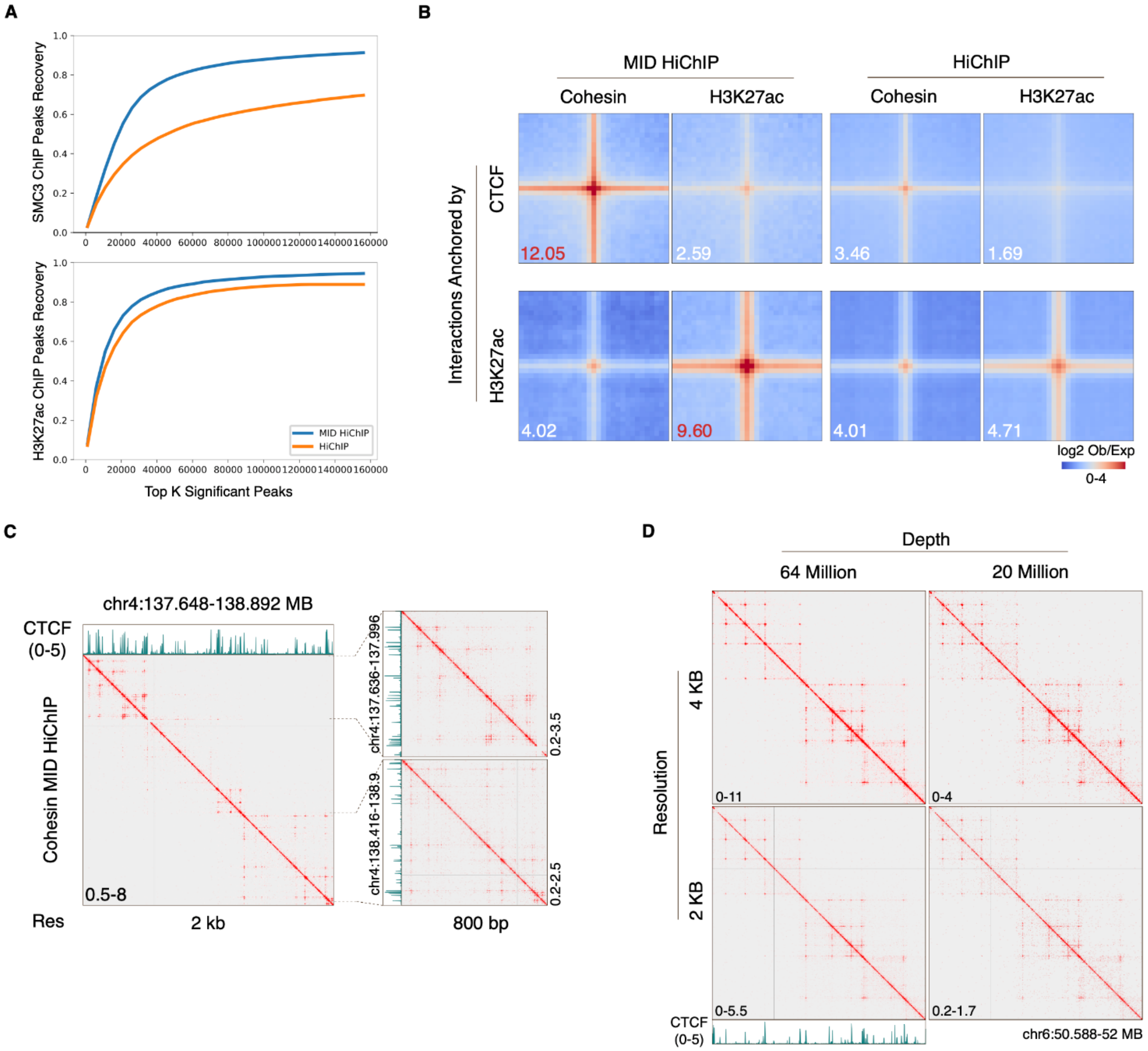

A
Cohesin
H3K27ac
ChIP
MID HiChIP
HiChIP
Cohesin (Smc3) Peaks
ChIP-seq, n=39003
H3K27ac ChIP Peaks
ChIP-seq, n=19986
-1.5
1.5 kb
0
0.8
0
0.5
C
chr19:5.75-6.69 MB
CTCF
(0-5)
Cohesin MID HiChIP
0-16
Res
2 kb
chr19:5.774-6.15
chr19:6.48-6.67
0.2-7
0.2-7
800 bp
E
Recall of Paired-cohesin Loops
(Combined MID Hi-C and Micro-C)
1
.8
.6
.4
.2
0
20k
40k
60k
80k
100k
Number of Top Rank Loops
(Ordered by Q-values)
MID HiChIP OBJECT
MID HiChIP hichipper
HiChIP OBJECT
HiChIP hichipper
B
MID HiChIP
HiChIP (Mumbach et al. 2016)
Cohesin
(Color Range: 0-8)
H3K27ac
(Color Range: 0-10)
64 Million
182 Million
90 Million
280 Million
Cohesin
H3K27ac
0-3
0-10
Res=800 bp
chr6:122,644,000-122,810,00
D
Anchors bound by:
CTCF
H3K27ac
MID HiChIP
HiChIP (Mumbach et al. 2016)
Cohesin
H3K27ac
Cohesin
H3K27ac
3.9x
3.7x
3.4x
1.1x
2.3x
2.5x
−2.5
0.0
2.5
5.0
7.5
10.0
12.5
15.0
Interaction Strength
(log2 Ob/Exp)

Figure 2.6. Superior signal of MID HiChIP. (A) Trend lines showing the recall rate for cohesin or H3K27ac ChIP-seq peaks by the number of top rank peaks detected from the MID HiChIP or HiChIP datasets. (B) APA displaying the enrichment signal of the MID HiChIP or HiChIP datasets at the CTCF-anchored interactions (first row) or H3K27ac-anchored interactions (second row). The enrichment scores are calculated as Figure 2.1E. (C) Example of high-resolution contact maps displaying superior signal-to-noise ratio of cohesin MID HiChIP. (D) Contact maps showing the signal qualities of the downsampled (20 million) cohesin MID HiChIP dataset.


A
SMC3 ChIP Peaks Recovery
1.0
0.8
0.6
0.4
0.2
0.0
0
20000
40000
60000
80000
100000
120000
140000
160000
H3K27ac ChIP Peaks Recovery
MID HiChIP
HiChIP
Top K Significant Peaks
B
MID HiChIP
HiChIP
Cohesin
H3K27ac
Cohesin
H3K27ac
Interactions Anchored by
CTCF
H3K27ac
12.05
2.59
3.46
1.69
4.02
9.60
4.01
4.71
log2 Ob/Exp
0-4
C
chr4:137.648-138.892 MB
CTCF
(0-5)
Cohesin MID HiChIP
0.5-8
Res
2 kb
chr4:137.636-137.996
0.2-3.5
chr4:138.416-138.9
0.2-2.5
800 bp
D
Depth
64 Million
20 Million
Resolution
4 KB
2 KB
0-11
0-4
0-5.5
0.2-1.7
CTCF
(0-5)
chr6:50.588-52 MB

2.6 New Loop Caller for HiChIP

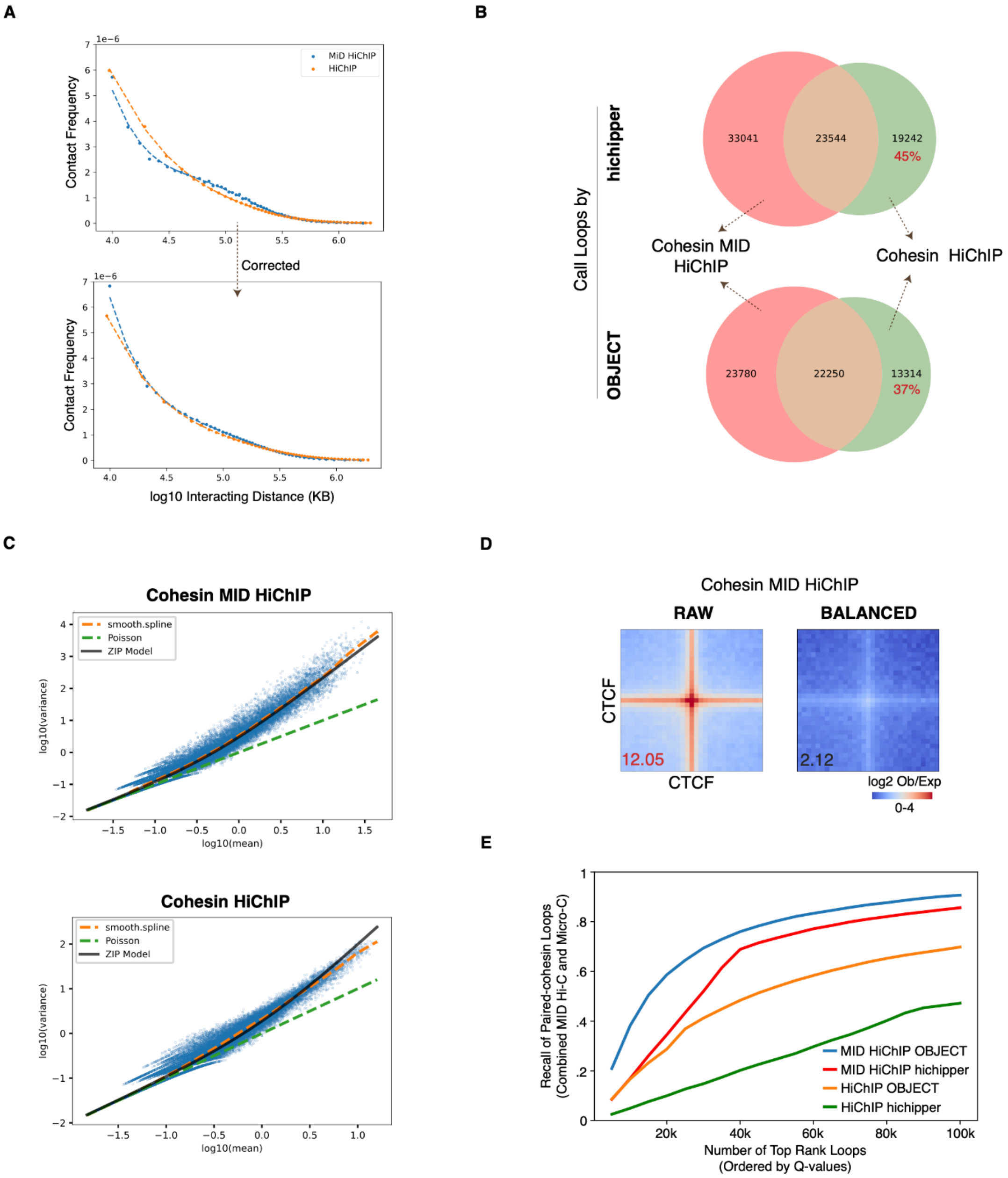
Upon further inspection of the MID HiChIP data, we observed two irregularities: a slower decay rate in the scaling profile of MID HiChIP and an overdispersed relationship between mean and variance (Figure 2.7A and 2.7C). The latter one is presented in HiChIP as well. Notice that these two irregularities affect the estimations of the background and count distribution, two most critical modeling components in HiChIP loop callers.

The slower decay rate of MID HiChIP is most prominent in the interacting range separated by 30 KB to 300 KB, where elevated looping signals are observed from MID HiChIP versus HiChIP. It needs to be corrected for a proper estimation of the background. Therefore, we employed an iterative loop calling approach that rebuilt the background model by removing significant interactions with which the loop call of last round was highly confident. We implemented the iterative version of the MANGO algorithm (Phanstiel et al., 2015) which is used in hichipper (Lareau and Aryee, 2018). We found the background model converged for all the tested HiChIP datasets using iterative MANGO. Specifically, the corrected scaling curve had a typical decaying pattern resembling that of HiChIP (Figure 2.7A**)**.

The overdispersion pattern we observed in HiChIP data (Figure 2.7C) implied that the underlying count distribution cannot be recapitulated by poisson distribution or binomial distribution that is assumed by existing HiChIP loop callers. Failing to address the inflated variance will increase the false positive rate of the model. Two alternative distributions are used to account for overdispersion in count data; that are zero-inflated poisson (ZIP) distribution and negative binomial (NB) distribution. To evaluate these distributions, we exploited the framework of general linear model (GLM) to simultaneously construct the background model and estimate the parameter of the count distribution. The model has two explanatory variables, interacting distance (L) and anchor coverages (D), that determined the expected count of the response variable. Then, the count values were modeled by ZIP or NB distributions. We found that the mean-variance relationship as modeled by ZIP distribution closely resembled the overdispersion trend in the data (Figure 2.7C). Motivated by these observations, we implemented the iterative procedures, as described above, for the ZIP-based GLM approach. We termed this new loop calling method as overdispersion and (Hi-C and ChIP) bias jointly emulated and corrected tool (OBJECT).

To benchmark OBJECT, we used it to call significant paired-cohesin loops from the cohesin-dependent MID HiChIP and HiChIP datasets. The consensus ratio, as revealed by intersection analysis of two loop sets, was increased relative to the ones that were called by hichipper (Figure 2.7B). Furthermore, the rates of recalling highly confident loops of paired cohesin binding sites were also higher when using OBJECT versus hichipper (Figure 2.5E). It is worth noting that the improvements of recall rates were more prominent for noisier inputs, such as calling loops from HiChIP data versus MID HiChIP data, or using HiChIP peak calls as anchor input versus ChIP-seq. This analysis highlights the notion that OBJECT is more robust against noise than hichipper, justifying our choice of ZIP distribution.

Figure 2.7. Call loops from HiChIP data using OBJECT. (A) Scaling profiles showing the elevated contact coverages at the mid-range before (top) and after (bottom) correction. (B) Venn diagram displaying the intersection of significant ($q < 0.01$) paired-cohesin loops detected from cohesin MID HiChIP and HiChIP by hichipper (top) or OBJECT (bottom). (C) Scatter plots displaying overdispersion patterns in HiChIP data. (D) APA at paired-CTCF interactions aggregated from raw or balanced contact matrices of cohesin MID HiChIP data. (E) Trend lines describing the recall rate (y axis) for the reference paired-cohesin loops, combined from HiCCUPS loop calls of MID Hi-C and Micro-C, as a function of the numbers (x axis) of top rank loops as arranged by hichipper or OBJECT using peaks directly called from HiChIP data as anchor inputs.


A
B
C
D
E
1e−6
Contact Frequency
MiD HiChIP
HiChIP
Corrected
log10 Interacting Distance (KB)
Call Loops by
hichipper
OBJECT
33041
23544
19242
45%
23780
22250
13314
37%
Cohesin MID HiChIP
Cohesin HiChIP
Cohesin MID HiChIP
Cohesin HiChIP
smooth.spline
Poisson
ZIP Model
log10(variance)
log10(mean)
Cohesin MID HiChIP
RAW
BALANCED
CTCF
CTCF
12.05
2.12
log2 Ob/Exp
0-4
Recall of Paired-cohesin Loops
(Combined MID Hi-C and Micro-C)
MID HiChIP OBJECT
MID HiChIP hichipper
HiChIP OBJECT
HiChIP hichipper
Number of Top Rank Loops
(Ordered by Q-values)
20k
40k
60k
80k
100k

2.7 Summary

In this chapter, we developed a new Hi-C approach, termed MID Hi-C, characteristic with finer but manageable chromatin fragmentation as well as better preservation of chromatin interactions via using second crosslinker and milder condition as opposed to in situ Hi-C. We showcased that MID Hi-C surpassed in situ Hi-C and is comparable to Micro-C in detecting chromatin interactions at high resolutions (< 1 KB). We have also adapted the pre-processing pipeline from HiC-Pro to cope with the modified workflow of MID Hi-C (Servant et al., 2015).

The parallel workflow of MID Hi-C and in situ Hi-C makes easy upgrades of the existing Hi-C derived assays such as Capture-C and HiChIP. As a proof of principle, we developed MID HiChIP. We found that the mild condition used in MID Hi-C substantially increased the subsequent ChIP efficiency in MID HiChIP. Combined with the richer Hi-C library, MID HiChIP generates fine mapping of chromatin interactions with high signal-to-noise ratio and cost efficiency. Additionally, we developed a new loop caller for HiChIP to effectively correct the bias from Hi-C and ChIP. Together, we built a toolkit for the next generation 3C that addresses the pressing challenge of the field to chart fine-scale chromatin interactions.

2.8 Material and Methods

mESC Cell Culture

E14TG2a mESCs were maintained in KnockOut DMEM (Thermo Fisher,10829018) medium, supplied with 15% FBS (R&D Systems, S10250H), 1mM MEM Non-Essential Amino Acids (Thermo Fisher, 11140-050), 2mM Glutamax (Thermo Fisher, 35050061), 100 U/mL Pen/Strep (Thermo Fisher, 15140122), 0.1 mM 2(β)-ME (Sigma, M-3148), and 1000 U/mL LIF (Cell Guidance Systems, GFM200) on 0.1% gelatin coated dishes at 37C and 5% CO2. Cells were passaged every two days. Medium was changed on a daily basis.

MID Hi-C Experiment

mESC cells were double crosslinked by formaldehyde and DSG then aliquoted as described previously (Hsieh et al., 2020). One aliquot of fixed cells (~5 million) is washed in 1.4 mL low SDS buffer (50 mM pH 7.5 Tris-HCL, 10 mM NaCl, 0.1% SDS) with 1X Protease Inhibitor Cocktails (PICs). Incubate on ice for 15 min. Spin down the nuclei at 2500 rcf for 5 min at 4C. Remove supernatant and wash one more time in the low SDS buffer. Then, resuspend nuclei in 1 mL digestion buffer (1X Cutsmart buffer, 1% Triton X-100, 1X PICs). Incubate in 37C for 20 min with gentle rotation. Add 250U MseI and 250U DdeI to digest nuclei at 37C for 6 hours or overnight with gentle rotation.

After digestion, spin down the nuclei and resuspend in 1 mL fill-in reaction mix (1X CutSmart buffer, 1% Triton X-100, 66 uM dTTP, 66 uM dCTP, 66 uM dGTP, 66 uM dATP-Biotin, 50 U Klenow DNA large polymerase I, 1X PICs). Rotate at room temp for 90 min. Spin down nuclei and resuspend in 1 mL ligation mix (1X T4 DNA ligase buffer, 1% Triton X-100, 0.1 mg/mL BSA, 15 U T4 ligase, 1X PICs). Incubate at 14C overnight with gentle rotation. Next day, spin down the nuclei. The proximal ligated nuclei can be used for library preparation or proceeded for MID HiChIP.

MID Hi-C Library Preparation

Resuspend proximal ligated nuclei in 300 uL ChIP elution buffer (1% SDS, 100 mM NaHCO3) and incubate at 65C overnight. Next day, add RnaseA and incubate at 37C for 45 min. Then, add Proteinase K and incubate at 55C for 2 hours. Purify DNA using Phenol:Chloroform:Isoamyl (PCI) method. Purified DNA can be stored in -80C till ready for library preparation.

Prepare MyOne Streptavidin T1 beads by washing 5 uL T1 beads in 600 ul Tween wash buffer (1X binding buffer, 0.05% Tween-20) at 55 C for 2 min for twice. Reclaim the beads on magnetic stand and resuspend in 50 ul 2X binding buffer (2 M NaCl, 10 mM pH 7.5 Tris-HCl, 1 mM EDTA). The beads are ready for immobilizing biotin labeled DNA.

We used NEB fragmentase to fragment DNA for library preparation. But sonication can be used to fulfill the same purpose. We treat 50 ng to 500 ng of DNA with NEB fragmentase in 20 uL reaction volume, per manufacture instructed. Incubate the reaction at 37C for 9 to 13 min, optimized to generate 100 – 500bp fragment size. To stop the reaction, add 250 ul 2X binding buffer to the DNA fragmentation reaction and incubate at 55C for 15 min. Then, add 280 uL Zymo DNA elution buffer and 50 ul prepared T1 beads to the stopped reaction (total 600 ul). Rotate the beads and DNA at room temperature for 15 min. After that, reclaim the beads and wash twice in the Tween wash buffer as above.

The DNA-bound T1 beads were used for NEBNext library preparation per manufacture instruction with gentle taps during reactions to keep beads in suspension. We typically amplify

8-10 PCR cycles for library amplification. The library is then double size selected using home-made SPRI beads to isolate fragments between 200 and 800 bp for next generation sequencing.

MID HiChIP

Proximal-ligated nuclei are processed for ChIP following Myers Lab ChIP-seq Protocol (v011014). In brief, proximal-ligated nuclei are resuspended in RIPA buffer followed by sonication using Bioruptor. Soluble chromatin fragments are immunoprecipitated by SMC3 (abcam ab9263) or H3K27ac (ActiveMotif 39133) antibody overnight. We used 3 to 5 mg of antibody for the starting material of 5 million cells. Next day, wash the immunoprecipitated chromatin extensively. Purify the DNA for library preparation after reverse crosslinking.

All ChIPed DNA is immobilized on T1 beads and the library is prepared on beads as MID Hi-C library preparation described above. We typically got 10 to 50 ng ChIPed DNA for library generation. We have successfully generated libraries for next generation sequencing started by as low as 5 ng DNA with less than 15 PCR amplification cycles.

DATA Analysis

MID Hi-C Sequencing Data Process

Adaptors in pair-end (PE) sequencing reads were trimmed using Trimmomatic. Trimmed PE reads were aligned to mm10 genome using HiC-Pro analysis pipeline (Servant et al., 2015). To improve mappability, HiC-Pro performs 'global' and 'local' mapping. In the global mapping stage, full-length reads are first aligned. Any unmapped reads will be split at the ligation sites (MID Hi-C ligation sites: TTATAA, TTATNAG, CTNATNAG, CTNATAA) and the 5' segment will be re-aligned in the local mapping stage. We denote this one pass of iterative alignment.

Instead of dumping the 3’ segment, more passes of iterative alignment can be done till no more segments of reads are left to align. The multi-pass iterative mapping (MPMAP) was designed for the MID Hi-C dataset for two purposes. First, MPMAP salvages the unmapped reads in the local mapping stage. Since chromatin fragments generated by MID Hi-C are smaller than in situ Hi-C, the 5’ segments, which are truncated during library preparation, are sometimes too small to be aligned. Therefore, we need to resort to inner fragments. MPMAP can increase 10 to 20 percent library yield for MID Hi-C. Second, given a fixed read length like 100 bp, it can sequence a greater number of fragments for MID Hi-C versus in situ Hi-C. MPMAP retrieves the multi-fragment information which can be leveraged to study multiway interactions (Quinodoz et al., 2018; Rao et al., 2017) in future studies. To validate the MPMAP pipeline, we compared the single pass mode with the HiC-Pro pipeline. They produced essentially the same output.

For multi-fragment interaction reads, we kept two longest separated fragments for Hi-C filtering. The output of validpairs was transformed to HIC (Rao et al., 2014) and COOL (Abdennur and Mirny, 2019) formats for downstream analysis. In brief, validpair data was first binned and normalized using JUICER toolkit (Durand et al., 2016a) and stored in the HIC file. The contact maps shown in the figures were directly visualized by JUICEBOX (Durand et al., 2016b). To call HiCCUPS loops, we used the HiCCUPS subcommand of JUICER toolkit with flags `-r 5000,10000 -f 0.1,0.1 -p 4,2 -I 7,5 -d 20000,20000`. For all other downstream analysis, we then transformed the HIC file to COOL file by dumping the raw and normalized matrices in HIC file to COOL storage. This allowed us to use COOLER and COOLTOOLS packages (Nora et al., 2020) to access Hi-C data and seamlessly integrate downstream analysis with Python.

Published in situ Hi-C and Micro-C Data Process

We processed the published in situ Hi-C data of mESCs (Bonev et al., 2017) using HiC-Pro analysis pipeline, setting fragment bed file and ligation site specific to Hi-C. For processing the published Micro-C data (Hsieh et al., 2020), we left these two parameters unset. Output validpairs were processed to HIC and COOL files as described above. Unless specified, downstream analyses were performed with the same parameters for all three datasets (MID Hi-C, in situ Hi-C, and Micro-C).

Contact Scaling Curves

We followed the example (https://cooltools.readthedocs.io/en/latest/notebooks/contacts_vs_distance.html) from COOLTOOLS to generate the contact scaling curves in Figure 1B.

A/B Compartment Analysis and Saddle Plot

A/B compartments were called by CscoreTool (Zheng and Zheng, 2017) with default parameters. The compartment scores generated by CscoreTool is on a uniformed scale between -1 to 1, allowing for direct comparison between datasets. Compartmental tracks were plotted by pyGenomeTracks. Scikit-learn (https://scikit-learn.org/stable/) was used for PCA analysis for the compartment score profiles of samples. Saddle plot and the associated metrics were generated as described previously (Gibcus et al., 2018; Zhu et al., 2020).

Topologically Associated Domain (TAD) Analysis

TADs were called by TopDom in 10 KB resolution and 20 window sizes. We took the boundary locations detected from the Micro-C dataset for pileup analysis (ADA analysis) on the MID Hi-C dataset.

Loop Annotation

Loop annotation is essentially annotating the anchors. For example, we checked whether the anchors were bound by CTCF, given a list of CTCF peaks. In cases of annotating E-P interactions, we either used the annotations from https://github.com/guifengwei/ChromHMM_mESC_mm10 by ChromHMM or ‘CTCF-devoid’ enhancers and promoters identified from ATAC-seq analysis as described in chapter 3.

Loop Intersection Analysis

We intersected loops called from different datasets using an anchor-centric approach. Briefly, anchors of loops were adjusted to 10 KB sizes. Then, anchors of loops to be merged were pooled and clustered based on overlapping by PyRanges, tolerating 7500 BP distance (`slack=7500`). Therefore, each loop can be represented by a pair of IDs of the anchor cluster. Loops with the same pair of IDs were considered intersecting. In essence, this approach merged loops within 25 KB Euclidean distances (Rao et al., 2014).

Aggregated Peak Analysis (APA, a.k.a. pileup analysis)

We performed pileup analysis similar to APA or ADA analysis as described previously (Rao et al., 2014).

MID HiChIP Analysis

Sequencing data of MID HiChIP was trimmed, mapped, filtered, and created HIC/COOL files in the same way as the MID Hi-C data process above. For the published HiChIP data, we processed in the same way as the in situ Hi-C data process above.

To assess the ChIP enrichments, we pooled all fragments below 1 MB separations to call peaks and generate tornado plots as described in ATAC-seq analysis in chapter 3. Visualization of the contact maps and APA analyses were similar as described above, except for using raw count values instead of normalized values. We found that MID HiChIP substantially violated the equal-visibility assumption for Hi-C matrix normalization (Imakaev et al., 2012) due to the high enrichment at target-bound regions. Applying normalization to contact matrices of MID HiChIP almost wiped out the interaction signals (Figure 2.7D). We used hichipper to call significant paired-cohesin loops with flags `--peak-pad 0 –skip-resfrag-pad –no-merge`.

Distance Scaling Curves for MID HiChIP and HiChIP

We generate the distance scaling curves the same way as described in MANGO (Phanstiel et al., 2015).

Mean-Variance Analysis for MID HiChIP and HiChIP

We first performed equal occupancy binning (Ay et al., 2014) in the distance (L) or depth (D) spaces. Depth for a pair of bins was calculated by the production of the coverages of the individual peak. Using these two series of break points, we formed a parametric grid in the 2D space (L and D). Each L-D bins were considered to have similar distance separations and depth coverages so that all paired peaks (regardless of being observed) fallen in a L-D bin were assumed to have the same count distribution. Mean and variance were approximated for every L-D bins. The resulting pairs of means and variances were plotted as scatters.

2.9 Acknowledgements

For their contributions in chapter 2, thanks to: Paula Pham, Han Han, Jeff Jiajing Zhou, Yina Zhu, Wenqi Wang, Ken Cho and Cornelis Murre. The material in this chapter is currently prepared for submission for publication.

CHAPTER 3

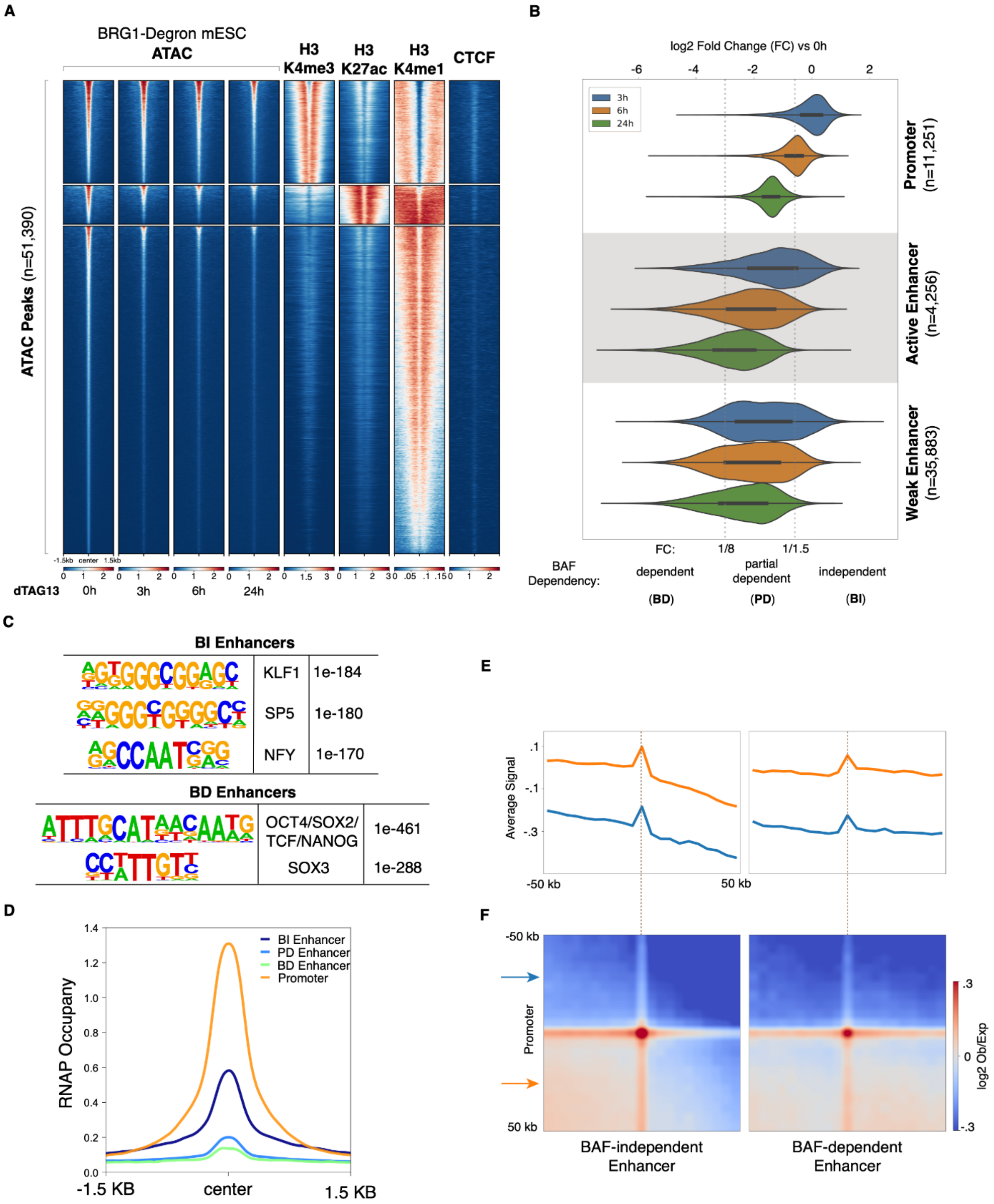
BAF Complex Instructs Chromatin Folding

In this chapter, we investigated the alterations in chromatin folding in response to the perturbation of BAF complex, the chromatin remodeler that controls chromatin accessibilities at cis-regulatory elements. We found enhancer syntax instructs the dependency of enhancer on the BAF complex as well as the contact configurations of enhancer-promoter (E-P) hubs. Loss of BAF occupancies weakened the strengths of a subset of E-P interactions. Lastly, we examined how global reduction in enhancer accessibilities affect the nuclear compartmentalization.

3.1 Enhancer Sequence Dictates Selective Recruitment of Chromatin Remodelers

The data described in chapter 2 indicates that MID Hi-C, unlike in situ HiC, permits the robust detection of E-P interactions and enables us to systematically determine how E-P interactions are regulated. Numerous previous studies have revealed that chromatin remodelers, including the BAF complex, promote chromatin accessibility across enhancers (Alver et al., 2017; Nakayama et al., 2017; Park et al., 2021; Wang et al., 2017). To determine how the BAF complex regulates enhancer accessibility we inserted a dTAG-inducible FKBPv degron (Nabet et al., 2018) into the SMARCA4 locus, encoding for the only ATPase BRG1, in mESCs (Figure 3.2A and 3.2B). We found that dTAG-13 treatment swiftly reduced BRG1 abundance to barely detectable levels (Figure 3.2C). We next conducted a time-course MID Hi-C to interrogate acute (3h) and prolonged (6h and 24h) effects of BAF complex perturbation (Figure 3.2B). We first examined for alterations in chromatin accessibility at sites enhancer and promoter repertoires excluding sites associated with CTCF-bound sites. ATAC-seq reads were overlaid and quantified signal fold changes were compared to t=0 control (Figures 3.1A and 3.1B). We found that chromatin accessibility across enhancers were more vulnerable to BAF depletion in comparison to promoters (Figures 3.1A and 3.1B). Depletion of BAF affected chromatin accessibility at active enhancers to a similar degree as weak enhancers (Figures 3.1A and 3.1B). Interestingly, enhancers demonstrated differential responses to BAF complex perturbation, as evidenced by a wide range of changes in chromatin accessibility (Figure 3.1B). To distinguish between the different responses, we segregated enhancers based on accessibility changes of BAF depletion (3h) into three subsets: BAF-dependent (BD) enhancer (5574, more than 8-fold reduction), BAF-partial-dependent (PD) enhancer (25828, more than 1.5-fold but less than 8-fold reduction), and BAF-independent (BI) enhancer (8970, less than 1.5-fold reduction) (Figure 3.1B). We found that the major reduction in chromatin accessibility of BD and PD enhancers occurred by 3h when the decline in BRG1 plateaued whereas accessibility changes of BI enhancers were moderate during the entire course of perturbation (Figure 3.2D). Consistently, BRG1 occupancies as determined by BRG1 ChIP-seq (Chronis et al., 2017) at BD enhancers were higher than BI enhancers (Figure 3.2E). Notably, DNA motif analysis for the BD and BI enhancer repertoires revealed striking differences in motif constitution (Figure 3.1C). Specifically, we found that BD enhancers were enriched for the motifs for OCT4/SOX2/TCF/NANOG, master pluripotent regulators whose functions are strongly dependent on the BAF complex (esBAF). In contrast, BI enhancers were enriched for the motifs of KLF1, SP5, and NFY that were reported to regulate naive pluripotency. In sum, enhancer sequence dictates selective recruitment of chromatin remodelers.

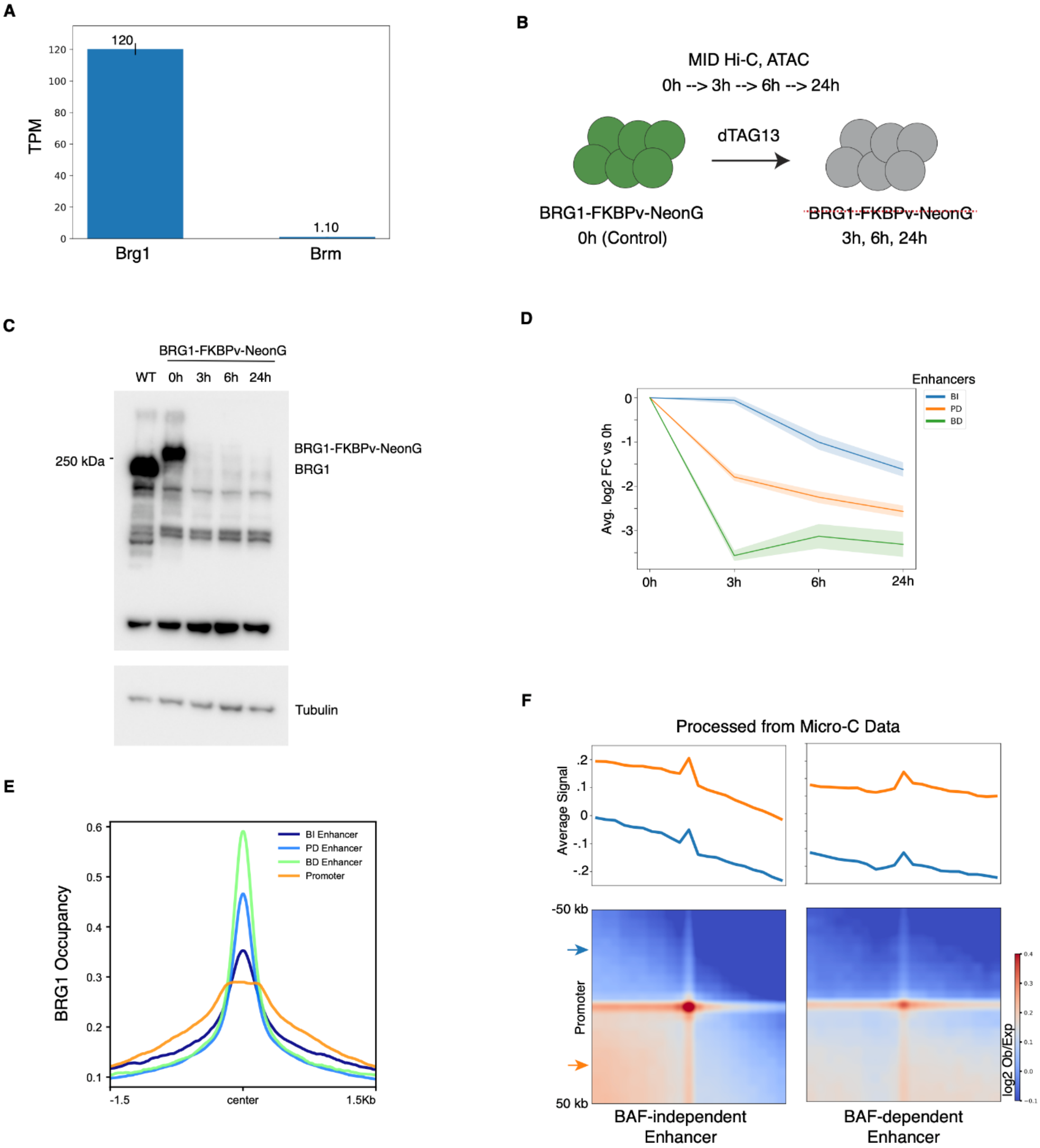
Figure 3.1. Enhancer dependencies on BAF correlate to the insulation potential. (A) Tornado plots showing the signals of the time-course ATAC-seq, H3K4me3, H3K27ac, H3K4me1, and CTCF at the accessible regions detected from the control (0h). (B) Violin plot showing the fold changes in the chromatin accessibilities for the treatment versus the control. (C) Enriched motifs at BI and BD enhancers. (D) Line plots displaying PolII occupancies at enhancers and promoters. (E) Line plots displaying the column-wise means calculated from the upper (blue) and lower (orange) halves of the APA matrices in (F). (F) APA at the E-P interaction subsets (by column) for MID Hi-C dataset.


A
BRG1-Degron mESC
ATAC
H3 K4me3
H3 K27ac
H3 K4me1
CTCF
ATAC Peaks (n=51,390)
-1.5kb center 1.5kb
dTAG13
0h
3h
6h
24h
B
log2 Fold Change (FC) vs 0h
-6
-4
-2
0
2
3h
6h
24h
Promoter (n=11,251)
Active Enhancer (n=4,256)
Weak Enhancer (n=35,883)
FC: 1/8 1/1.5
BAF Dependency:
dependent (BD)
partial dependent (PD)
independent (BI)
C
BI Enhancers
KLF1 1e-184
SP5 1e-180
NFY 1e-170
BD Enhancers
OCT4/SOX2/TCF/NANOG 1e-461
SOX3 1e-288
D
RNAP Occupancy
BI Enhancer
PD Enhancer
BD Enhancer
Promoter
-1.5 KB
center
1.5 KB
E
Average Signal
-50 kb
50 kb
F
-50 kb
Promoter
50 kb
log2 Ob/Exp
.3
0
-.3
BAF-independent Enhancer
BAF-dependent Enhancer

3.2 BRG1 Permits Genomic Interactions Beyond the Enhancer Boundary

The data described above indicates that enhancer repertoires can be segregated into BI-enhancers and BD-enhancers. These findings raise the question whether BI-enhancers and BD-enhancers instruct distinct patterns of chromatin folding. As a first approach to address this question, we aggregated MID Hi-C contacts at paired promoters and BI- and BD-enhancers. We found that BI-enhancers were associated with a distinctively insulated chromatin domain structure (Figure 3.1F). The BI-enhancer instructed insulation was characterized by elevated contact frequencies within the domain and marked by stripes emanating from enhancers and promoters that intersected at focal points associated with promoter-enhancer interactions (Figure 3.1F). We found that BI enhancers were able to insulate chromatin interactions crossing the enhancer boundary, albeit weaker when compared to promoters (Figures 3.1E and 3.1F). In contrast, we found that chromatin domains associated with BD enhancers did not insulate against genomic interactions emanating from the promoter or genomic regions located between the promoter and enhancers (Figures 3.1E and 3.1F). Likewise, using Micro-C rather than MID Hi-C we found that BD enhancers differ from BI-enhancers in their ability to enforce insulation beyond the enhancer region (Figure 3.2F). Recent studies suggested that RNA polymerase II occupancy at transcription start sites interferes with loop extrusion resulting in unidirectional stripes and insulations of genomic interactions emanating from the promoter. Consistent with this model, we found that RNA polymerase II occupancy was enriched at BI-enhancers but not at BD-enhancers (Figure 3.1D). Taken together, these data indicate that the chromatin remodeler ATPase BRG1 enables genomic interactions emanating from promoters beyond enhancer boundary elements.

Figure 3.2. BAF Complex Perturbation Uncovers BAF-Dependent and BAF-Independent Enhancers. (A) Bar plot showing Brg1 is the only expressed BAF ATPase in mESCs quantified by RNA-seq. (B) Schematic showing the design of experiments. (C) Western blots showing BRG1 protein is completely degraded after applying dTAG13. (D) Trend lines showing the fold changes (vs. 0h) in the chromatin accessibilities (ATAC-seq) during BRG1 depletion. (E) Line plots displaying BRG1 occupancies at enhancers and promoters. (F) APA at the E-P interaction subsets (by column) for Micro-C dataset (heatmap at the bottom). Column-wise mean (along the x axis centering on enhancer) calculated from the upper (blue) and lower (orange) halves of the APA matrices were plotted on the top panel.


A
TPM
120
100
80
60
40
20
0
120
1.10
Brg1
Brm
B
MID Hi-C, ATAC
0h --> 3h --> 6h --> 24h
dTAG13
BRG1-FKBPv-NeonG
0h (Control)
~~BRG1-FKBPv-NeonG~~
3h, 6h, 24h
C
BRG1-FKBPv-NeonG
WT
0h
3h
6h
24h
250 kDa
BRG1-FKBPv-NeonG
BRG1
Tubulin
D
Enhancers
BI
PD
BD
Avg. log2 FC vs 0h
0
-1
-2
-3
0h
3h
6h
24h
E
BRG1 Occupancy
0.6
0.5
0.4
0.3
0.2
0.1
BI Enhancer
PD Enhancer
BD Enhancer
Promoter
-1.5
center
1.5Kb
F
Processed from Micro-C Data
Average Signal
.2
.1
0
-.1
-.2
-50 kb
Promoter
50 kb
BAF-independent Enhancer
BAF-dependent Enhancer
log2 Ob/Exp
0.4
0.3
0.2
0.1
0.0
−0.1

3.3 BAF Complex Orchestrates the Assembly of Enhancer-Promoter Interaction Hubs

To investigate how BAF orchestrates the E-P interaction hub, we performed a PCA analysis for changes in E-P interaction strength during the course of BRG1 depletion. We found that PCA values associated with loss of BRG1 (3h, 6h, and 24h) clustered but segregated from cells not depleted for BRG1 (0h) (Figure 3.3A). Accordingly, E-P strengths segregated for BI, PD, and BD enhancer repertoires at 0h significantly differed from that observed for 3h, 6h and 24h BRG1-depletion (Figure 3.3B). Conversely, BRG1-depletion for 3 hours when compared to 6 and 24 hours did not reveal significant differences in E-P strength (Figure 3.3B). Thus, a new state of the E-P interactome was established at 3h and maintained beyond 3h, indicating that the changes in E-P interactions were tightly associated with the loss of BAF abundance.

To determine how during the course of BRG1 depletion the EP interaction hub is altered, we generated aggregated contact maps for BD, PD, and BI enhancer repertoires interacting with promoters. We found that the degree of reduction in E-P interaction strengths (intensity at the E-P focal point) correlated well with that of acute changes in chromatin accessibility across enhancers at 3h (Figures 3.3C and 3.4A). Specifically, E-P interactions at BD enhancers were nearly abrogated (Figures 3.3D and 3.4B). Conversely, we found that PE interaction strength at BI enhancers was only modestly attenuated (Figure 3.3C). Despite considerable reductions in the strength upon loss of BRG1 abundance, E-P interactions were retained across the PD enhancer repertoire (Figure 3.3C). Thus, abolishment of E-P interaction requires eradication of factors assembled at the enhancer platform that sequester the BAF complex.

We next focused our analysis on the role of BAF in instructing genomic interactions that span the entire E-P domain. To increase the coverage, we combined the MID Hi-C datasets of samples depleted for BAF during the course of 3, 6 and 24 hours. We found that although genomic interactions emanating from BD enhancers were virtually undetectable under BAF complex perturbation condition, genomic interactions involving the promoter regions were largely intact (Figure 3.3E). Additionally, the contact frequencies of genomic interactions among genomic regions that span the promoter stripe remained elevated above the background (Figure 3.3E). In summary, we found that chromatin accessibility across enhancer repertoires, reflecting transcription factor occupancy and recruitment of associated cofactors, instructs enhancer-communication but is dispensable for genomic interaction across the E-P hub.

Figure 3.3. BAF complex orchestrates the assembly of E-P interaction neighborhoods. (A) PCA analysis on the profiles of E-P interaction strengths. (B) Pairwise Mann-Whitney U rank tests among time points (by column) for the E-P interaction subsets (by row). (C) APA at the E-P interaction subsets (by column) for time-course MID Hi-C datasets (by row). (D) Virtual 4C displaying the contact frequencies, at 0h and 3h, from the viewpoint (eye symbol at the enhancer). The yellow shade highlights the interacting promoter. (E) APA at the E-P interaction subsets (by column) for combined BRG1-depleted MID Hi-C dataset. Line profiles are generated as described in Figure 3.1.

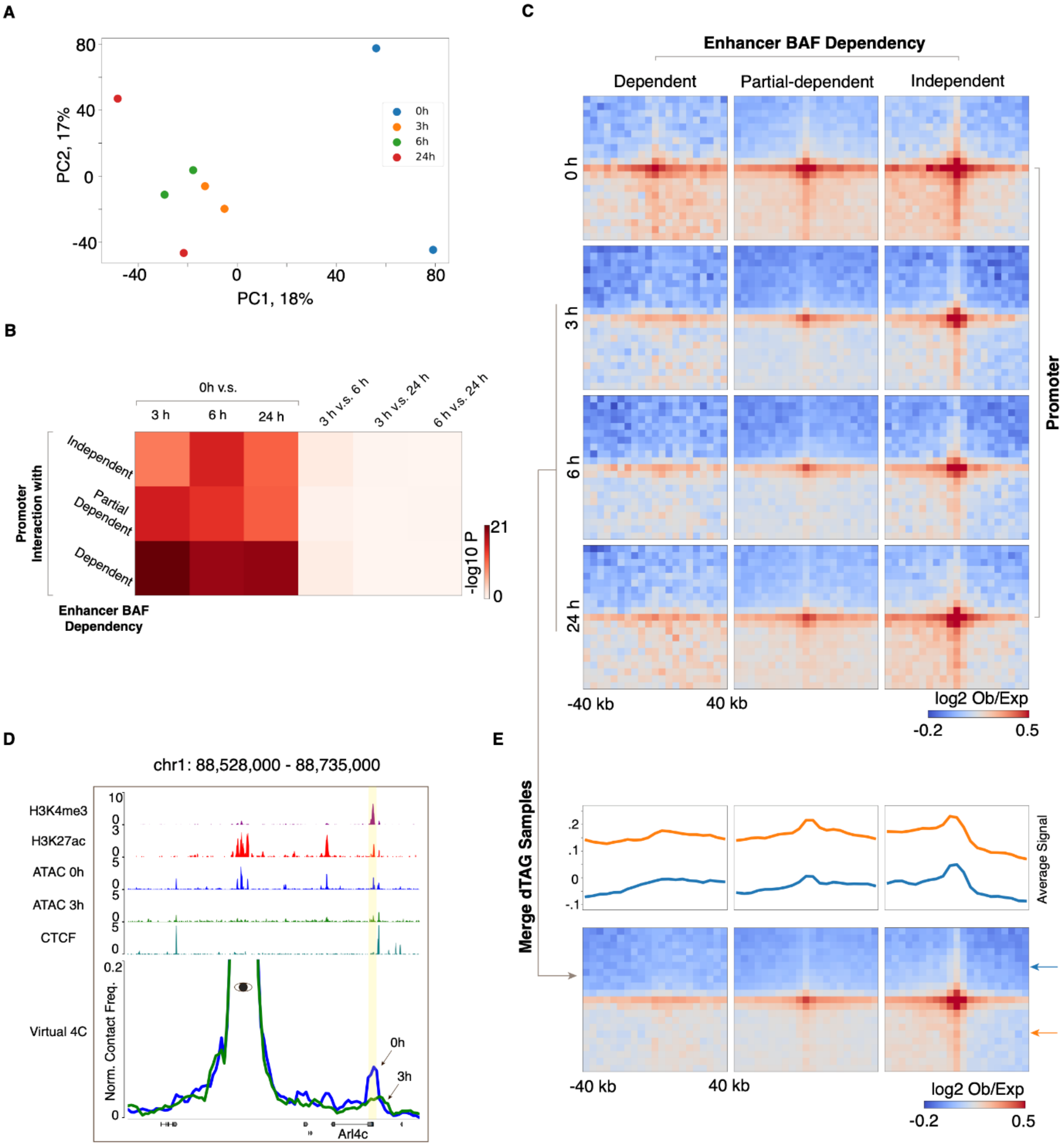
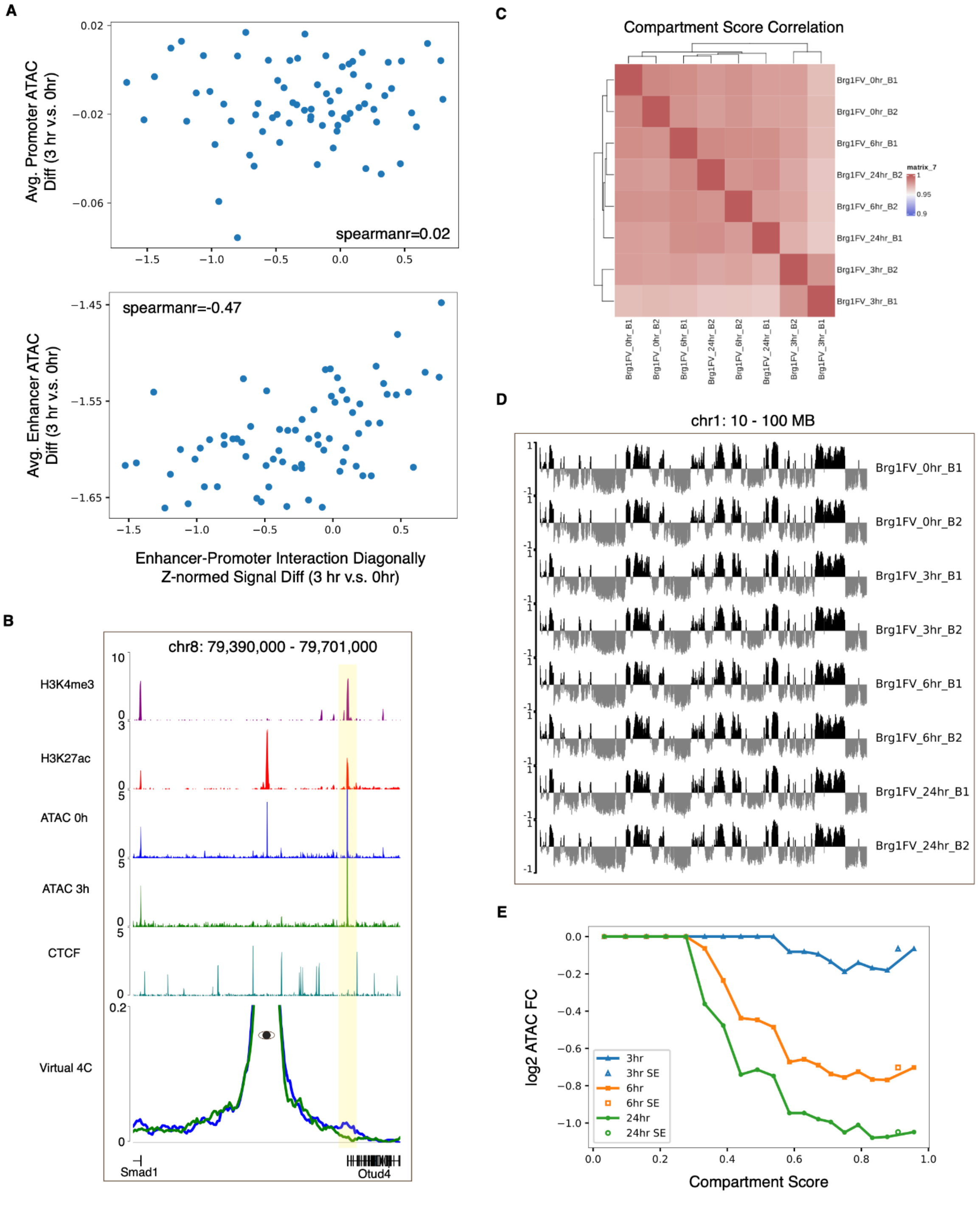

A
80
40
0
-40
PC2, 17%
-40
0
40
80
PC1, 18%
0h
3h
6h
24h
B
0h v.s.
3 h
6 h
24 h
3 h v.s. 6 h
3 h v.s. 24 h
6 h v.s. 24 h
Promoter Interaction with
Independent
Partial Dependent
Dependent
Enhancer BAF Dependency
-log10 P
21
0
C
Enhancer BAF Dependency
Dependent
Partial-dependent
Independent
0 h
3 h
6 h
24 h
Promoter
-40 kb
40 kb
log2 Ob/Exp
-0.2
0.5
D
chr1: 88,528,000 - 88,735,000
H3K4me3
H3K27ac
ATAC 0h
ATAC 3h
CTCF
Virtual 4C
10
0
3
0
5
0
5
0
5
0
0.2
0
Norm. Contact Freq.
0h
3h
Arl4c
E
Merge dTAG Samples
.2
.1
0
-.1
Average Signal
-40 kb
40 kb
log2 Ob/Exp
-0.2
0.5

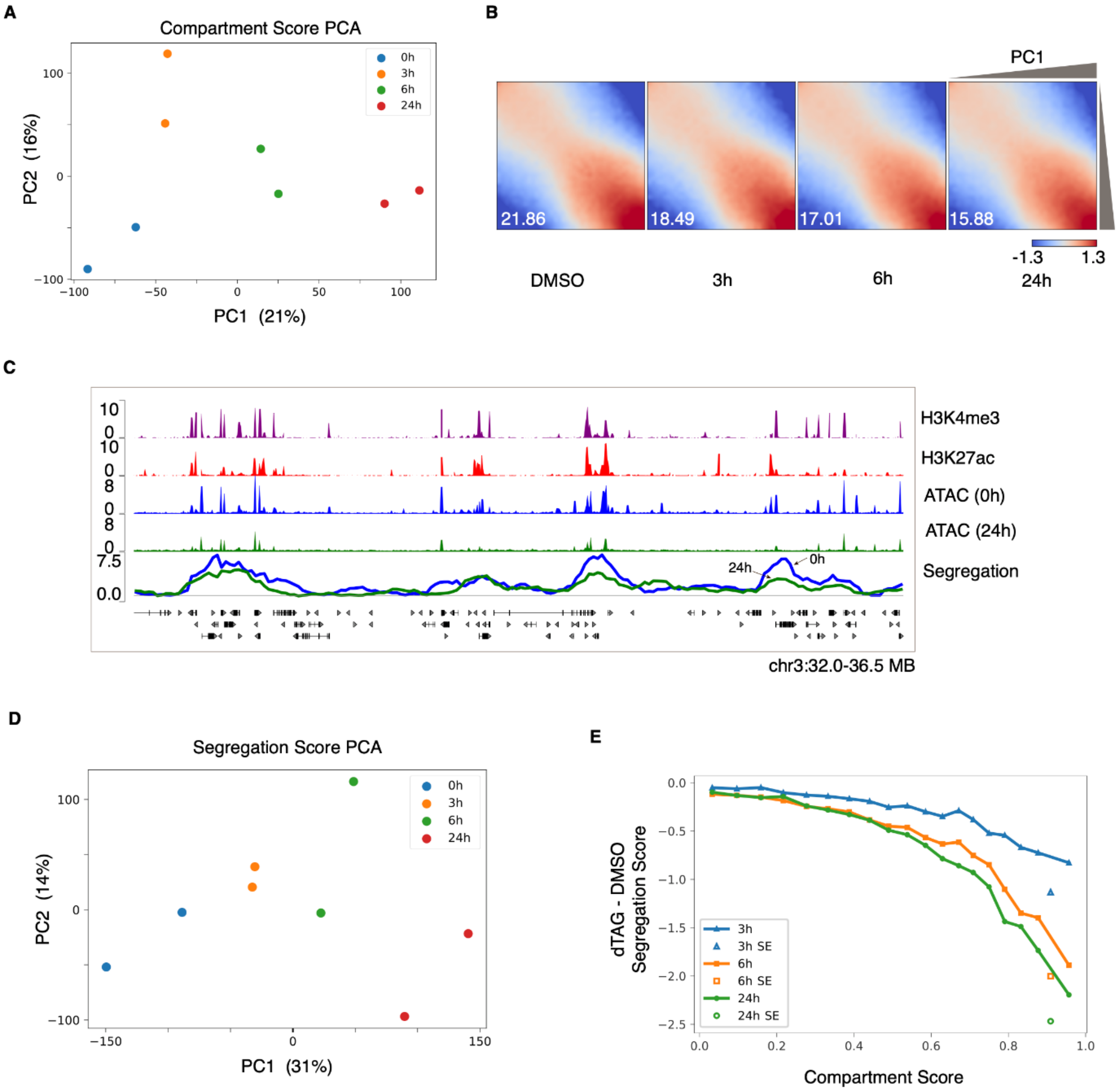
Figure 3.4. BAF perturbation affects E-P interactions and compartmental segregations. (A) Scatter plots showing the acute changes in E-P interaction strengths at 3h versus the acute changes of the anchors (top panel for promoters and bottom panel for enhancers). (B) Virtual 4C displaying the contact frequencies, at 0h and 3h, from the viewpoint (eye symbol at the enhancer). The yellow shade highlights the interacting promoter. (C) Heatmap showing the pairwise spearman correlations among the compartmentalizations for time-course MID Hi-C datasets. (D) Example tracks of the compartment scores. (E) Line plots displaying the acute fold changes (at 3h) in chromatin accessibilities versus the compartmental scores.


A
Avg. Promoter ATAC Diff (3 hr v.s. 0hr)
0.02
−0.02
−0.06
spearmanr=0.02
−1.5
−1.0
−0.5
0.0
0.5
Avg. Enhancer ATAC Diff (3 hr v.s. 0hr)
−1.45
−1.55
−1.65
spearmanr=-0.47
Enhancer-Promoter Interaction Diagonally Z-normed Signal Diff (3 hr v.s. 0hr)
B
chr8: 79,390,000 - 79,701,000
H3K4me3
10
0
H3K27ac
3
0
ATAC 0h
5
0
ATAC 3h
5
0
CTCF
5
0
Virtual 4C
0.2
0
Smad1
Otud4
C
Compartment Score Correlation
Brg1FV_0hr_B1
Brg1FV_0hr_B2
Brg1FV_6hr_B1
Brg1FV_24hr_B2
Brg1FV_6hr_B2
Brg1FV_24hr_B1
Brg1FV_3hr_B2
Brg1FV_3hr_B1
matrix_7
1
0.95
0.9
D
chr1: 10 - 100 MB
1
-1
Brg1FV_0hr_B1
Brg1FV_0hr_B2
Brg1FV_3hr_B1
Brg1FV_3hr_B2
Brg1FV_6hr_B1
Brg1FV_6hr_B2
Brg1FV_24hr_B1
Brg1FV_24hr_B2
E
log2 ATAC FC
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
3hr
3hr SE
6hr
6hr SE
24hr
24hr SE
0.0
0.2
0.4
0.6
0.8
1.0
Compartment Score

3.4 The BAF Complex Instructs Compartmental Segregation

At the coarse 3D genomic scale, chromatin is segregated into active (A) and inactive (B) compartments (Lieberman-Aiden et al., 2009). Since activate enhancers are primarily located in the A compartment, we examined whether the global reduction of chromatin accessibility caused by BAF depletion modulates compartmental organization. We found that during the course of BAF depletion (3-24 hours) A/B compartmental remained largely intact (pairwise spearman > 0.94, Figures 3.4C and 3.4D). However, upon closer inspection using PCA analysis we found significant changes in A/B compartmentalization during the course of BAF depletion (Figure 3.5A). To determine whether these changes were orchestrated by changes in the compartmental segregation we generated saddle plots (Gibcus et al., 2018). This analysis captured a progressive increase in contact frequencies at the inter-compartmental interactions, indicating weakening in compartmental segregation that became more prominent during the course of BAF depletion (Figure 3.5B). To determine to what degree compartmental segregation changed upon BAF depletion, segregation scores, which quantifies the strength difference between the intra- and inter-compartmental interaction, were computed for genomic bins spanning 25 Kbp. PCA analysis on the segregation score profiles also detected time-dependent compartmental segregation differences among samples but a higher percentage of sample-wise variation were captured compared to compartment scores (Figure 3.5D). We found upon BAF depletion a significant decline in compartmental segregation across genomic regions that were associated with a significant decline in chromatin accessibilities (Figure 3.5C). To determine how changes during the course of BAF depletion in compartmental segregation relate to the positioning in the A compartment, we computed segregation differences as a function of compartment score. We found that upon BAF depletion the loss of segregation was increased overtime for regions of higher compartmental scores (Figures 3.5E), correlated well with the reduction of chromatin accessibility (Figures 3.4E). Taken together, these data indicate that the BAF complex enforces compartmental segregation.

Figure 3.5. BAF complex perturbation reduces compartmental segregation. (A) PCA analysis on the profiles of compartmental scores. (B) Saddle plot displaying the segregation pattern for the time-course MID Hi-C datasets. The summarized statistics (bottom left corner) were calculated from the difference of top 20% A-A and top 20% B-B interactions (intra-compartment, diagonal) versus A-B interactions (inter-compartment, off-diagonal) (Gibcus et al., 2018). (C) Example track of the compartmental segregation and epigenetic marks. (D) PCA analysis on the profiles of segregation scores. (E) Line plots displaying the acute fold changes (at 3h) in compartmental segregation versus the compartmental scores.


A
Compartment Score PCA
PC2 (16%)
PC1 (21%)
0h
3h
6h
24h
B
PC1
21.86
18.49
17.01
15.88
-1.3
1.3
DMSO
3h
6h
24h
C
H3K4me3
H3K27ac
ATAC (0h)
ATAC (24h)
Segregation
24h
0h
chr3:32.0-36.5 MB
D
Segregation Score PCA
PC2 (14%)
PC1 (31%)
E
dTAG - DMSO Segregation Score
Compartment Score
3h
3h SE
6h
6h SE
24h
24h SE

3.5 Summary

In this chapter, we identified two subsets of enhancers based on their dependencies on BAF complex as revealed by the acute changes in the chromatin accessibility after BRG1 depletion. They are not only enriched for different transcription factor motifs, but the underlying E-P interacting neighborhoods also assume distinct configurations. Lastly, we observed global reduction in chromatin accessibilities weaken the compartmental segregation in the nucleus.

3.6 Material and Methods

FKBPv Cell Line Generation

The constructions of the guide RNA (sgRNA) and the donor plasmid for inserting FKBPv to Brg1 endogenous locus were described in our previous paper (Zhu et al., 2020). For generation of FKBP12F36V cell lines, mESCs cells were seeded the day before transfection on a 10 cm tissue culture plate. 12 ug of plasmid DNA comprising equimolar amounts of sgRNA and donor plasmid were transfected using 40 ul of Lipofectamine 3000 with 24 ul of P3000 reagent (Thermo Fisher, L3000008). Two days after transfection, the transfected cells were selected in 10 ug/mL of Blasticidin (Cayman Chemical, NC1445974) for a week. Single cells from the EYFP positive population with successful insertions were then sorted using a BDFACSAria Fusion Sorter in 96-well round bottom plates. Single cells were grown for ~7 days to form single clones. Genotyping primers (Zhu et al., 2020) were used to screen for homozygous clones. The established cell lines were treated with dTAG13 (CAS 2064175-41-1). Efficiency of acute degradation of BRG1 proteins was examined by western blots as previously described (Zhu et al., 2020).

ATAC-seq Analysis

Adaptors in pair-end (PE) sequencing reads were trimmed using Trimmomatic (Bolger et al., 2014). Trimmed PE reads were aligned to mm10 genome using bowtie2 (Langmead and Salzberg, 2012) with flags `--very-sensitive -I 10 -X 700 –dovetail`. Only mapped fragments with sizes between 0 to 200 BP were kept for further analysis. Peaks were called by MACS2 with flags `-g mm -f BEDPE -q 0.01 –SPMR –keep-dup auto`. Then, functionalities of deeptools (Ramírez et al., 2014) were used to analyze enrichment signals. Briefly, density tracks were generated by `bamCoverage` to be visualized in UCSC genome browser or by pyGenomeTracks (Lopez-Delisle et al., 2020). Clustering of peaks, based on H3K4me3, H3K27ac, and CTCF signal, and tornado plots were generated by `computeMatrix` and `plotHeatmap` to identify enhancers and promoters devoid of CTCF occupancies. Summarized signal profiles were plotted by `plotProfile`. In addition, quantification of peaks was performed through counting the number of reads overlapping with peaks using PyRanges (Stovner and Sætrom, 2019). Homer (Heinz et al., 2010) was used to search enriched motifs within 200 BP from the peak centers.

Compartmental Segregation Analysis

We calculated the averaged interaction strengths with A and B compartments separately for a 25 Kbps bin. Then, according to the compartment that the bin belongs to, we computed the ratio of intra- versus inter-compartmental interaction strengths, a.k.a compartmental segregation score. Segregation tracks were plotted by pyGenomeTracks and PCA analysis was performed as described above. Also, we calculated the averaged changes in segregation scores by stratifying bins according to the compartmental scores.

3.7 Acknowledgements

For their contributions in chapter 3, thanks to: Paula Pham, Han Han, Jeff Jiajing Zhou, Yina Zhu, Wenqi Wang, Ken Cho and Cornelis Murre. The material is in this chapter currently prepared for submission for publication.

CHAPTER 4

Chromatin Structural and Transcriptional Responses of Neutrophils to Stimuli

4.1 Responses of Human Neutrophils to Microbial Challenges

Upon microbial challenge, human neutrophils undergo rapid changes in nuclear architecture and chromatin folding to orchestrate an immediate inflammatory gene program

Matthew Denholtz,[1,5] Yina Zhu,[1,5] Zhaoren He,[1] Hanbin Lu,[1] Takeshi Isoda,[1,4] Simon Döhrmann,[2] Victor Nizet,[2,3] and Cornelis Murre[1]

[1]Division of Biological Sciences, Department of Molecular Biology, University of California at San Diego, La Jolla, California 92039, USA; [2]Department of Pediatrics, University of California at San Diego School of Medicine, La Jolla, California 92093, USA; [3]Skaggs School of Pharmaceutical Sciences, University of California at San Diego, La Jolla, California 92093, USA

Differentiating neutrophils undergo large-scale changes in nuclear morphology. How such alterations in structure are established and modulated upon exposure to microbial agents is largely unknown. Here, we found that prior to encounter with bacteria, an armamentarium of inflammatory genes was positioned in a transcriptionally passive environment suppressing premature transcriptional activation. Upon microbial exposure, however, human neutrophils rapidly (<3 h) repositioned the ensemble of proinflammatory genes toward the transcriptionally permissive compartment. We show that the repositioning of genes was closely associated with the swift recruitment of cohesin across the inflammatory enhancer landscape, permitting an immediate transcriptional response upon bacterial exposure. We found that activated enhancers, marked by increased deposition of H3K27Ac, were highly enriched for cistromic elements associated with PU.1, CEBPB, TFE3, JUN, and FOSL2 occupancy. These data reveal how upon microbial challenge the cohesin machinery is recruited to an activated enhancer repertoire to instruct changes in chromatin folding, nuclear architecture, and to activate an inflammatory gene program.

[*Keywords*: gene expression; inflammation; neutrophils; nuclear architecture]

Supplemental material is available for this article.

Received October 11, 2019; revised version accepted December 12, 2019.

The organization of the human genome within the nucleus is central to the control of gene expression and thus cell identity and function. At the largest scale, the genome is folded into chromosome territories that, with the exception of nucleoli, rarely intermingle. However, chromosomes are not randomly distributed across the nucleus. Large and gene-poor chromosomes are predominantly positioned at the lamina, whereas small and gene-rich chromosomes concentrate in the nuclear interior (Fritz et al. 2016). Chromosomes themselves fold into loop domains that physically associate to establish the transcriptionally repressive or inert heterochromatic B compartment or transcriptionally permissive euchromatic A compartment (Lieberman-Aiden et al. 2009; Dixon et al. 2012). The heterochromatic compartment is highly enriched at the nuclear lamina, whereas the euchromatic compartment is positioned in the nuclear interior (Kosak et al. 2002).

Loop domains are established in part by the CTCF protein (Dixon et al. 2012; Rao et al. 2014). Convergently oriented pairs of CTCF-bound loci can form CTCF-anchored loops generated by recruitment of the cohesin complex (Nora et al. 2017; Rao et al. 2017). Cohesin is loaded onto transcribed regions located throughout loop bodies (Busslinger et al. 2017). Once sequestered, the cohesin complex extrudes chromatin in a progressive manner until a pair of convergent CTCF-bound sites are reached, a process termed loop extrusion (Fudenberg et al. 2016). Gene promoters connected to transcriptional enhancers

---

[4]**Present address: Department of Pediatrics and Developmental Biology, Tokyo Medical and Dental University, Tokyo, Japan.**
[5]**These authors contributed equally to this work.**
Corresponding author: cmurre@ucsd.edu
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.333708.119.

by CTCF-mediated loops tend to be highly expressed (Rao et al. 2014), and CTCF occupancy at nearby sites contributes to the maintenance of gene expression and stable chromatin structure (Nora et al. 2017; Rao et al. 2017; Schwarzer et al. 2017; Bintu et al. 2018). Finally, recent studies have demonstrated that cohesin-dependent looping is closely associated with the control of inducible gene expression (Bruno et al. 2018).

Human neutrophils are abundant, short-lived circulating white blood cells that are critical first-responders to infection and tissue damage. Upon injury or infection, neutrophils exit the circulation via extravasation, migrate toward damaged tissues or infectious foci, phagocytose small pathogens, release reactive oxygen and nitrogen species, and extrude their chromatin as cytotoxic granule-laced extracellular traps (NETs). In addition to their direct role in killing invading pathogens, activated neutrophils rapidly induce the expression of a wide range of cytokines and chemokines to orchestrate an immediate inflammatory response (Ley et al. 2018).

The nuclei of polymorphonuclear (PMN) neutrophils are composed of multiple distinct but internally continuous lobes allowing them to swiftly migrate between (paracellular route) or through (transcellular route) endothelial cells that line blood vessels and interstitial spaces of tissues while maintaining their nuclear integrity (Olins et al. 2009; Muller 2013; Rowat et al. 2013). The Lamin B receptor (LBR) is an important determinant for imposing a multi-lobular nuclear architecture on neutrophils (Hoffmann et al. 2002; Shultz et al. 2003). Neutrophils of mice deficient in the *Lbr* gene fail to adopt a multi-lobular nuclear shape (Shultz et al. 2003), and mouse neutrophilic cell lines lacking *Lbr* cannot form characteristic toroidal nuclei during differentiation (Zhu et al. 2017). Similarly, humans with *LBR* mutations manifest the Pelger-Huët anomaly, characterized by a reduction in nuclear lobe number in granulocytes (Hoffmann et al. 2002).

Chromatin folding in murine neutrophils is highly enriched for remote genomic interactions, primarily involving heterochromatic regions. These interactions span vast genomic distances resulting in large-scale chromosome condensation. Terminal differentiation of murine neutrophils is also associated with the relocation of centromeres, pericentromeres, telomeres, LINE elements, and ribosomal DNA from the nuclear interior to the nuclear lamina, a process that requires the *Lbr* gene (Zhu et al. 2017). As neutrophils differentiate, the LBR deforms the malleable nuclear envelope by wrapping it around the heterochromatic component of the neutrophil genome, resulting in its characteristic lobed shape.

Upon reaching a tissue site of infection, neutrophils neutralize bacteria in multiple ways: (1) engulfment through phagocytosis, (2) degranulation to release microbicidal factors into the extracellular space, (3) release of extracellular traps or NETs that are composed of extruded chromatin fibers and antimicrobial factors, and (4) rapid induction of cytokine gene expression to coordinate a broader immune response (Brinkmann et al. 2004; Ley et al. 2018). To detect and respond appropriately to diverse invading pathogens, neutrophils express a variety of pattern recognition receptors including cell-surface and endolysosomal Toll-like receptors (TLRs), C-type lectin receptors, and formyl peptide receptors, among others. Once activated, a variety of downstream signaling pathways converge on the NF-κB and AP1 transcription factors to induce an inflammatory gene program including the cytokines and chemokines IL-8/CXCL8, TNFα, IL-1β, IL-17, and IFNγ (Garcia-Romo et al. 2011; Thomas and Schroder 2013; Tecchio et al. 2014).

The mechanisms by which pathogen-sensing pathways interface with the neutrophil genome to induce a rapid and stimulant-appropriate inflammatory gene expression program remain unclear. Here we found that human neutrophil genomes display highly segmented compartments and contracted heterochromatin when compared with human embryonic stem cells. Upon microbe encounter, a specific subset of modestly euchromatic subdomains, spatially segregated from the highly euchromatic A compartment, displayed strengthening of their euchromatic character, and relocalized from a perinuclear envelope position toward the nuclear interior. Prominent among the regions that repositioned during human neutrophil activation were gene loci associated with an activated neutrophil-specific gene expression program. Microbial-induced changes in long-range chromatin interactions were globally associated with rapid loss of insulation at euchromatic subdomain boundaries, as well as the formation of de novo chromatin loops linking immune response genes to pre-existing and de novo formed transcriptional enhancers. The loop-mediated juxtaposition of inflammatory genes to transcriptional enhancers upon microbial exposure was closely associated with the deposition of histone 3 lysine 27 acetylation (H3K27ac), an enhancer-associated histone modification and rapid loading (<3 h) of the cohesin complex at the subset of enhancer elements that control an inflammatory gene program. Based on these observations, we propose that the microbe-induced transcriptional signature of activated neutrophils is driven by activated enhancer repertoires. Activated enhancers marked by elevated levels of H3K27Ac, in turn, rapidly recruit the cohesin machinery to dictate changes in chromatin folding and nuclear positioning of genes associated with an inflammatory gene program.

Results

Human neutrophil development is associated with segmented compartments and contracted genomes

Neutrophil nuclei undergo dramatic morphological changes during differentiation from multipotent progenitors, with terminally differentiated neutrophil nuclei having three to five internally continuous but spatially distinct lobes (Supplemental Fig. S1A). To characterize the genomic interactions established during the development of PMN cells, neutrophils were isolated from human peripheral blood, formaldehyde-fixed, and analyzed using in situ HiC (Supplemental Table S1; Rao et al. 2014). The genomes of human neutrophils were slightly enriched for interchromosomal interactions when

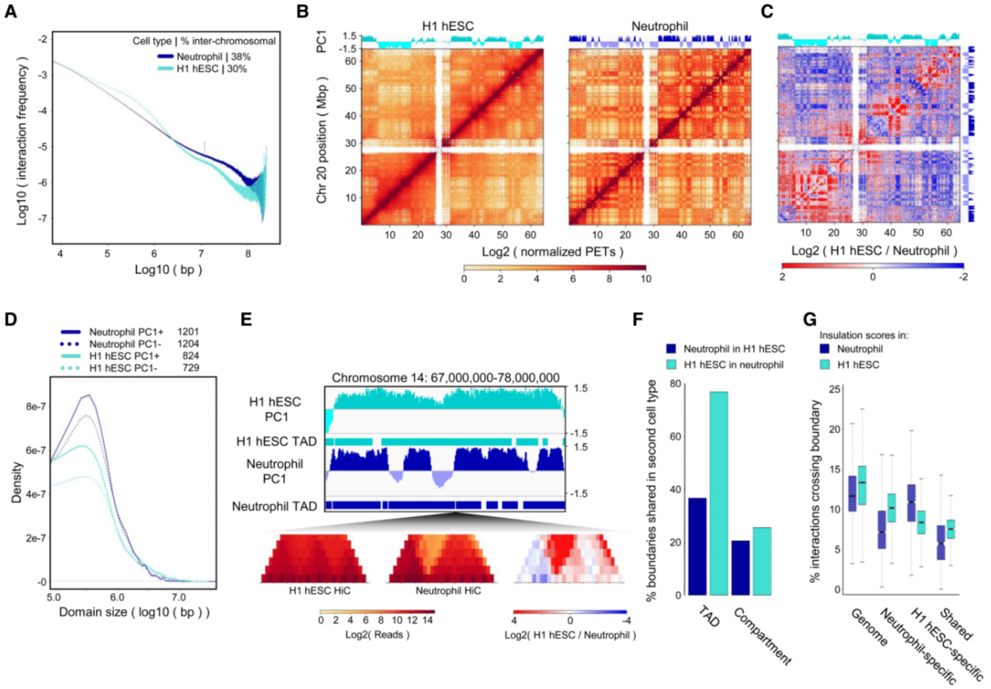


Figure 1. Heterochromatic super-contraction and segmentation of the neutrophil genome during acquisition of the PMN shape. (*A*) Neutrophil (blue) and H1 hESC (teal) chromatin interaction frequencies as a function of linear genomic distance. Percent interchromosomal paired end tags (PETs) are indicated. (*B*) Normalized HiC contact matrices for H1 hESC (*left*) and neutrophil (*right*) chromosome 20. First principal component 1 eigenvector (PC1) for each HiC matrix is displayed *above* its respective matrix. Positive PC1 values correspond to the gene-rich A compartment, negative values to the gene-poor B compartment, based on human genome build hg38. (*C*) HiC contact matrix showing the $\log_2$ fold change in normalized interactions between the H1 hESC and neutrophil matrices in B, illustrating changes in the organization of chromosome 20 during terminal differentiation from the pluripotent state and acquisition of the PMN shape. PC1 values for H1 hESCs and neutrophils are shown *above* and at the *right* of the matrix, respectively. (*D*) Density plot showing the distribution of A (solid lines) and B (dotted lines) compartment domain sizes in H1 hESCs (teal) and neutrophils (blue). Total number of domains for each data set are listed at *top*. Domains <100,000 bp were not considered. (*E*) Example of a new TAD and compartment boundary formed during differentiation and hypercompartmentalization of the neutrophil genome. (*Top*, from *top* to *bottom*) IGV tracks showing H1 hESC PC1 values, H1 hESC TADs, neutrophil PC1 values, and neutrophil TADs. (*Bottom*) Normalized HiC contact matrix of H1 hESC and neutrophil HiC matrices and a $\log_2$ fold change difference matrix at a new TAD/compartment boundary on chromosome 14. (*F*) Percent of TAD and compartment boundaries shared between H1 hESCs and neutrophils. Domain boundaries within 100 kb (< 3 windows) were considered shared. (*G*) Insulation scores in H1 hESCs and neutrophils calculated for each 40-kb bin genome-wide, at neutrophil-specific, H1 hESC-specific, and shared PC1 compartment/TAD boundaries for both cell types. Grouped pairs are all significantly different by the Wilcoxon rank sum test with log(*P*) values $<1 \times 10^{-11}$.

compared with human embryonic stem cells (hESCs) (Fig. 1A). Chromosome territories remained intact and we found no evidence of individual chromosomes being split across multiple lobes (Supplemental Fig. S1B). Notably, compared with hESCs, human neutrophils were depleted for genomic interactions that spanned <3 Mb but were enriched for interactions that covered >3 Mb (Fig. 1A).

We next constructed contact matrices for hESCs cells and human neutrophils (Fig. 1B). We found that a larger fraction of the neutrophil genome was sequestered in the B compartment when compared with hESCs (Supplemental Fig. S1C). The stereotypic plaid pattern, resulting from the spatial segregation of the A and B compartments, was much more pronounced in human neutrophils compared with hESCs (Fig. 1B). Intrachromosomal and interchromosomal interactions between A and B compartments were both less prevalent in neutrophils versus hESCs (Fig. 1C; Supplemental Fig. S1D). Conversely, long-range genomic interactions across the B compartment were significantly more extensive in human neutrophils than hESCs (Fig. 1C; Supplemental Fig. S1D). During differentiation, large genomic regions that exhibited a continuum of either positive or negative PC1 values in hESCs fragmented into smaller genomic regions that switched PC1 values in neutrophils (Fig. 1C,D). Many of the genomic regions that switched from negative to positive PC1 values during development were associated with a neutrophil-specific transcription signature, whereas those regions switching from positive to negative PC1 values were associated with silencing of lineage-inappropriate genes (Supplemental Fig. S1E–G; Supplemental Table S2). Notably, the hypersegmentation of compartment domains in the neutrophil genome established de novo loop domain and compartment boundaries (Fig. 1E). Specifically, although >75% of loop domain boundaries identified in hESCs were conserved in neutrophils, <40% of loop domain boundaries in neutrophils were present in hESCs (Fig. 1F). Overall compartment boundaries were poorly conserved between these two cell types (Fig. 1F). Genome-wide analysis of cell type-specific loop domain and compartment boundary element insulation strength confirmed this finding, indicating the existence of cell type-specific boundaries that were associated specifically with either hESCs or human neutrophils, in addition to shared boundaries (Fig. 1G). Taken together our data reveal that human neutrophils, when compared with hESCs, are characterized by a contracted genome with increased enforcement of compartmentalization and highly segmented A and B compartments.

PMA-induced activation of neutrophils rapidly modulates nuclear architecture

Upon detecting invading microbes, neutrophils rapidly activate an inflammatory-specific transcription signature. As a first approach to examine whether and how the nuclear architecture of neutrophils responds to inflammatory signals, HiC was performed on neutrophils cultured in both the absence and presence of the canonical neutrophil activator phorbol 12-myristate 13-acetate (PMA), a protein kinase C. PMA stimulation of human neutrophils resulted in a global decrease in short-range intrachromosomal interactions and a global increase interchromosomal interactions (Fig. 2A), while exerting minimal effects on A-B compartmentalization and loop domain boundaries (Fig. 2B). Likewise, PMA-induced activation did not trigger large-scale switching of genes or regulatory elements between the A and B compartments (Fig. 2C). However, further scrutiny of chromatin folding across the A compartment revealed a small but significant number of discrete genomic regions that underwent significant PMA-dependent changes from low but positive PC1 values to highly positive PC1 values, indicating an increase in euchromatic character (Fig. 2C). We refer to these regions as PMA ΔPC1 domains (Fig. 2C; Materials and Methods). Notably, PMA ΔPC1 domains were strongly enriched for genes implicated in the neutrophil defense response, including genes downstream from key innate immune receptors such as the complement receptors, FCγ receptor, and dectin-2, as well as genes implicated in cell migration and regulation of lysosomal pH (Fig. 2D; Supplemental Table S2).

Analysis of intrachromosomal HiC contact matrices revealed few significant changes across chromosome 18 following PMA-induced activation (Fig. 2E). PMA ΔPC1 domains, however, often displayed changes in interactions with euchromatin, both in their immediate vicinity, as well as across chromosome 18 (Fig. 2F). Direct measure of genome-wide changes in chromatin organization showed that in activated neutrophils PMA ΔPC1 domains showed large-scale changes in contact frequencies (Fig. 2G). Specifically, ∼25% of PMA ΔPC1 domains fell within the top 10% most differentially interacting genomic regions (Fig. 2G). Likewise, PMA ΔPC1 domains on average displayed significantly lower chromatin interaction correlation with unstimulated neutrophils when compared with the genome as a whole (Fig. 2H). Taken together, these data indicate that in PMA-activated neutrophils genic and intergenic domains associated with innate immune genes increase their euchromatic character and undergo alterations in remote genomic interactions.

Upon microbial exposure, a subset of neutrophil inflammatory genes increase their euchromatic character

To validate the alterations in neutrophil euchromatic character using a physiologically relevant stimulus, human neutrophils were cultured in the presence of live *Escherichia coli* bacteria for a period of 3 h. *E. coli* cocultured neutrophils were isolated, formaldehyde cross-linked, and analyzed using HiC. Genomes of human neutrophils cultured in the presence of *E. coli* only displayed minor alterations in contact frequencies, maintained overall compartment and loop domain structures (Fig. 3A,B), and remained essentially free from detectable A–B compartment switching (Fig. 3C). However, similar to PMA-activated neutrophils, a distinct subset of genomic regions positioned in the A compartment displayed a substantial increase in euchromatic character upon *E. coli* encounter (Fig. 3C, *E. coli* ΔPC1 domains). Notably, the *E. coli* ΔPC1 domains included genes encoding for cytokines and chemokines, genes associated with neutrophil degranulation, and genes linked with the inflammatory response (Fig. 3D; Supplemental Table S2).

Similar to PMA-activated neutrophils, *E. coli* cocultured neutrophils showed few large-scale changes in chromatin organization compared with unstimulated neutrophils (Fig. 3E). *E. coli* ΔPC1 domains, however, showed dramatic increases in genomic interactions involving neighboring euchromatic regions, as well as the remainder of the chromosome upon coculture with *E. coli* (Fig. 3F). Similar to PMA ΔPC1 domains, *E. coli* ΔPC1 domains were among the most restructured genomic regions in response to *E. coli*, with 25% of *E. coli* ΔPC1 domains assigned to the top 15% of the most differentially interacting regions globally (Fig. 3G). *E. coli* ΔPC1 domains overall displayed significantly lower correlation with unstimulated neutrophil genome structure than the remainder of the genome (Fig. 3H). These data indicate that upon microbial exposure, a subset of genes associated with an inflammatory response increase their euchromatic character.

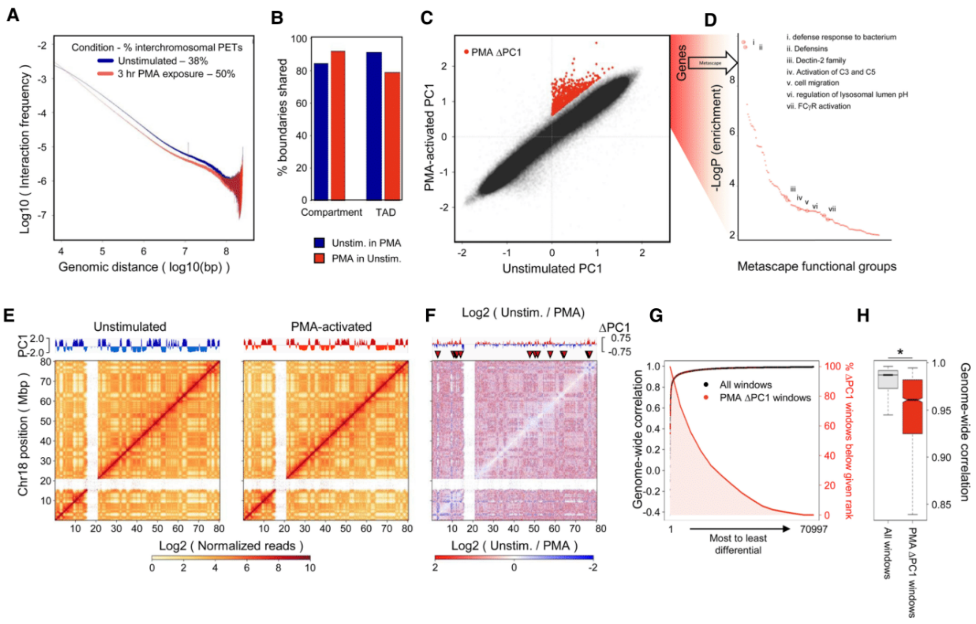


Figure 2. PMA activation increases euchromatic character at distinct genomic loci encoding for neutrophil activation genes. (*A*) Unstimulated (blue) and PMA-activated neutrophil (red) (3 h) chromatin interaction frequencies as a function of linear genomic distance. Percent interchromosomal PETs are indicated. (*B*). Percentages of TAD and compartment boundaries shared between unstimulated and PMA-activated neutrophils. (*C*) Scatter plot comparing 10-kb-windowed PC1 values between unstimulated and PMA-activated neutrophils. PMA ΔPC1 domains are marked in red. (*D*) Metascape-defined functional groups enriched for genes found within 100 kb of PMA ΔPC1 domains. Complete metascape results can be found in Supplemental Table S2. (*E*) HiC contact matrices of chromosome 18 for unstimulated (*left*) and PMA-activated (*right*) neutrophils, 100-kb resolution. Respective PC1 values are shown *above* each matrix. (*F*) Difference matrix showing the $\log_2$ fold change in normalized interactions between unstimulated and PMA-activated neutrophils. PC1 differences are shown at *top*, PMA ΔPC1 domains are marked by red triangles. (*G*) Black and red points indicate genome-wide interaction correlations for each 40-kb bin in the genome, ranked from most to least differential, *left* to *right*. Bins containing PMA ΔPC1 domains are marked with red points. The red line with red shading indicates the proportion of total PMA ΔPC1 domains found at a given rank or lower, showing a preference for PMA ΔPC1 domains to fall in genomic regions with the most differential chromatin interactions upon PMA stimulation. (*H*) Box plots of genome-wide interaction correlation values for PMA ΔPC1 domains and the remainder of the genome during PMA stimulation. Box plot outliers are not shown. (*) Wilcoxon rank-sum test *P*-value $<2\times10^{-16}$.

We next sought to ascertain to what degree ΔPC1 domains differed between stimuli. *E. coli* ΔPC1 domains only partially overlapped with PMA ΔPC1 domains (Supplemental Fig. S2A). The identities of genes in ΔPC1 domains also depended on the stimulus that neutrophils encountered. *E. coli*-specific ΔPC1 domains were highly enriched for chemokine and cytokine genes as well as genes involved in chemotaxis (Supplemental Fig. S2B). In contrast, PMA-specific ΔPC1 regions were enriched for defensin gene clusters (Supplemental Fig. S2B). These data suggest that the changes in euchromatic character regulate stimulant-appropriate inflammatory responses. Supporting this hypothesis, genes residing in stimulus-specific ΔPC1 domains underwent stimulus-specific changes in gene expression. Genes in *E. coli* ΔPC1 domains were more highly expressed upon *E. coli* encounter than upon PMA stimulation, whereas genes in PMA ΔPC1 domains were more highly expressed upon PMA stimulation than during *E. coli* coculture (Supplemental Fig. S2C). Taken together, these data indicate that neutrophil activation enhances the euchromatic character of a subset of inflammatory response gene loci in a stimulus-dependent manner.

Rapid assembly and relocalization of a CXCL transcriptional hub upon E. coli *encounter*

To determine how euchromatic character is strengthened upon microbial activation, we focused on an archetypal *E. coli* ΔPC1 domain containing inflammatory-specific genes encoded within the extended *CXCL* locus. The *CXCL* locus spans a cluster of genes encoding a class of chemokines that include *CXCL8* (IL8), *CXCL1*, and *CXCL2* (MIP2α), each of which is rapidly induced when

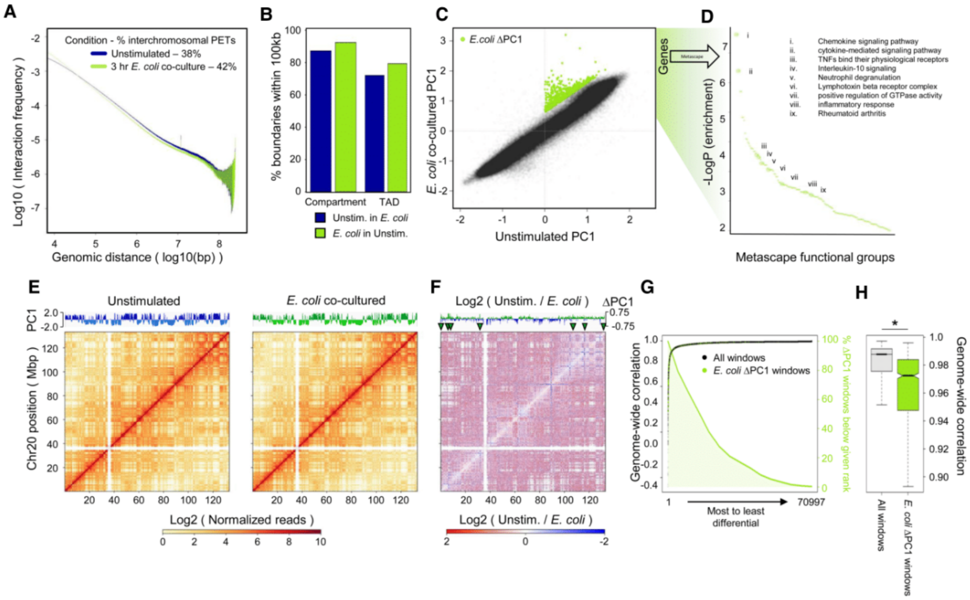


Figure 3. Neutrophil *E. coli* coculture increases euchromatic character at distinct loci encoding for neutrophil pathogen response genes. (*A*) Unstimulated (blue) and *E. coli* cocultured neutrophil (green) chromatin interaction frequencies as a function of linear genomic distance. Percent interchromosomal PETs are indicated. Note: Unstimulated neutrophil data is identical to that shown in Figure 2A and is shown here to illustrate differences between unstimulated and *E. coli* cocultured neutrophil data. Neutrophils were cultured in the presence of *E. coli* for a 3-h period. (*B*) Percent of TAD and compartment boundaries shared between unstimulated and *E. coli* cocultured neutrophils. (*C*) Scatter plot comparing 10-kb-windowed PC1 scores between unstimulated and *E. coli* cocultured neutrophils. *E. coli* ΔPC1 domains are marked in green. (*D*) Metascape-defined functional groups enriched for genes found within 100 kb of *E. coli* ΔPC1 domains. Complete metascape results are described in Supplemental Table S2. (*E*) HiC contact matrices of chromosome 20 for unstimulated (*left*) and *E. coli* cocultured (*right*) neutrophils, 100-kb resolution. Respective PC1 values are shown *above* each matrix. (*F*) Difference matrix showing the $\log_2$ fold change in normalized interactions between unstimulated and *E. coli* cocultured neutrophils. PC1 differences are shown at the *top* with *E. coli* ΔPC1 domains marked with green triangles. (*G*) Black and green points indicate genome-wide interaction correlations for each 40-kb bin in the genome, ranked from most to least differential, *left* to *right*. Bins containing *E. coli* ΔPC1 domains are marked with green points. The green line with green shading indicates the proportion of total *E. coli* ΔPC1 domains found at a given rank or lower, showing a preference for *E. coli* ΔPC1 domains to fall within genomic regions with the most differential chromatin interactions upon *E. coli* encounter. (*H*) Box plots of genome-wide interaction correlation values for *E. coli* ΔPC1 domains and the remainder of the genome during *E. coli* encounter. Box plot outliers are not shown. (*) Wilcoxon rank-sum test *P*-value $<2 \times 10^{-16}$.

exposed to microbial agents. We found that in unstimulated neutrophils the *CXCL* locus exists as a loop domain associated with a modestly positive PC1 score that is insulated from neighboring euchromatin (Fig. 4A). Notably, within 3 h of exposure to *E. coli*, the euchromatic character of the *CXCL* locus was significantly strengthened (Fig. 4A), accompanied by large scale changes in chromatin folding, with genomic interactions and transcriptional activation spreading into neighboring regions (Fig. 4B,C).

To determine whether the alterations in genome folding were associated with gene expression, activated neutrophils were analyzed for transcript abundance as well as CTCF and SMC3 occupancy (Fig. 4C). As expected, CXCL8, CXCL1, and CXCL2 transcript abundance was significantly elevated upon *E. coli* encounter (Fig. 4C). Notably, a recently described noncoding genomic region located immediately upstream of *CXCL8* was also transcriptionally induced upon exposure to bacteria (Fig. 4C; Fanucchi et al. 2019). While CTCF occupancy was elevated at a site closely linked with the *CXCL8* locus, other CTCF-bound sites in the locus were not modulated upon activation (Fig. 4C). In contrast, we found that *E. coli* encounter substantially enriched cohesin occupancy across the locus (Fig. 4C). Cohesin occupancy was particularly prominent at sites closely associated with de novo loops that linked the *CXCL8*, *CXCL1*, and *CXCL2* gene bodies, promoter regions, and SMC3-enriched intergenic regions into a shared transcriptional hub (Fig. 4C).

To validate these findings in single cells, we performed fluorescence in situ hybridization (FISH) using a probe

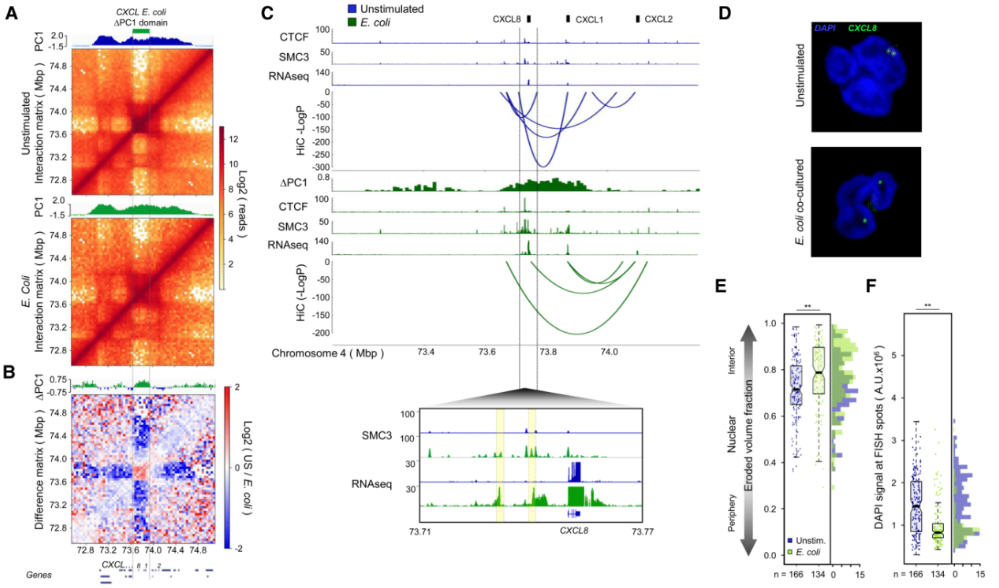


Figure 4. *E. coli* co-culture-induced topological changes at the *CXCL* sub-domain are associated with noncoding transcription, cohesin recruitment, and locus repositioning. (*A*) HiC contact maps of the extended *CXCL* gene cluster in unstimulated (*top*) and *E. coli* cocultured (*bottom*) human neutrophils. PC1 scores are shown *above* their respective matrices, *CXCL E. coli* ΔPC1 domain position is noted. (*B*) Log$_2$ difference matrix comparing HiC contacts between unstimulated and *E. coli* cocultured neutrophils within the extended *CXCL* gene cluster. PC1 differential values (*E. coli* cocultured–unstimulated PC1 values) are shown above matrix, protein-coding genes in the CXCL gene locus are shown *below*. (*C*, *top*) Linear genomic features and significant chromatin interactions at the *CXCL E. coli* ΔPC1 domain. CTCF and SMC3 ChIP-seq, RNA-seq, and HOMER-defined chromatin interactions with −log(*P*) values less than −50 are shown for unstimulated (blue, *top*) and *E. coli* cocultured neutrophils (green, *bottom*) with PC1 differential values (*E. coli* cocultured–unstimulated PC1 values) shown between. The *bottom* panel displays SMC3 ChIP-seq and RNA-seq tracks at the *CXCL8* (IL8) gene, demonstrating transcription-associated recruitment of SMC3 to the *CXCL8* locus. Peaks of transcription and associated SMC3 recruitment are highlighted in yellow. (*D*) Representative FISH image (z-projection) showing the euchromatic *CXCL8* locus (green) in unstimulated (*top*) and *E. coli* cocultured (*bottom*) neutrophils. (*E*) Quantification of the proportion of nuclear volume between the FISH signal and the nuclear periphery (Eroded volume fraction) in unstimulated and *E. coli* co-cultured neutrophils. Number of alleles analyzed are listed *below* the box plots. Wilcoxon rank sum test *P*-value for the data distributions: (**) $P < 0.00005$. (*F*) Quantification of DAPI signal intensity at the FISH spots identified in *B*. Wilcoxon rank sum test *P*-values: (**) $P < 0.00005$.

corresponding to the *E. coli*-specific *CXCL* ΔPC1 domain (Fig. 4D). In unstimulated neutrophils, the *CXCL E. coli* ΔPC1 domain localized near the nuclear periphery (Fig. 4D). Upon *E. coli* encounter the *CXCL E. coli* ΔPC1 domain rapidly relocated away from the heterochromatic nuclear periphery toward the nuclear interior, concomitant with its change in euchromatic character and elevated transcript levels (Fig. 4D). Specifically, the *E. coli* ΔPC1 domain relocated from the DAPI-dense portion of the nucleus near the nuclear periphery to the DAPI-sparse nuclear interior (Fig. 4E,F). This change in nuclear positioning was not an indirect result of changes in nuclear morphology, nor activation-induced loss of nuclear integrity, as heterochromatic control probes remained tightly associated with the nuclear periphery during *E. coli* encounter (Supplemental Fig. S3A). Collectively, these observations indicate that upon microbial exposure human neutrophils rapidly remodel nuclear architecture to assemble a *CXCL* transcriptional hub in the nuclear interior.

Neutrophil activation is associated with global loss of insulation at inflammatory genes

The data described above reveal that when human neutrophils encounter bacteria, a subset of inflammatory genes undergo large-scale changes in chromatin folding that spread into neighboring loop domains. To quantitatively describe this loss of subdomain insulation, we computed the insulation scores for genomic regions that surrounded the boundaries of *E. coli* ΔPC1 domains and upon microbial exposure gained euchromatic character to merge with surrounding euchromatin. We found that upon *E. coli* encounter the gain of euchromatic character across

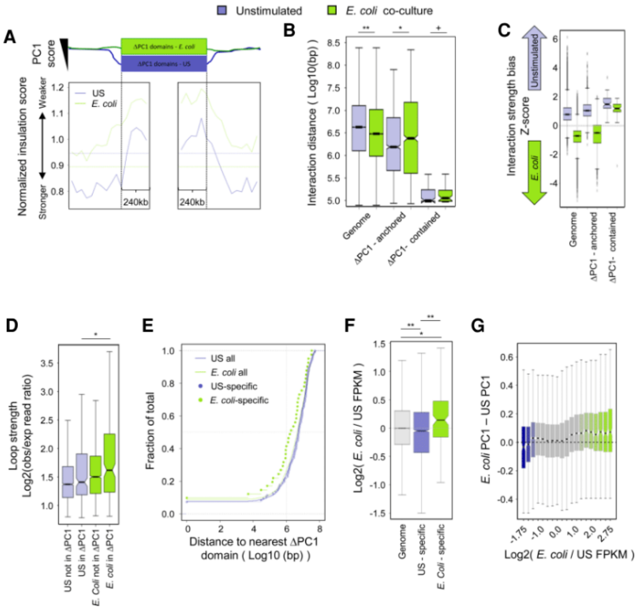


Figure 5. *E. coli* ΔPC1 domains lose local spatial insulation and are associated with increased cohesin-bound chromatin loop strength. (*A*) Insulation score meta-analysis. Insulation scores were calculated for genomic regions surrounding *E. coli* ΔPC1 domain boundaries. Normalized insulation scores for unstimulated (blue) and *E. coli* cocultured (green) neutrophils are shown. Vertical dotted lines demarcate meta-domain boundaries. Horizontal dotted lines show median normalized insulation scores at PC domain boundaries genome-wide for each condition. Calculations shown are for all *E. coli* ΔPC1 domains larger than 100 kb. (*B*) Distribution of linear genomic distances between chromatin interaction anchor points genome-wide, for interactions anchored in ΔPC1 domains, and for interactions fully contained within *E. coli* ΔPC1 domains. Wilcoxon rank sum test: (**) $P < 2.2 \times 10^{-16}$; (*) $P < 1 \times 10^{-5}$; (+) not significant. (*C*) Interaction strength bias (see the Materials and Methods) of HOMER-defined chromatin interactions for unstimulated and *E. coli* cocultured neutrophils genome-wide, for those interactions with a single anchor in an *E. coli* ΔPC1 domain, and for those interactions contained entirely within *E. coli* ΔPC1 domains. (*D*) HICCUPS defined loop strength with respect to *E. coli* ΔPC1 domains in unstimulated and *E. coli* cocultured neutrophils. (*E*) Distance between HICCUPS loops and *E. coli* ΔPC1 domains. Differences are not significant by the Kolmogorov-Smirnov test. Five out of 45 (11%) *E. coli* coculture-specific loops are in *E. coli* ΔPC1 domains, and 19/49 are within 1 Mb of *E. coli* ΔPC1 domains (45%). *E. coli* ΔPC1 domains make up ~0.3% of the human neutrophil genome. (*F*) Log_2 (normalized *E. coli* cocultured/unstimulated) FPKM values for genes at chromatin interaction anchors shared between unstimulated and *E. coli* cocultured neutrophils, or at chromatin interactions specific to one condition. Wicoxon rank sum test: (**) $P < 1 \times 10^{-4}$; (*) $P < 5{\times}10^{-3}$. (*G*) Change in PC1 values at genes with the given mRNA expression differential during *E. coli* encounter.

E. coli ΔPC1 domains was closely associated with decreased insulation strength at *E. coli* ΔPC1 domain boundaries (Fig. 5A). Although globally the genomic distances separating chromatin interaction anchor points were significantly decreased in activated versus unstimulated neutrophils, the distance separating anchor points of chromatin interactions with *E. coli* ΔPC1 domains increased (Fig. 5B). This loss of insulation at *E. coli* ΔPC1 domain boundaries suggests de novo formation of regulatory interactions with the surrounding area (Fig. 5B). Additionally, chromatin interactions contained entirely within *E. coli* ΔPC1 domains were found on average to be significantly stronger in unstimulated neutrophils as compared with *E. coli* cocultured neutrophils, suggesting a loss of subdomain structure and self-association during microbial encounter (Fig. 5C).

Given the loss of insulation at *E. coli* ΔPC1 domain boundaries, we next sought to determine the relationship between ΔPC1 domains and gene regulatory chromatin interactions. Although chromatin interactions within *E. coli* ΔPC1 domains were on average weakened during *E. coli* encounter (Fig. 5C), *E. coli* encounter-specific chromatin loops within ΔPC1 domains were significantly stronger than chromatin loops found only in unstimulated neutrophils (Fig. 5D). This finding suggested a gene regulatory role for *E. coli* encounter-dependent loops, and a tight link between these loops and *E. coli* ΔPC1 domains. Supporting this finding, *E. coli*-dependent chromatin loops were generally closer to *E. coli* ΔPC1 domains than were unstimulated neutrophil-specific chromatin loops, and 11% of *E. coli*-dependent chromatin loops were identified in *E. coli* ΔPC1 domains, which make up only 0.3% of the genome (Fig. 5E). Importantly, genes near *E. coli* coculture-specific chromatin loop anchors were significantly more highly expressed than genes at chromatin loop anchors found only in unstimulated neutrophils (Fig. 5F).

Given the enrichment of neutrophil inflammatory response genes in *E. coli* ΔPC1 domains (Fig. 3D) and the link between expression levels and an increase in euchromatic character (Supplemental Fig. S2C), we next determined the relationship between euchromatic character (PC1 score) and transcript levels during microbe encounter. Notably, we found a strong correlation between PC1 score dynamics and transcriptional dynamics, with the most highly induced genes also showing the largest increases in PC1 score, and the most repressed genes showing the largest decreases in PC1 score (Fig. 5G).

These phenomena are readily visible at a number of inflammatory loci, wherein the tight self-association of *E. coli* ΔPC1 domains in unstimulated cells is lost in favor of distal regulatory interactions and transcriptional activation during *E. coli* encounter (Supplemental Fig. S4).

Taken together these data indicate that neutrophil transcriptional state, euchromatic character, and spatial localization of genes are closely linked.

Microbial exposure rapidly recruits cohesin to inflammatory enhancers

Examination of gene regulatory interactions associated with *E. coli* ΔPC1 domains (Fig. 4; Supplemental Fig. S4) hinted that a large number of *E. coli*-dependent interactions were associated with the recruitment of the cohesin complex to *cis* regulatory elements. To study this phenomenon and understand its role in *E. coli*-dependent changes in gene expression, we analyzed unstimulated and *E. coli*-exposed neutrophils for SMC3 and CTCF occupancy, as well as changes in H3K27Ac-marked enhancer repertoires and transcription. We then focused our analysis on a specific subset of SMC3-amassed enhancers: those H3K27ac-defined enhancers present in *E. coli* cocultured neutrophils that gained substantial SMC3 occupancy during *E. coli* encounter (Fig. 6A; Materials and Methods).

SMC3-amassed enhancers are characterized by modestly increased CTCF binding, and substantially increased H3K27ac deposition and polyadenylated RNA abundance (Fig. 6A). Supporting the importance of SMC3 and H3K27ac deposition at new regulatory interactions, *E. coli* coculture-specific chromatin interaction anchors were found to be enriched for SMC3 occupancy and H3K27ac deposition, with only modest changes in CTCF occupancy compared with unstimulated neutrophil-specific interactions (Supplemental Fig. S5A–C). Similarly, H3K27ac-defined enhancers enriched for polyadenylated RNA signal were likewise enriched for SMC3 occupancy and H3K27ac deposition (Supplemental Fig. S5D).

Supporting the importance of SMC3-amassed enhancers in *E. coli* ΔPC1 domain behavior, SMC3-amassed enhancers were enriched at *E. coli* ΔPC1 domains (Fig. 6B). SMC-amassed enhancers were also, on average, localized closer to ΔPC1 domains when compared with the global enhancer repertoire, and were more enriched in ΔPC1 domains than the entire enhancer repertoire (Fig. 6C).

To understand the mechanism of cohesin targeting to SMC3-amassed enhancers, we identified transcription factor binding motifs enriched within both SMC3-amassed enhancers and non-SMC3-amassed enhancers found in *E. coli* cocultured neutrophils. We then computed the enrichment of transcription factor motif density in SMC3-amassed enhancers compared with all enhancers found in *E. coli* cocultured neutrophils. Validating this approach, we found that DNA-binding motifs associated with known inflammatory regulating factor AP1 including JUND and FOSL2, as well as transcriptional regulators that orchestrate neutrophil differentiation and physiology including CEBPB, CEBP homolog CHOP, and PU.1, were significantly enriched across the bacterial-induced SMC3-amassed enhancer repertoire compared with the entire enhancer repertoire (Fig. 6D [top], E). Notably, transcript abundance associated with these factors was elevated in neutrophils exposed to *E. coli* (Fig. 6D, bottom). Apart from known inflammatory and myeloid regulatory transcription factors, we also found that DNA sequences associated with TFE3 occupancy were enriched at SMC3-amassed enhancers (Fig. 6D,E).

Recent studies revealed that in activated macrophages and microglial cells TFE3 orchestrate an inflammatory program of gene expression (Pastore et al. 2016). Notably, upon activation TFE3 relocated from the cytoplasm to the nuclear interior (Pastore et al. 2016). As a first approach to determine whether likewise TFE3 translocates from the cytoplasm to the nucleus we examined naïve and microbial-exposed neutrophils for TFE3 localization using immunofluorescence. We found that in naive neutrophils TFE3 was predominantly localized in the cytoplasm. Notably, however, we found that within 3 h upon exposure to *E. coli* TFE3 repositioned from the cytoplasm to the nuclear interior in the vast majority of neutrophils (Supplemental Fig. S6).

Genes interacting with SMC3-amassed enhancers were next analyzed for functional group enrichment. Notably, bacterial-induced SMC3-amassed enhancers were closely associated with genes involved in neutrophil activation, including cytokine signaling and response, chemotaxis, and degranulation (Fig. 6F; Supplemental Fig. S7; Supplemental Table S2). Analysis of RNA-seq data revealed that genes interacting with enhancers in general showed little preference to be induced upon *E. coli* encounter as compared with any other gene in the genome. In contrast, genes linked to SMC3-amassed enhancers showed a significant increase in gene expression during *E. coli* encounter compared with genes globally, or to genes linked to enhancers in general (Fig. 6G). This phenomenon appears to depend on both SMC3 occupancy and H3K27ac deposition, as interactions with either SMC3-amassed *E. coli*-specific enhancers or SMC3-amassed pre-existing enhancers were both associated with increased gene expression, whereas interactions with enhancers only found in unstimulated cells were not associated with increased gene expression, regardless of SMC3 occupancy (Supplemental Fig. S8).

Taken together, these data indicate that upon bacterial exposure human neutrophils rapidly sequester the cohesin machinery at a specific subset of enhancers to modulate chromatin folding and activate an inflammatory gene transcription program.

Discussion

The unique morphology of neutrophils has been an enigma since its discovery more than a century ago (Cavaillon 2011). How neutrophil genomes are folded into three-dimensional space and how neutrophil nuclear architecture is altered upon microbial exposure has remained largely unknown. Here we used a genome-wide

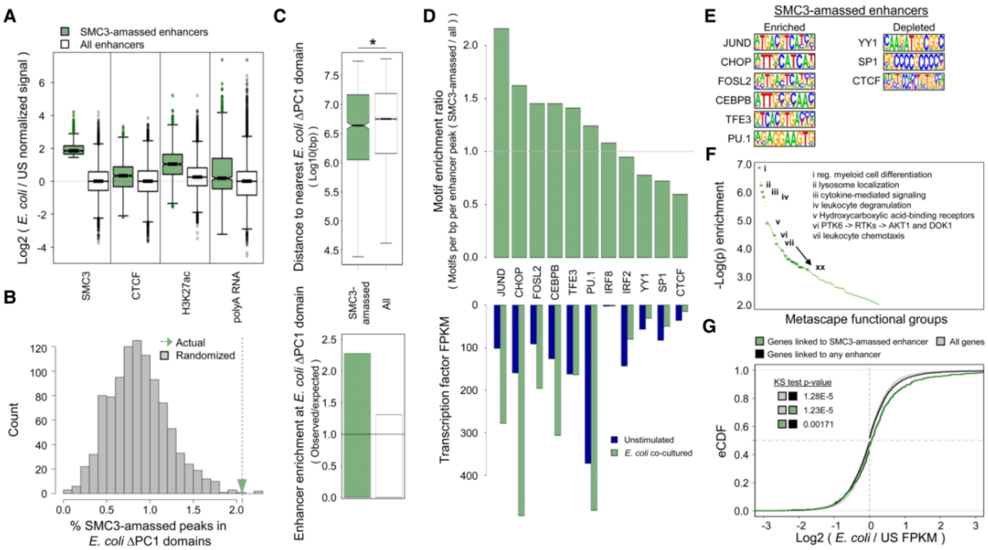


Figure 6. *E. coli* coculture induces cohesin recruitment to a subset of H3K27ac-defined enhancers. (*A*) Log_2 ratio (*E. coli* cocultured/unstimulated) of the normalized ChIP-seq and RNA-seq signals at H3K27ac-defined enhancers that amass SMC3 as well as all enhancers genome-wide. (*B*) Histogram showing the percent of SMC3-amassed enhancers falling in *E. coli* ΔPC1 domains randomly positioned within the A compartment (gray), and the percent of SMC3-amassed enhancers falling within actual *E. coli* ΔPC1 domains (green arrow). Out of 1000 random permutations, 997 resulted in lower overlap between SMC3-amassed enhancers and ΔPC1 domains than were observed in the empirical data. (*C*) The *top* panel shows distance distribution between *E. coli* ΔPC1 domain boundaries and SMC3-amassed enhancers (green) and all enhancers (white). The *bottom* panel shows observed enrichment of SMC3-amassed enhancers or all enhancers in ΔPC1 domains divided by the expected enrichment of these enhancers in *E. coli* ΔPC1 domains based on 1000 random permutations of *E. coli* ΔPC1 domain positions within the A compartment. (*) Wilcoxon rank sum test P-value < 0.005. (*D*) The *top* panel indicates ratio of mean transcription factor motif density (motifs per base pair per peak, SMC3-amassed enhancers/all enhancers) for representative transcription factors. The *bottom* panel shows gene expression values (FPKM) of representative transcription factors in unstimulated and *E. coli* cocultured neutrophils. (*E*) Known transcription factor motifs identified in *D*. (*F*) Metascape gene functional analysis for genes interacting with SMC3-amassed enhancers. Full metascape analysis results are shown in Supplemental Table S2. (*G*) Empirical cumulative distribution of $\log_2$(*E. coli* cocultured/unstimulated FPKM) values for all genes, genes interacting with any enhancer, and genes interacting with *E. coli* coculture-dependent SMC3-amassed enhancers.

chromosome conformation capture approach (HiC) to address these questions. We found that human neutrophil nuclei, when compared with embryonic stem cells, displayed a distinct nuclear architecture: (1) a decline in genomic interactions across loop domains (<3 Mb); (2) a segmentation of large, continuous A and B compartments into numerous small compartments, resulting in the establishment of new compartment and loop domain boundaries; and, (3) an increase in remote chromosomal interactions across loop domains (>3 Mb). This increase in long-range genomic interactions primarily involved heterochromatic regions indicating a key role for heterochromatic interactions in influencing human neutrophil genome topology. Our data are consistent with previous studies involving murine neutrophils that also displayed a highly contracted genome when compared with progenitor cells and show that key features of neutrophil genome structure are conserved between the murine and human genomes (Zhu et al. 2017).

The neutrophil genome undergoes large-scale alterations in morphology upon bacterial encounter. Using genome-wide chromosome conformation capture studies, we found that such changes involve the repositioning of euchromatic *E. coli* ΔPC1 domains enriched for cytokine and other immune response genes. Upon encountering activating stimuli, these domains gained euchromatic character, repositioning themselves from the nuclear periphery to the more euchromatic nuclear interior. During this process, the boundaries of these domains lost insulation, allowing the domain to merge with neighboring highly euchromatic regions, and further allowing for new chromatin interactions to form and activate an inflammatory gene program. These subdomains resemble a previously identified euchromatic A2 spatial subcompartment positioned between the nuclear periphery and the nuclear interior (Rao et al. 2014; Chen et al. 2018). Based on our observations, we propose that the A2 subcompartment is associated with genes or regulatory elements that need to be transcriptionally repressed, but accessed quickly, precluding both their sequestration to the fully heterochromatic B compartment, as well as their presence in the transcriptionally active A1 compartment.

Our data further provide mechanistic insight as to how neutrophils instruct changes in nuclear positioning and domain insulation upon bacterial encounter. Alterations in chromatin topology both at ΔPC1 domains and across the genome are closely associated with the rapid recruitment of cohesin to a subset of H3K27ac-defined enhancers. While cohesin occupancy is substantially enriched at these enhancers, CTCF binding is only modestly elevated upon bacterial encounter. These observations imply that changes in nuclear architecture are predominantly activated by cohesin-dependent loop extrusion. This finding then raises the question as to how cohesin is being recruited to inflammatory genes upon bacterial encounter. We found that the increase in cohesin occupancy at SMC3-amassed enhancers was closely accompanied by substantial enrichment for the enhancer mark H3K27ac. Hence, we suggest that upon bacterial encounter, human neutrophils activate a signaling response that involves the Toll-like receptor pathway. Motif analysis suggests that Toll-like receptor mediated signaling modulates the expression and/or biochemical activities of key neutrophil-associated transcriptional regulators such as PU.1, CEBP/β, CEBP homolog CHOP, AP1 factors JUN and FOS, as well as TFE3. The activities of such regulators, in turn, would promote the assembly of an active enhancer repertoire as evidenced by the deposition of H3K27Ac, which then rapidly sequesters cohesin at inflammatory response enhancer-gene promoter clusters. Once recruited to SMC3-amassed enhancers, cohesin may act to extrude chromatin until convergent CTCF sites are reached, removing insulation at ΔPC1 domain boundaries by forming de novo loop domains in which activated enhancers are placed within close spatial proximity of gene promoters, altogether facilitating the rapid activation of an inflammatory response gene program (Fig. 7).

Why has such an elaborate mechanism of gene activation, including loop extrusion, evolved in human neutrophils? We suggest that segregating enhancers and promoters in spatially distinct loop domains ensures efficient silencing and prevents stochastic activation of an inflammatory-specific gene program in unstimulated neutrophils. Only upon exposure to activating stimuli are unstimulated neutrophils instructed to juxtapose the inflammatory enhancer repertoire with their target gene promoters, thus facilitating enhancer-promoter communication and the induction of an inflammatory-specific gene program. We hypothesize that the specificity of this response is likely governed by transcription factors downstream from activated receptors that bind their target enhancers, allowing cohesin and histone acetyl transferase recruitment, juxtaposition of target gene promoters, and stabilization of transcription units.

As documented here for human neutrophils during a microbial encounter, enhancers and promoters may be spatially segregated from each other in distinct loop domains until an appropriate environmental signal is received in order to prevent inappropriate or pathological activation of gene expression. Previous studies have documented a related mechanism that orchestrates the developmental progression of lymphoid cells. Specifically, regulatory regions associated with key developmental regulators such as EBF1 and Bcl11b are, in progenitor cells, positioned at the nuclear lamina to suppress premature activation during developmental progression. Upon reaching the appropriate developmental stage, alterations in chromatin folding readily reposition such enhancers away from the transcriptionally repressive environment at the lamina into the euchromatic nuclear interior, leading to assembly of transcriptionally productive enhancer–promoter interactions. The repositioning also directs the enhancer into a single loop domain to facilitate enhancer–promoter communication. Once placed within the euchromatic compartment and within spatial proximity to EBF1 and Bcl11b, enhancers and promoters then act to establish B or T cell identity, respectively (Lin et al. 2012; Isoda et al. 2017). Thus, the inflammatory gene response and activation of a developmental-specific gene expression programs share a common

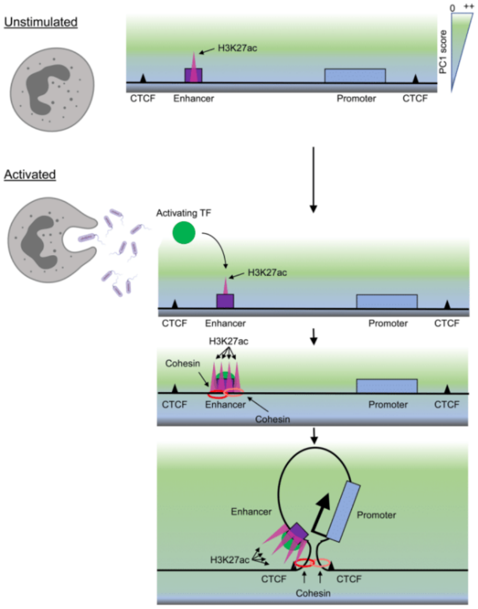


Figure 7. Microbial-induced human neutrophil activation instructs rapid changes in nuclear architecture to orchestrate an inflammatory gene program. Activation-induced transcription factor binding results in H3K27ac deposition, cohesin recruitment, and formation of de novo chromatin loops linking enhancers to inflammatory genes to orchestrate an inflammatory gene program.

mechanism that assures appropriate timing of gene expression.

In sum, here we demonstrate that in human neutrophils, prior to encounter with bacteria, an armamentarium of inflammatory genes was positioned in a transcriptionally passive environment suppressing premature transcriptional activation. Upon microbial exposure, however, human neutrophils rapidly (<3 h) repositioned the ensemble of proinflammatory genes towards the transcriptionally permissive compartment. We found that the repositioning of genes was closely associated with the swift recruitment of cohesin across the inflammatory enhancer landscape, permitting an immediate transcriptional response upon bacterial exposure. These data reveal at the mechanistic level how upon microbial challenge human neutrophils undergo rapid changes in nuclear architecture to orchestrate an immediate inflammatory gene program.

Materials and methods

Human subject details

Blood for neutrophil isolation was obtained via venopuncture from healthy human volunteers under written informed consent approved by the University of California at San Diego Human Research Protection Program (#131002X).

Blood draws and neutrophil isolation

Whole blood was layered onto Polymorphprep reagent (Accurate Chemical and Scientific Corp. AN1114683), centrifuged for 45 min at 500*g*, 25°C, and allowed to stop without braking. The granulocyte layer was extracted and contaminating red blood cells were lysed as needed (generally one to three times) with brief resuspensions in sterile H_2O followed by immediate flooding with 1× phosphate-buffered saline (PBS) and centrifugation at 500*g* for 7 min at 25°C. Cells were checked for purity via Wright-Giemsa staining; the final granulocyte fraction was generally >95% neutrophils. For RNA-sequencing experiments, neutrophils were further purified to homogeneity using an EasySep human neutrophil enrichment kit (Stem Cell Technologies 19257) as per the manufacturer's protocol.

For all experiments, neutrophils were cultured in HBSS +Ca/+Mg/−Phenol red (Thermo Fisher 14025092) with the addition of 0.5% endotoxin-free BSA (Akron AK8917-0100) at 37°C in a 5% CO_2 humidified incubator.

Wright-Giemsa staining

Neutrophils (1×10^5) were spun onto cover slips using a Cytospin3 (Shandon 74010121 GB) and flooded with Wright stain (Sigma WS16-500ML) for 3 min. Cover slips were then washed with six consecutive dips in water baths. Cover slips were then allowed to air dry and were then flooded with Giemsa stain (Sigma GS500-500ML) and allowed to incubate for 7 min before being washed as above and allowed to air dry.

Neutrophil activation

Neutrophils were plated at the desired cell numbers and treated with 25 nM phorbol 12-myristate 13-acetate (PMA, Promega V1171) or cocultured in the presence of *E. coli* strain K1 at a multiplicity of infection (MOI) of 5. Stimulations were performed for 3 h and cells were harvested as detailed below.

Chromatin immunoprecipitation with sequencing (ChIP-seq)

Neutrophils were plated at 10×10^6 to 20×10^6 cells/10 mL for each ChIP experiment. At the completion of each experiment cells were washed with fresh media, formaldehyde was added to the culture to a final concentration of 1%, and cells were cross linked with agitation for 10 min at room temperature. Fixation was then quenched for 5 min with glycine at a final concentration of 0.13 M. Fixed cells were scraped from the plate and washed three times in ice cold 1× phosphate buffered saline (PBS) with 0.1 mM EDTA and 1×EDTA-free complete protease inhibitors (Roche 05056489001). Cell pellets were snap frozen in liquid nitrogen and stored at −80°C until processing.

To bind antibody to Protein G Dynabeads (Invitrogen 10004D), beads were washed three times with 1 mL of bead wash buffer (1×PBS, 5 mg/mL BSA, Roche complete EDTA-free protease inhibitor, 0.22 µM filtered) and resuspended in 500 µL of the same. One to five micrograms of antibody was added and allowed to bind beads overnight at 4°C with rotation. The following day, beads were washed three times with 1 mL of bead wash buffer and resuspended in 100 µL of RIPA 150 (50 mM Tris at pH 8.0, 150 mM NaCl, 0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100, 1 mM EDTA).

For each ChIP, cells were thawed and lysed for 10 min on ice with inversion in Farnham lysis buffer (5 mM PIPES at pH 8.0, 85 mM KCl, 0.5% NP-40, 10 mM EDTA, protease inhibitors) with or without 20 draws through an 18-gauge needle. Nuclei were spun down at 2000 rpm for 5 min at 10°C in a benchtop microfuge, supernatant was discarded, and nuclei were resuspended in 300 µL of RIPA 150. Chromatin was then sonicated in a Diagenode Bioruptor 300 chilled to 4°C three times for eight cycles of 30 sec on and 30 sec off, set on high with 5 min of cooling time between each set of eight cycles. The insoluble fraction was spun down at maximum speed for 20 min at 4°C in a benchtop microfuge. Input and IP samples were split to separate new tubes, IP volume was adjusted to 900 µL with RIPA 150, and 100 µL of Protein G dynabeads bound to the antibody of interest in RIPA 150 was added to each IP. Chromatin was allowed to bind to antibody-bead conjugates overnight at 4°C while rotating. Following binding, beads were washed twice for 5 min in RIPA 150, twice for 5 min in RIPA 500 (50 mM Tris at pH 8.0, 500 mM NaCl, 0.1% SDS, 01% sodium deoxycholate, 1% Triton X-100, 1 mM EDTA), twice for 3 min in LiCl wash (10 mM Tris at pH 8.0, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA), and once in 1× TE. Beads were transferred to clean tubes at the start of each new wash buffer. DNA was eluted from beads with 200 µL of elution buffer (1 mM sodium carbonate, 1% SDS) for 1 h at 65°C with shaking, at which point beads were removed and cross-links were reversed overnight at 65°C. Eluted DNA was purified using a ChIP DNA clean and concentrator kit (Zymo D5205).

DNA for ChIP and other high-throughput sequencing approaches was processed as follows: End repair was performed using an Epicenter End-It kit (Lucigen ER0720), according to manufacturer's instructions and column purified in a Zymo Minelute column (Zymo D4013). A-tails were added by incubating DNA in 1×NEB buffer 2 (New England Biolabs B7002S) with the addition of 200 µM dATP and 7.5 units of Klenow (exo-) (New England Biolabs M0212L) for 45 min at 37°C. NEB Next adaptors (New England Biolabs E7337A) were ligated using an NEB quick ligation kit (New England Biolabs M2200L) for 30 min at benchtop temperature followed by treatment with 2 µL of USER enzyme (New England Biolabs M5505L) for 15 min at 37°C. DNA was purified using an AmpureXP bead-analogous two-step SPRI bead protocol (Rohland and Reich 2012), resulting in purification of DNA fragments between ~200 and 800 bp.

PCR amplification of final libraries for sequencing was performed with Phusion hot start polymerase II system (Thermo Fisher F549L) in conjunction with the NEB Next indexing system (New England Biolabs E7335L and E7500S). Final size selection for all high-throughput sequencing libraries was performed using a home-made two-step SPRI bead-based DNA purification system, resulting in final DNA fragment sizes of ∼200–800 bp.

RNA sequencing

At specified time points neutrophils were washed once with PBS and lysed in the RLT buffer component of the Qiagen RNeasy mini kit (Qiagen 74106) with the addition of 10 µL/mL 2-mercaptoethanol, homogenized via Qiashredder (Qiagen 79654), and snap-frozen in liquid nitrogen. Total RNA was purified via RNeasy mini kit (Qiagen 74106) according to the manufacturer's instructions, including the RNase-free DNase (Qiagen 79254) treatment step. RNA was eluted in H_2O, Turbo DNase kit buffer (Thermo Fisher/Ambion AM1907) was added to a 1× concentration, and RNA was treated with 4 U of Turbo DNase for 30 min at 37°C. Turbo DNase was then treated with inactivation reagent per manufacturer's specifications. mRNA was purified from total RNA using a Dynabead mRNA purification kit (Life Technologies 61006). First strand synthesis was performed using the SuperScript III first strand synthesis system (Thermo Fisher 18080051) as follows: 100–500 ng of RNA, 0.5 µL of oligo(dT) primer, 0.8 µL of random hexamer, 1 µL of 10 mM dNTP, and H_2O to 9.5 µL. The mixture was incubated for 10 min at 70°C and then snap frozen. First strand synthesis mix composed of 2 µL of 10× RT buffer, 4 µL of 25 mM $MgCl_2$, 2 µL of 0.1 M DTT, 0.5 µL of 120 ng/µL ActinomycinD, 40 U of RNaseOUT, and 200 units of SuperScriptIII was added to the mixture, which was then incubated for 10 min at 25°C, 45 min at 42°C, 25 min at 50°C, and15 min at 75°C. Unincorporated nucleotides were removed from the mixture using a ProbeQuant G-50 column (Sigma GE28-9034-08). First strand synthesis reaction was then brought to 51 µL with H_2O and cooled on ice. Twenty-four microliters of second strand mixture composed of 1 µL of 10× RT buffer; 2 µL of 25 mM $MgCl_2$; 1 µL of 0.1 M DTT; 2 µL of 10 mM dATP, dGTP, dCTP, and dUTP mix; 15 µL of 5× second strand synthesis buffer (New England Biolabs B6117S); 0.5 µL *E. coli* ligase (New England Biolabs M0205S); 2 µL of DNA polymerase I (New England Biolabs M0209S); and 0.5 µL of RNase H was added and the mixture was incubated for 2 h at 16°C. DNA was purified using a DNA clean and concentrator kit (Zymo D4013) and sonicated on a Covaris E220 with the following settings; duty cycle 10%; intensity 5; cycle per burst 200; time (seconds) 180. Sonicated DNA was purified using a DNA clean and concentrator kit. DNA was prepared for high-throughput sequencing using the methodology described above for ChIP-seq, with the addition of 1 µL of UNG (Thermo Fisher/Applied Biosystems N8080096) during USER enzyme treatment.

Whole-genome bisulfite sequencing

Neutrophils were washed twice with PBS and genomic DNA was isolated using a DNeasy Blood and Tissue kit (Qiagen 69504). One microgram of genomic DNA mixed with unmethylated λ DNA at a concentration of 0.5% of total DNA was sonicated by Biorupter 300 with 20 cycles (30 sec on/30 sec off at low power). Fragmented DNA was end-repaired and A-tailed as described above. TruSeq adapters (Illumina FC-121-2001) were ligated to fragmented DNA, which was then purified by running on a 2% agarose gel. Bisulfite conversion was performed using the MethylCode kit as described by the manufacturer (Invitrogen MECOV-50). Bisulfite-treated DNA was amplified by using a TruSeq PCR primer mixture and Pfu Turbo Cx Polymerase, agarose gel-purified, and sequenced on an Illumina HiSeq 2500 sequencer with paired-end 150-bp reads.

E. coli *culture and MOI determination*

E. coli strain K1 was grown in LB at 37°C with shaking overnight, and diluted into a fresh culture and grown to exponential phase the day of each experiment. *E. coli* was then pelleted at 3000 rpm for 10 min at 10°C on a benchtop centrifuge, washed in cell culture medium, and added to neutrophil cultures at an MOI of ∼5 in HBSS +Ca/+Mg/−Phenol red with 0.5% endotoxin free BSA. Nine 1:10 serial dilutions of *E. coli*-containing media were plated on LB agar and grown overnight at 37°C. The resulting colonies were counted in order to assess MOI for individual experiments.

In situ HiC

In situ HiC was performed as described (Rao et al. 2014), modifying only the MboI restriction enzyme digest time to assure proper digestion of chromatin. Generally, HiC libraries prepared from activated neutrophils were digested for 2–4 h with 50 U of MboI to avoid overdigesting the chromatin. The remainder of the library preparation adhered to the published protocol and reagents exactly. HiC library DNA was prepared for high throughput sequencing using the NEB Next platform according to manufacturer's instructions, and sequenced using paired-end 100-bp reads.

Fluorescence in situ hybridization (FISH)

Cover slips were incubated overnight in 1%HCl in 70% ethanol, washed 3× with H_2O, once in 70% ethanol, and stored in 100% ethanol. Coverslips were allowed to air dry prior to adding cells. Cells were incubated on cover slips in 24-well plates as described above. At the completion of incubation times, cells were washed three times for 3 min in PBS and fixed for 30 min in 6% paraformaldehyde (Electron Microscopy Sciences 15710) in 1× PBS. PFA was flushed out with >5 volumes of PBS/0.05% Tween-20 (PBST), ensuring that cells never contact the air. Residual PFA was quenched via incubation with fresh 20 mM glycine in PBS for 15 min at room temperature. Cells were permeabilized in PBS + 0.5% Triton X-100 for 20 min at room temperature, washed twice with PBST, and incubated in PBS + 100 µg/mL RNase A (Qiagen 19101) for 1 h at 37°C. Cells were then treated with 0.1 N HCl for 5 min at room temperature, washed twice for 3 min with 1× PBS and twice for 5 min with 2× SSC, and then incubated for >48 h in 2× SSC/50% formamide at 4°C. Coverslips were then blotted dry and 5 µL of probe containing 75–200 ng of labeled DNA was added to each coverslip. Cover slips were then sealed on top of glass slides along with probe using rubber cement. Probes and genomic DNA were denatured together for 5 min at 78°C on a heat block and allowed to hybridize for 16–48 h at 37°C. Following hybridization cover slips were washed 1 × 15 min in SSC/50% formamide prewarmed to 37°C, three times for 15 min in 2× SSC prewarmed to 37°C, three times for 7 min in 0.1× SSC prewarmed to 60°C, three times for 7 min in 4× SSC/0.02% Tween-20 prewarmed to 42°C, once for 5 min with 2× SSC prewarmed to 37°C, and twice for 5 min in 1× PBS. Cells were then postfixed in 4% PFA in 1× PBS for 10 min at room temperature, and PFA was flushed out as above. Cells were washed once for 10 min in PBST + DAPI, four times for 5 min in 1× PBS, and mounted in Prolong Gold mounting medium (Thermo Fisher P36930).

FISH probes were prepared from bacterial artificial chromosomes (BACs) using nick/translation (Roche 11745808910). One microgram of BAC DNA was used in each 20-µL nick/translation reaction along with the following fluorophores, as needed: ChromaTide Alexa fluor 488-5-dUTP (Thermo Fisher/Life Technologies C11397), Cy3-dUTP (VRW 42501), or Alexa fluor 647-aha-dUTP (Thermo Fisher/Life Technologies A32763). Nick/Translation was performed for 5–16 h at 15°C and terminated by addition of 1 µL of 0.5 M EDTA. Unincorporated nucleotides were removed with ProbeQuant G-50 columns per manufacturer's instructions. One-hundred nanograms of labeled probe DNA was run on a 1.5% agarose gel following each nick/translation reaction to ensure that the majority of probe fragments were in the 300- to 800-bp range. Up to 200 ng of total probe per cover slip was combined with 10 µg of salmon sperm DNA (Thermo Fisher 15632011), 4 µg of human Cot1 DNA (Thermo Fisher 15279011), 1/10 volume of 3 M sodium acetate (pH 5.2), and 2.5 volumes of 100% ethanol. Probes were allowed to precipitate for 30 min at −20°C, centrifuged at maximum speed for 20 min at 4°C, washed twice with 70% ethanol and once with 100% ethanol, air-dried, and resuspended in 6 µL of 100% formamide at 56°C. Six microliters of 2× hybridization buffer (40% dextran sulfate in 8× SSC [20× SSC: 3 M NaCl, 0.3 M sodium citrate]) was then added to each probe. Probes were denatured for 5 min at 80°C and snap-cooled on ice. Probes were then added to cover slips and denatured and hybridized to genomic DNA as noted above. The *CXCL* locus FISH probe utilized BAC RP11-243E9, the heterochromatic control probe used BAC RP11-134J16.

Imaging of FISH samples was performed at the Waitt Biophotonics Center at the Salk Institute. FISH samples were imaged on Zeiss Airyscan 880 microscopes using the Airyscan Fast mode (Huff 2016) at a resolution of 40 nm in the *x* and *y* axes. Z sections were imaged every 160 nm. Quantification of FISH data was performed using TANGO (Ollion et al. 2013) for FIJI (Schindelin et al. 2012). Nuclei and spot detection were performed with built-in tools in TANGO. Image metrics analyzed in TANGO include "eroded volume fraction" and "signal quantification layer" in Figure 4, and "distances" in Supplemental Figure S3. Metrics were exported from TANGO as text files and statistical analysis and figure generation were performed in R using built-in tools (https://www.R-project.org).

HiC analysis

Raw HiC library read alignment to human genome build hg38, valid read pair filtering, matrix assembly at various resolutions, and ICE normalization of said matrices were performed using HiC-pro with default settings (Servant et al. 2015). Biological replicates were pooled following valid read pair filtering, and pooled data sets were used for analysis except where noted.

For all direct comparisons of HiC data (topological domain boundary location comparisons, insulation scores, plotted contact matrices, $\log_2$ differential matrices) ICE normalized sparse matrix files were created containing only the subset of interacting bins that recorded reads in all data sets being compared. Read numbers at these bins were then quantile normalized in R using the normalize.quantiles() function in the preprocessCore package (https://github.com/bmbolstad/preprocessCore), allowing direct comparison of chromatin interactions between libraries with different read distributions and sequencing depths (Hsu et al. 2017).

Topological domain boundaries were called on normalized HiC data at 40-kb resolution using the domain calling software published in Dixon et al. (2012).

HiC-Pro defined valid read pairs were used in conjunction with HOMER (Heinz et al. 2010) to run principal component analysis (PCA, runHiCpca.pl -res 10000), generate distance versus interaction frequency plots (makeTagDirectory), define compartment boundaries (findHiCCompartments.pl), determine interaction correlations (getHiCcorrDiff.pl -res 40000 -superRes 40000), define distance-normalized chromatin interactions (analyzeHiC -res 20000 -superRes 40000 -minDist 100000), and to generate whole-chromosome pairing plots (analyzeHiC -res 400000000). CTCF anchored-type loops were called using HICCUPS (Rao et al. 2014).

Insulation scores were determined as follows: The genome was divided into 40-kb segments. Insulation scores for each segment were defined as the number of normalized (ICE and quantile) (see above) valid read pairs within a 500-kb window centered on the segment of interest whose ends map to opposite sides of the segment of interest divided by the total number of valid read pairs whose ends both map within the 500-kb window.

ΔPC1 domains were identified as follows: PCA was run at 10,000-bp resolution on pooled HiC data using the runHiCpca.pl command in HOMER with the following settings: -res 10000 -superRes 10000 -genome hg38. Visual inspection showed that positive PC1 values corresponded to the gene-rich A compartment, and negative PC1 values corresponded to the gene-poor B compartment on all chromosomes and across all conditions. Genomic regions with PC1 score differentials between conditions greater than three standard deviations above the mean PC1 score differential between conditions were identified as potential ΔPC1 domains. PCA was then run on individual HiC biological replicates and only those potential ΔPC1 domains with a reproducible gain in PC1 value in each biological replicate were retained. Finally, reproducible ΔPC1 domains within 100 kb of each other were merged into single continuous ΔPC1 domains which were used for downstream analysis.

ChIP-seq analysis

Raw fastq files were aligned to the human genome build hg38 using Bowtie (Langmead et al. 2009) with the following parameters: -m1 –best –strata. Downstream processing of ChIP-seq data was performed using HOMER, except where noted. Uniquely mapped reads from high quality biological replicates were pooled for downstream analysis (Landt et al. 2012). Sequencing data was reorganized as a HOMER-formatted tag directory for each replicate and multiple reads mapping to the same base pair were collapsed to a single read using the makeTagDirectory command in HOMER with the following parameters: *-tbp 1*. ChIP peaks were called using the findPeaks command in HOMER with default parameters. Genes at ChIP peaks were identified using annotatePeaks.pl in HOMER, and the GenomicRanges package (Lawrence et al. 2013) in R.

SMC3-amassed enhancers were defined as follows: Enhancers were defined as H3K27ac peaks called as above. In order to identify enhancers with activation-dependent cohesin recruitment (SMC3-amassed enhancers), total unique SMC3 ChIP-seq reads mapping to enhancers were calculated using annotatePeaks.pl in HOMER. To directly compare binding strength between conditions, read numbers at enhancers were quantile normalized across conditions using the preprocessCore R package. Those reads with a $\log_2$(normalized activated/normalized unstimulated read numbers) value >1.5 were defined as SMC3-amassed. The GenomicRanges package in R was used to identify genes in contact with SMC3-amassed enhancers in conjunction with HOMER-defined chromatin interactions (detailed below). Enhancer-gene pairs were called as interacting if the center of one interaction anchor was within 10 kb of an enhancer and the center of the other interaction anchor was within 50 kb of a gene promoter.

RNA-seq analysis

RNA-seq data was analyzed using the Tuxedo tools, except where noted. Raw fastq files were aligned to the human genome build hg38 using tophat2 (Kim et al. 2013) with the following parameters: –library-type fr-firststrand -a 15. Duplicated reads were removed using Picard tools command MarkDuplicates REMOVE_DUPLICATES=T, and RNA-seq quality metrics were assessed using Picard tools command CollectRnaSeqMetrics (http://broadinstitute.github.io/picard). Gene expression values were computed for each replicate across each condition using cuffdiff with an hg38 refflat file as reference with the following parameters: –library-type fr-firststrand. Subsequent analysis of gene expression and integration of gene expression data with other data types was performed in R.

Metascape analysis

Genes associated with various genomic features were identified using the GenomicRanges package in R and were analyzed for functional enrichment in the Metascape Web portal (http://metascape.org) using "Express Analysis" on default settings (Tripathi et al. 2015). Metascape gene set enrichment visualizations were performed in R.

Bisulfite-seq analysis

Bisulfite converted DNA sequencing data was processed using the BSseeker2 software suite (Guo et al. 2013). A bisulfite-sequencing amenable hg38 reference genome was built using the bs_seeker2-build.py command. DNA sequences were aligned to the hg38 bisulfite sequencing-amenable genome build using the bs_seeker2-align.py command with the following options: -m 6 -I 0 -X 800. Cytosine methylation levels were determined using bs_seeker2-call_methylation.py with default settings. Awk was used to convert CGmap files to HOMER-compatible allC formatted files. HOMER-formatted tag directories were built using HOMER's makeTagDirectories command with the following options: -format allC -minCounts 0 -genome hg38. Due to sequencing coverage-induced biases in DNA methylation meta-analysis (data not shown), awk was used to create HOMER-formatted tag directories containing only those cytosine residues covered by both unstimulated and PMA-activated neutrophil data sets. HOMER's annotatePeaks.pl command was used with the -ratio option to determine DNA methylation levels at particular genomic features.

Data visualization

Normalized HiC contact matrices presented in this study were generated using HiCPlotter (Akdemir and Chin 2015). HiC interactions and ChIP-seq data in Figure 4 and Supplemental Figures S4 and S6 were visualized using Sushi (Phanstiel et al. 2014). The remainder of linear genomic data was visualized using the Integrated Genomics Viewer (Robinson et al. 2011; Thorvaldsdóttir et al. 2013). FISH images were processed in FIJI. All other data were visualized using R.

Data availability

Data sets generated in this study are available as a series in the GEO database under accession number GSE126758.

Acknowledgments

We thank Alex Bortnick and other members of the Murre laboratory for editing the manuscript, and Yolanda Markaki and Kathrin Plath for assistance with the FISH protocol. This study was supported by funding from the Center for Computational Biology and Bioinformatics (UL1TRR001442), the California Institute for Regenerative Medicine (RB5-07025), and the National Institutes of Health to C.M. (AI082850, AI00880, and AI09599) and V.N. (1 U01 AI124316); from the Frontiers of Innovation Scholars Program to M.D.; and by the Waitt Advanced Biophotonics Core Facility of the Salk Institute with funding from National Institutes of Health-National Cancer Institute Cancer Center Support Grant (P30 014195) and the Waitt Foundation. T.I. was supported by the Uehara Memorial Foundation. High-throughput sequencing was performed at the Institutes for Genomic Medicine Genomics Center, University of California at San Diego.

Author contributions: M.D. performed the majority of the experiments and analysis. Y.Z. processed HiC samples. A.H., H.L., T.I., and S.D. provided technical support and advice. M.D. and C.M. wrote the manuscript. V.N. and C.M. supervised the study.

References

Akdemir K, Chin L. 2015. HiCPlotter integrates genomic data with interaction matrices. *Genome Biol* **16:** 198. doi:10.1186/s13059-015-0767-1

Bintu B, Mateo LJ, Su J-HH, Sinnott-Armstrong NA, Parker M, Kinrot S, Yamaya K, Boettiger AN, Zhuang X. 2018. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science (New York, NY)* **362:** eaau1783. doi:10.1126/science.aau1783

Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, Weinrauch Y, Zychlinsky A. 2004. Neutrophil extracellular traps kill bacteria. *Science (New York, NY)* **303:** 1532–1535. doi:10.1126/science.1092385

Bruno L, Ramlall V, Studer RA, Sauer S, Bradley D, Dharmalingam G, Carroll T, Ghoneim M, Chopin M, Nutt SL, et al. 2018. Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nat Immunol* **9:** 932–941. doi:10.1038/s41590-019-0471-5

Busslinger GA, Stocsits RR, van der Lelij P, Axelsson E, Tedeschi A, Galjart N, Peters J-MM. 2017. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544:** 503–507. doi:10.1038/nature22063

Cavaillon J-MM. 2011. The historical milestones in the understanding of leukocyte biology initiated by Elie Metchnikoff. *J Leuk Biol* **90:** 413–424. doi:10.1189/jlb.0211094

Chen Y, Zhang Y, Wang Y, Zhang L, Brinkman EK, Adam SA, Goldman R, van Steensel B, Ma J, Belmont AS. 2018. Mapping 3D genome organization relative to nuclear compartments using TSA-seq as a cytological ruler. *J Cell Biol* **217:** 4025–4048. doi:10.1083/jcb.201807108

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485:** 376–380. doi:10.1038/nature11082

Fanucchi S, Fok ET, Dalla E, Shibayama Y, Börner K, Chang EY, Stoychev S, Imakaev M, Grimm D, Wang KC, et al. 2019. Immune genes are primed for robust transcription by proximal long noncoding RNAs located in nuclear compartments. *Nat Genet* **51:** 138–150. doi:10.1038/s41588-018-0298-2

Fritz AJ, Barutcu A, Martin-Buley L, van Wijnen AJ, Zaidi SK, Imbalzano AN, Lian JB, Stein JL, Stein GS. 2016. Chromosomes at Work: organization of chromosome territories in the interphase nucleus. *J Cell Biochem* **117:** 9–19. doi:10.1002/jcb.25280

Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. 2016. Formation of chromosomal domains by loop extrusion. *Cell Rep* **15:** 2038–2049. doi:10.1016/j.celrep.2016.04.085

Garcia-Romo GS, Caielli S, Vega B, Connolly J, Allantaz F, Xu Z, Punaro M, Baisch J, Guiducci C, Coffman RL, et al. 2011. Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus. *Sci Transl Med* **3:** 73ra20.

Guo W, Fiziev P, Yan W, Cokus S, Sun X, Zhang MQ, Chen PY, Pellegrini M. 2013. BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics* **14:** 774. doi:10.1186/1471-2164-14-774

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *MolCell* **38:** 576–589.

Hoffmann K, Dreger CK, Olins AL, Olins DE, Shultz LD, Lucke B, Karl H, Kaps R, Müller D, Vayá A, et al. 2002. Mutations in the gene encoding the lamin B receptor produce an altered nuclear morphology in granulocytes (Pelger-Huët anomaly). *Nat Gen* **31:** 410–414. doi:10.1038/ng925

Hsu SC, Gilgenast TG, Bartman CR, Edwards CR, Stonestrom AJ, Huang P, Emerson DJ, Evans P, Werner MT, Keller CA, et al. 2017. The BET protein BRD2 cooperates with CTCF to enforce transcriptional and architectural boundaries. *Mol Cell* **66:** 102–116.e7. doi:10.1016/j.molcel.2017.02.027

Huff J. 2016. The Fast mode for ZEISS LSM 880 with Airyscan: high-speed confocal imaging with super-resolution and improved signal-to-noise ratio. *Nat Meth* **13:** i–ii. doi:10.1038/nmeth.f.398

Isoda T, Moore AJ, He Z, Chandra V, Aida M, Denholtz M, Piet van Hamburg J, Fisch KM, Chang AN, Fahl SP, et al. 2017. Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. *Cell* **171:** 103–119.e18. doi:10.1016/j.cell.2017.09.001

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14:** R36. doi:10.1186/gb-2013-14-4-r36

Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, Singh H. 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science (New York, NY)* **296:** 158–162. doi:10.1126/science.1068768

Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, et al. 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22:** 1813–1831. doi:10.1101/gr.136184.111

Langmead B, Trapnell C, Pop M, Salzberg S. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25. doi:10.1186/gb-2009-10-3-r25

Lawrence M, Huber W, Pagès P, Aboyoun P, Carlson M, Gentleman R, Morgan M, Carey V. 2013. Software for computing and annotating genomic ranges. *PLoS Comp* **9:** e10003118.

Ley K, Hoffman HM, Kubes P, Cassatella MA, Zychlinsky A, Hedrick CC, Catz SD. 2018. Neutrophils: new insights and open questions. *Sci Immunol* **3:** eaat4579. doi:10.1126/sciimmunol.aat4579

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (New York, NY)* **326:** 289–293. doi:10.1126/science.1181369

Lin YC, Benner C, Mansson R, Heinz S, Miyazaki K, Miyazaki M, Chandra V, Bossen C, Glass CK, Murre C. 2012. Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nat Immunol* **13:** 1196–1204. doi:10.1038/ni.2432

Muller W. 2013. Getting leukocytes to the site of inflammation. *Vet Pathology* **50:** 7–22. doi:10.1177/0300985812469883

Nora EP, Goloborodko A, Valton A-LL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. 2017. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169:** 930–944.e22. doi:10.1016/j.cell.2017.05.004

Olins AL, Hoang TV, Zwerger M, Herrmann H, Zentgraf H, Noegel AA, Karakesisoglou I, Hodzic D, Olins DE. 2009. The LINC-less granulocyte nucleus. *European Cell Biol* **88:** 203–214. doi:10.1016/j.ejcb.2008.10.001

Ollion J, Cochennec J, Loll F, Escudé C, Boudier T. 2013. TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinformatics* **29:** 1840–1841. doi:10.1093/bioinformatics/btt276

Pastore N, Brady OA, Diab HI, Martina JA, Sun L, Huynh T, Lim JA, Zare H, Raben N, Ballabio A, et al. 2016. TFEB and TFE3 cooperate in the regulation of the innate immune response in activated macrophages. *Autophagy* **12:** 1240–1258. doi:10.1080/15548627.2016.1179405

Phanstiel D, Boyle A, Araya C, Snyder MP. 2014. Sushi.R: flexible, quantitative and integrative genomic visualizations for publication-quality multi-panel figures. *Bioinformatics* **30:** 2808–2810. doi:10.1093/bioinformatics/btu379

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159:** 1665–1680. doi:10.1016/j.cell.2014.11.021

Rao SS, Huang S-CC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon K-RR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, et al. 2017. Cohesin loss eliminates all loop domains. *Cell* **171:** 305–320.e24. doi:10.1016/j.cell.2017.09.026

Robinson J, Thorvaldsdóttir H, Winckler W, Guttman M, Lander E, Getz G, Mesirov J. 2011. Integrative genomics viewer. *Nat Biotech* **29:** 24–26. doi:10.1038/nbt.1754

Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res* **22:** 939–946. doi:10.1101/gr.128124.111

Rowat AC, Jaalouk DE, Zwerger M, Ung W, Eydelnant IA, Olins DE, Olins AL, Herrmann H, Weitz DA, Lammerding J. 2013. Nuclear envelope composition determines the ability of neutrophil-type cells to passage through micron-scale constrictions. *J Biol Chem* **288:** 8610–8618. doi:10.1074/jbc.M112.441535

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9:** 676–682. doi:10.1038/nmeth.2019

Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, Fonseca NA, Huber W, Haering CH, Mirny L, et al. 2017. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551:** 51–56. doi:10.1038/nature24281

Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-JJ, Vert J-PP, Heard E, Dekker J, Barillot E. 2015. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol* **16:** 259. doi:10.1186/s13059-015-0831-x

Shultz LD, Lyons BL, Burzenski LM, Gott B, Samuels R, Schweitzer PA, Dreger C, Herrmann H, Kalscheuer V, Olins AL, et al. 2003. Mutations at the mouse ichthyosis locus are within the lamin B receptor gene: a single gene model for human Pelger-Huët anomaly. *Hum Mol Genet* **12:** 61–69. doi:10.1093/hmg/ddg003

Tecchio C, Micheletti A, Cassatella MA. 2014. Neutrophil-derived cytokines: facts beyond expression. *Front Immunol* **5:** 508. doi:10.3389/fimmu.2014.00508

Thomas CJ, Schroder K. 2013. Pattern recognition receptor function in neutrophils. *Trends Immunol* **34:** 317–328. doi:10.1016/j.it.2013.02.008

Thorvaldsdóttir H, Robinson JT, Mesirov J. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinformatics* **14:** 178–192. doi:10.1093/bib/bbs017

Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM, DeJesus P, Che J, Mulder LC, et al. 2015. Meta- and orthogonal integration of influenza 'OMICs' data defines a role for UBR4 in virus budding. *Cell Host Microbe* **18:** 723–735. doi:10.1016/j.chom.2015.11.002

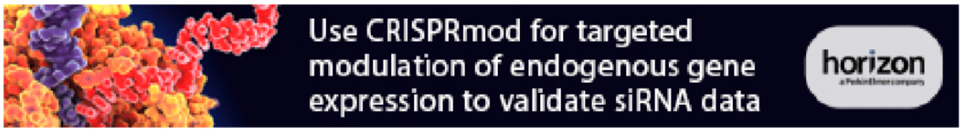
Zhu Y, Gong K, Denholtz M, Chandra V, Kamps MP, Alber F, Murre C. 2017. Comprehensive characterization of neutrophil genome topology. *Genes Devel* **31:** 141–153. doi:10.1101/gad.293910.116

Genes & Development

Upon microbial challenge, human neutrophils undergo rapid changes in nuclear architecture and chromatin folding to orchestrate an immediate inflammatory gene program

Matthew Denholtz, Yina Zhu, Zhaoren He, et al.

Genes Dev. 2020, **34:** originally published online January 9, 2020
Access the most recent version at doi:10.1101/gad.333708.119



4.2 Responses of Neutrophils to Calcium Signaling

Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization

Yina Zhu,[1] Matthew Denholtz,[1] Hanbin Lu, and Cornelis Murre

Division of Biological Sciences, Department of Molecular Biology, University of California at San Diego, La Jolla, California 92093, USA

During developmental progression the genomes of immune cells undergo large-scale changes in chromatin folding. However, insights into signaling pathways and epigenetic control of nuclear architecture remain rudimentary. Here, we found that in activated neutrophils calcium influx rapidly recruited the cohesin-loading factor NIPBL to thousands of active enhancers and promoters to dictate widespread changes in compartment segregation. NIPBL recruitment to enhancers and promoters occurred with distinct kinetics. The induction of NIPBL-binding was coordinate with increased P300, BRG1 and RNA polymerase II occupancy. NIPBL-bound enhancers were associated with NFAT, PU.1, and CEBP *cis* elements, whereas NIPBL-bound promoters were enriched for GC-rich DNA sequences. Using an acute degradation system, we found that the histone acetyltransferases P300 and CBP maintained H3K27ac abundance and facilitated NIPBL occupancy at enhancers and that active transcriptional elongation is essential to maintain H3K27ac abundance. Chromatin remodelers, containing either of the mutually exclusive BRG1 and BRM ATPases, promoted NIPBL recruitment at active enhancers. Conversely, at active promoters, depletion of BRG1 and BRM showed minimal effect on NIPBL occupancy. Finally, we found that calcium signaling in both primary innate and adaptive immune cells swiftly induced NIPBL occupancy. Collectively, these data reveal how transcriptional regulators, histone acetyltransferases, chromatin remodelers, and transcription elongation promote NIPBL occupancy at active enhancers while the induction of NIPLB occupancy at promoters is primarily associated with GC-rich DNA sequences.

[*Keywords*: chromatin remodelers; cohesin; enhancers; NIPBL; neutrophils; P300 and CBP; promoters]

Supplemental material is available for this article.

Received August 6, 2020; revised version accepted October 22, 2020.

It is now well established that the mammalian genome is highly structured. Chromatin is organized as loop domains or topologically associating domains, that fold into either the transcriptionally active (A) compartment or repressive (B) compartment (Lieberman-Aiden et al. 2009; Dixon et al. 2012; Gibcus and Dekker 2013; Rao et al. 2014; Rowley and Corces 2018). Loop domains are established by cohesin-mediated extrusion of chromatin. Cohesin is loaded onto chromatin primarily by NIPBL and MAU2 (Ciosk et al. 2000; Nasmyth and Haering 2009; Murayama and Uhlman 2014; Petela et al. 2018; Parenti et al. 2020). Once loaded, cohesin progressively extrudes chromatin loops until a pair of convergent CTCF sites is reached (Kueng et al. 2006; Rao et al. 2014; Sanborn et al. 2015; Fudenberg et al. 2016; Haarhuis et al. 2017; Wutz et al. 2017; Vian et al. 2018). Release of cohesin from chromatin is orchestrated by WAPL (Kueng et al. 2006; Haarhuis et al. 2017; Wutz et al. 2017). NIPBL and cohesin mediated formation of loop domains suppress compartmental segregation allowing for fine-scale control over local chromatin interactions and regulation of gene expression (Haarhuis and Rowland 2017; Rao et al. 2017; Schwarzer et al. 2017; Nuebler et al. 2018).

Neutrophils are innate antimicrobial effector cells that rapidly respond to a diverse range of extracellular stimuli. Upon microbial encounter, neutrophils activate inflammatory gene programs, engulf bacteria through phagocytosis, discharge microbicidal factors into the extracellular space, and release an extracellular network of chromatin fibers to trap invading pathogens (Brinkmann et al. 2004; Tecchio et al. 2014; Ley et al. 2018). The

---

[1] **These authors contributed equally to this work.**
Corresponding author: cmurre@ucsd.edu
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.343475.120.

activation of neutrophils involves both calcium-independent and calcium-dependent mechanisms (Kenny et al. 2017). Intracellular calcium levels are tightly controlled and elevated upon stimulation (Hogan et al. 2010). A prominent target of calcium is nuclear factor of activated T cells (NFAT) that comprises a family of transcription factors (TFs) important for regulating T and B cell activation as well as innate immune responses (Fric et al. 2012; Vaeth et al. 2018). Inactive NFAT proteins are generally excluded from the nucleus through the phosphorylation of specific serine residues. Calcium influx activates the phosphatase calcineurin, which removes the regulatory phosphate groups and allows the translocation of NFAT into the nucleus (Hogan et al. 2003, 2010).

In recent studies we demonstrated that upon activation human neutrophils rapidly sequester cohesin at H3K27ac-marked enhancers (Denholtz et al. 2020). Here we examined in neutrophils whether and how calcium signaling modulates nuclear architecture. We found that calcium signaling triggered rapid changes in compartment segregation. Calcium-induced alterations in compartmentalization were closely associated with calcium-induced NIPBL occupancy. Calcium-induced recruitment of NIPBL involved a wide spectrum of TFs, including NFAT and PU.1. The induction of NIPBL-binding was coordinate with increased co-occupancy of the histone acetyltransferase (HAT) P300, the chromatin remodeler BRG1, and RNA polymerase II (RNAPII). To delineate the pathway that underpins calcium-induced NIPBL recruitment we generated an armamentarium of cell lines that carried inducible degrons. We found that acute degradation of P300- and CREB-binding protein (CBP) or SPT5 reduced H3K27ac abundance and decreased genome-wide NIPBL occupancy. Acute degradation of both BRG1 and BRM eliminated calcium-induced NIPBL occupancy at active enhancers but not at promoters. Finally, we found that calcium signaling orchestrated lineage-specific NIPBL occupancy in primary innate and adaptive immune cells. Collectively, these data reveal that active transcriptional elongation maintains H3K27ac abundance and show in mechanistic terms how calcium influx dictates NIPBL occupancy at active enhancers and promoters to orchestrate changes in nuclear architecture.

Results

Calcium influx triggers a weakening of compartment segregation

In previous studies we demonstrated that the genomes of mature neutrophils are characterized by a highly segregated compartment structure (Zhu et al. 2017; Denholtz et al. 2020). We also showed that exposure of human neutrophils to microbial agents readily induced recruitment of cohesin to H3K27ac marked enhancers (Denholtz et al. 2020). As a model system to determine how signals emanating from the cell surface instruct alterations in nuclear morphology we used differentiated neutrophils derived from an immortalized neutrophil progenitor cell line, ECOMG (Zhu et al. 2017). To study the impact of activation on neutrophil genome topology, ECOMG-derived neutrophils were cultured in the presence of the calcium ionophore A23187 for different durations and using different concentrations. Activated neutrophils were examined for changes in nuclear morphology using DAPI, Lamin B1, and HP1γ staining. We found that when treated with 5 µM A23187 for a duration of 4 h, neutrophils displayed a loss of lobular nuclear structures (data not shown). Treatment with higher concentrations of A23187 (20 µM) for 30 min resulted in a robust and homogeneous loss of lobular nuclear structures and spreading of condensed chromatin from the nuclear periphery across the entire nucleoplasm (Fig. 1A).

To examine how calcium signaling modulates genome topology we used in situ Hi-C. More than 180 million valid contacts were generated for each condition (Supplemental Table S1). We found that contact frequencies as a function of genomic distance scaled similarly for resting and activated neutrophils (Supplemental Fig. S1A). However, we found that contact matrices revealed substantial differences in long-range genomic interactions when comparing naïve versus activated neutrophils (Fig. 1B). Resting neutrophils were associated with a plaid-pattern involving long-range *cis* contacts in Hi-C contact matrices that were weakened in activated neutrophils (Fig. 1B). We next examined A and B compartment segregation using principal component analysis of normalized Hi-C matrices. Whereas resting neutrophil genomes were associated with highly segregated compartments, activated neutrophils revealed at the boundaries of small compartment domains a striking loss of compartmentalization, with small B regions converging into A regions and small A regions converging into B regions (Fig. 1C,D; Supplemental Fig. S1B). The loss of finely structured compartments resulted in a decrease in the numbers of principal domains (PDs) in A and B compartments (Supplemental Fig. S1C). PDs that switched compartments upon neutrophil activation were on average of smaller size than those that did not undergo changes in compartmentalization (Supplemental Fig. S1D,E; Supplemental Table S2). Upon calcium-induced neutrophil activation, PDs that switched compartments showed intermediate PC1 values (Supplemental Fig. S1F; Supplemental Table S2). The number of genomic regions that converted from the B to the A compartment was significantly higher than those that repositioned from the A to B compartment. We quantified the weakening of compartment segregation by calculating compartment strength. We found that calcium signaling depleted for AA and BB interactions whereas AB interactions were enriched (Fig. 1E). We validated this analysis by plotting the ratios for average interaction frequencies involving the same compartment (AA and BB) versus those involving different compartments (AB). This analysis validated the notion that calcium signaling diminishes segregation of A and B compartments (Fig. 1F).

Previously we found that during neutrophil differentiation A and B PDs that span large genomic distances fragmented into smaller domains (Zhu et al. 2017). Here we found that calcium entry in neutrophils induced a reversed pattern: highly segmented compartments that merged into singular compartments resembling that of

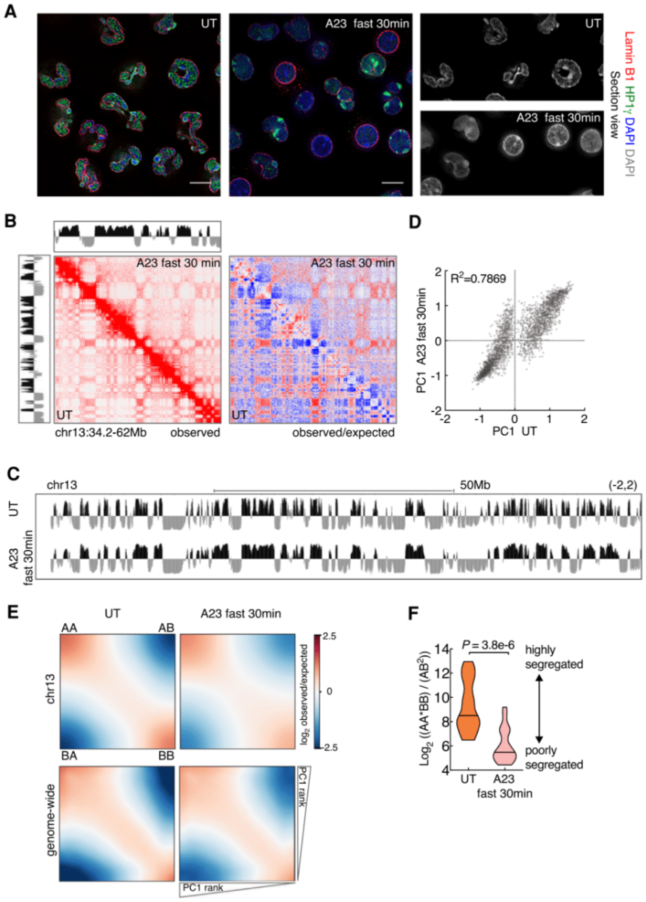


Figure 1. Calcium entry triggers a weakening of compartmentalization. (*A*) Immunofluorescence images comparing nuclear morphology of untreated (UT) and A23187-activated (A23187 fast 30 min) neutrophils. Neutrophils were activated using fast activation condition (20 µM A23187) for 30 min. Nuclear lamina was stained using anti-LaminB1. Anti-HP1γ stained HP1γ that binds both euchromatin and heterochromatin. Counterstain with DAPI highlighted heterochromatin. Representative image sections are presented as digitally magnified images. Original magnification, ×100. Scale bars, 5 µm. (*B*) Hi-C contact matrices for chromosome 13 (34.2–62 Mb) in untreated and activated neutrophils. The intensity of each pixel represents the normalized number of contacts between a pair of loci. (*Left*) Observed contact maps were balanced normalized at 50-kb resolution with a color range of 0–20. (*Right*) Observed/expected contact maps were balanced normalized and distance normalized at 50-kb resolution with a color range of −4 to 4 ($\log_2$ ratio). Corresponding compartment profiles (PC1 tracks) at 25-kb resolution, indicating genomic regions positioned in the A (black) and B (gray) compartments are shown *above* and to the *left* of the matrix for activated and untreated neutrophils, respectively. (*C*) Genome browser tracks for PC1 spanning chromosome 13 in untreated and activated neutrophils. (*D*) Scatter plot for PC1 values of untreated versus activated neutrophils, defined at PC1 domain (PD) level is shown. Signals away from the diagonal indicate PDs that switched compartments. Pearson coefficients are indicated. (*E*) Saddle plots for compartmentalization strength in untreated (*left* panels) and A23187-activated (*right* panels) neutrophils. Average distance-normalized contact frequencies between *cis* pairs of 50-kb bins arranged by descending PC1 values. The *top* shows saddle plots for chr13. The *bottom* shows aggregated map constructed by averaging all chromosomes. (*F*) Violin plots comparing distributions of $\log_2$ ratios for average genomic interaction frequencies involving different compartments (AB) compared with those involving the same compartments (AA and BB). The median values are shown as central black lines. *P*-value, two-tailed Wilcoxon rank-sum test.

neutrophil progenitors. To explore whether and how changes in compartmentalization and epigenetic marking are related, we examined resting and activated neutrophils for the deposition of H3K27me3 and H3K36me3. We found that genomic regions that switched from A to B compartments were initially associated with actively transcribed regions (H3K27me3^{low}H3K36me3^{hi}), while genomic regions that switched from B to A compartments were associated with transcriptionally silent genomic regions (H3K27me3^{hi}H3K36me3^{low}) (Supplemental Fig. S1G–J). Finally, we compared the PC1 values of PDs that upon activation switched compartments to that of other cell types. We found that PDs that upon neutrophil activation repositioned from A to B compartments were predominantly located in the B compartment in other immune cells (Supplemental Fig. S1K). Conversely, PDs that switched from B to A compartments were primarily localized in the A compartments in other immune cells (Supplemental Fig. S1K). In sum, these data indicate that in neutrophils calcium entry induced a weakening of compartment segregation to establish a topology akin to that of other immune cells.

Calcium signaling elevates NIPBL and SMC1 occupancy at active enhancers and promoters

Recent studies demonstrated that chromatin organization emerges from competition of nonequilibrium active loop extrusion and epigenetically defined compartment structure (Haarhuis and Rowland 2017; Rao et al. 2017; Schwarzer et al. 2017; Nuebler et al. 2018). To explore the possibility that calcium-induced loop extrusion suppressed compartment segregation we examined for NIPBL and SMC1 occupancy. We found that upon calcium entry, genome-wide NIPBL occupancy greatly increased (Supplemental Fig. S2A,B). NIPBL-bound sites in resting and activated neutrophils, using both fast and slow activating conditions, partially overlapped, indicating that calcium entry elevates NIPBL occupancy at pre-existing as well as de novo sites (Supplemental Fig. S2A). Among the 55,075 NIPBL-binding sites that were identified and merged from the aforementioned conditions, NIPBL occupancy at 22,269 sites was significantly elevated using either fast or slow activation conditions ($FC > 4$, $P < 0.0001$).

To investigate how NIPBL occupancy is distributed across the promoter and enhancer landscape, resting and activated neutrophils were examined for the deposition of epigenetic modifications that mark promoters (H3K4me3), poised enhancers (H3K4me1) and active promoters and enhancers (H3K27ac). Additionally, we examined resting and activated neutrophils for nucleosome density using ATAC-seq. K-means clustering revealed that calcium-induced NIPBL-bound sites were distributed as groups across the promoter (clusters P1 and P2) and enhancer landscapes (clusters E1, E2, and E3) (Fig. 2A; Supplemental Fig. S2B–D; Supplemental Table S3). NIPBL-bound promoters segregated into two clusters, P1 and P2, reflecting differences in transcriptional direction, while NIPBL-bound enhancers segregated into three clusters reflecting differences in the deposition of H3K27ac across active enhancers (E1 and E2) versus poised enhancers (E3) (Fig. 2A). Prior to activation, enhancers and promoters exhibited relatively low levels of NIPBL occupancy (Fig. 2A). Upon calcium entry NIPBL-binding strength was elevated at both active enhancers and promoters but to a much greater degree at enhancers (Fig. 2A). We found that induced NIPBL levels at enhancers were proportional to pre-existing levels of NIPBL ($E1 > E2 > E3$) and proportional to pre-existing H3K27ac abundance and chromatin accessibility (Fig. 2A). As expected, calcium-induced NIPBL occupancy was coordinate with increased SMC1 occupancy (Fig. 2B).

Consistent with previous studies, SMC1 bound sites were associated either with CTCF or NIPBL occupancy (Kagey et al. 2010; Vian et al. 2018). Using four-cluster K-means clustering analysis we identified two distinct groups of CTCF (C1 and C2) and NIPBL (N1 and N2) bound sites (Fig. 2C). Upon calcium influx, SMC1 occupancy at CTCF-bound sites essentially remained the same (Fig. 2C). In contrast, calcium-induced SMC1 occupancy was significantly elevated at NIPBL-bound sites ($P = 1.82 \times 10^{-12}$) (Fig. 2C). Finally, SMC1 levels at CTCF/SMC1-bound sites positively correlated with CTCF motif densities ($C1 > C2$) (Fig. 2D). SMC1 levels at NIPBL/SMC1 bound sites positively correlated with NIPBL-binding strength ($N1 > N2$) coordinate with enhancer and promoter activity (Fig. 2C). Taken together, these data indicate that upon calcium entry NIPBL and SMC1 are primarily sequestered at H3K27ac-marked promoters and H3K27ac-marked enhancers but not at H3K4me1-marked poised enhancers.

Calcium signaling instructs with distinct kinetics P300, BRG1, RNAPII, and NIPBL occupancy at enhancers versus promoters

To examine whether and how upon calcium influx NIPBL coordinately binds with P300, BRG1, and RNAPII to promoter and enhancer regions we used ChIP-seq. We found that prior to activation P300, BRG, RNAPII, and NIPBL showed low occupancy (Fig. 3A–D). Upon calcium influx, the induction of NIPBL occupancy was coordinate with P300, BRG1, and RNAPII binding (Fig. 3A–D). To determine the kinetics associated with NIPBL, BRG1, P300, and RNAPII Ser2 recruitment, we used slow activation conditions. We observed a gradual and sequential induction of NIPBL, BRG1, P300, and RNAPII Ser2 occupancy at active enhancers and promoters (Fig. 3D; Supplemental Fig. S3A,B). We found that NIPBL, BRG1, and P300 accumulated with faster kinetics at enhancers than at promoters (Fig. 3D; Supplemental Fig. S3A,B). At active enhancers, calcium influx activated a gradual recruitment of NIPBL that was initiated within 1 h to reach maximum levels 3–4 h after activation (Fig. 3D). Conversely at promoters, calcium-induced NIPBL recruitment did not occur until 3–4 h after stimulation (Fig. 3D). P300 and BRG1 occupancy mirrored the kinetics of NIPBL binding (Supplemental Fig. S3A,B). The kinetics of H3K27ac abundance at enhancers and promoters reached maximum levels 1 h after activation, which was maintained for the duration of calcium influx (Fig. 3E). RNAPII (RNAPII Ser2) rapidly associated with both active promoters and enhancers (Supplemental Fig. S3C). RNAPII Ser2 accumulation at active enhancers was particularly prominent using fast activation conditions (15 min) when compared with slow activation conditions (4 h), plausibly imposed by time-limited transcriptional elongation (Fig. 3C; Supplemental Fig. S3C). Taken together, these data indicate that calcium entry recruits NIPBL, P300, BRG1, and RNAPII Ser2 at active enhancers and promoters but with different kinetics.

Upon calcium entry NIPBL is sequestered at active promoters and translocated coordinately with elongating RNA polymerase

To determine whether and how nascent transcript abundance relates to NIPBL recruitment at promoters, neutrophils were activated for 4 h and analyzed using GRO-seq. We found that calcium signaling elevated the abundance of 671 transcripts, while 1554 transcripts showed decreased abundance (Fig. 4A; Supplemental Table S4). Prominent pathways that were affected included the NF-κB, inflammatory and cytokine mediated signaling cascades as well as TFs involved in the immediate stress response, including *nr4a1-3*, *fos*, and *fosb* (Supplemental Table S5). To determine how calcium entry modulates

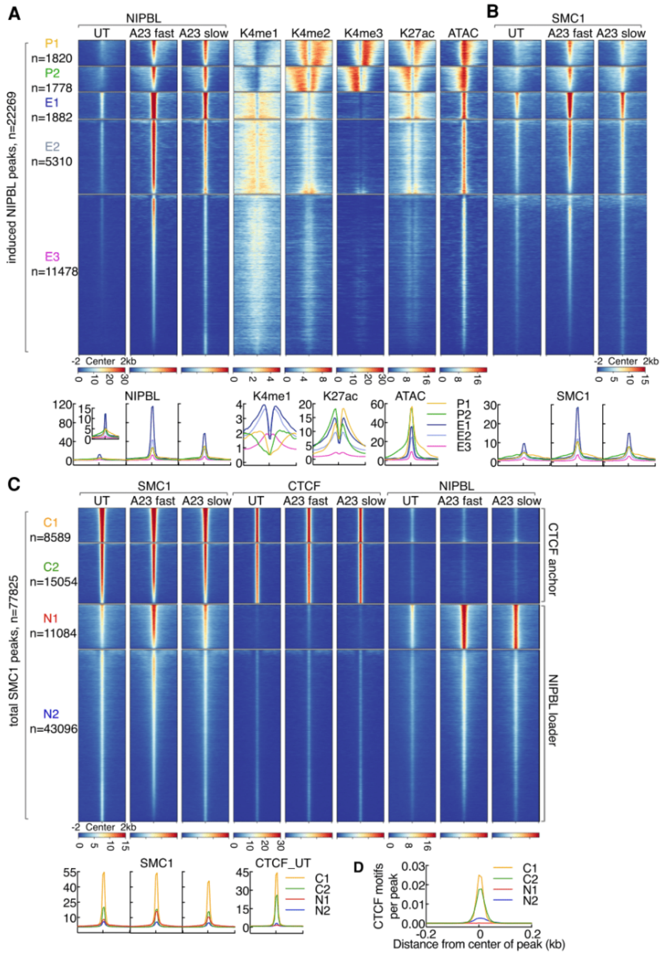


Figure 2. Calcium signaling rapidly elevates NIPBL occupancy at active enhancers and promoters. (*A*) Heat map of K-means clustering analysis for NIPBL and epigenetic marks as revealed by ChIP-seq and ATAC-seq signals for ±2-kb windows centered on induced NIPBL-binding sites (n = 22,269, FC > 4, K = 5). Each cluster was manually classified as promoters (P1 and P2) or enhancers (E1, E2, and E3) according to epigenetic features. Numbers of NIPBL-binding sites for each cluster are denoted at the *left*. The five clusters were saved as a reference bed file to compute matrices for the following analysis. Heat maps were generated by sorting maximum signal intensities associated with NIPBL occupancy for neutrophils cultured in the presence of A23187 (fast activation conditions) and calculated across the entire 4-kb interval around peak summits. Color scale intensities represent normalized reads scores (reads per 10 million uniquely mapped reads per base pair). Note that the blue-yellow-red color scheme was used for ChIP-seq, GRO-seq, and ATAC-seq heat maps throughout the manuscript except for the heat map shown in Figure 5G. Corresponding average enrichment plots with mean signal intensities (per base pair per peak) are shown *below*. Neutrophils were activated either in fast activation condition (20 µM A23187 for 15 min), or in slow activation condition (5 µM A23187 for 4 h). These two activation conditions were used throughout the manuscript unless otherwise mentioned. (*B*) Heat map of SMC1 ChIP-seq signals plotted for five clusters defined in *A*. Heat map was sorted by the maximum signal intensity of SMC1 occupancy using A23187 fast activation conditions. Corresponding average enrichment plots are shown *below* the heat maps. (*C*) Heat map of K-means clustering analysis for SMC1, CTCF, and NIPBL ChIP-seq signals across a region of ±2 kb centered on total SMC1 peaks (n = 77,825, K = 4). Each cluster was manually classified as anchors or loaders according to its characteristic features. Heat map was sorted by the maximum signal intensity associated with SMC1 occupancy in A23187 fast activation conditions. Corresponding average enrichment plots are shown *below* the heat maps. (*D*) Average enrichment plots for mean motif densities (motifs per base pair per peak) associated with CTCF motifs, centered on CTCF-binding sites (±200 bp) are shown for clusters C1 and C2.

transcript abundance during the course of 1–4 h after calcium influx we performed RPB1 and RNAPIISer2 ChIP-seq. As expected, RBP1 and RNAPII Ser2 abundance differed when comparing naïve versus activated neutrophils (Supplemental Fig. S4A). For a duration of 1–4 h after activation, RBP1 and RNAPII Ser2 abundance remained equivalent (Supplemental Fig. S4A). At activated genes calcium entry prompted productive RNAPII elongation along gene bodies while RNAPII and RNAPII Ser2 were depleted from gene-bodies and promoters at repressed genes (Supplemental Fig. S4B). RBP1 and RNAPII Ser2 abundance reached maximum levels already 1 h after activation (Supplemental Fig. S4B,C). Meta-gene analysis revealed tightly correlated RPB1, RNAPII Ser2, and NIPBL occupancy (Fig. 4B). Upon activation, the deposition of H3K27ac at promoters rapidly increased to reach maximum levels at 1 h, while NIPBL occupancy steadily increased during the course of activation (Fig. 4C). NIPBL occupancy at promoters and gene bodies also correlated well with global mRNA abundance (Fig. 4D). To determine whether and how transcriptional initiation and elongation modulates calcium-induced NIPBL-

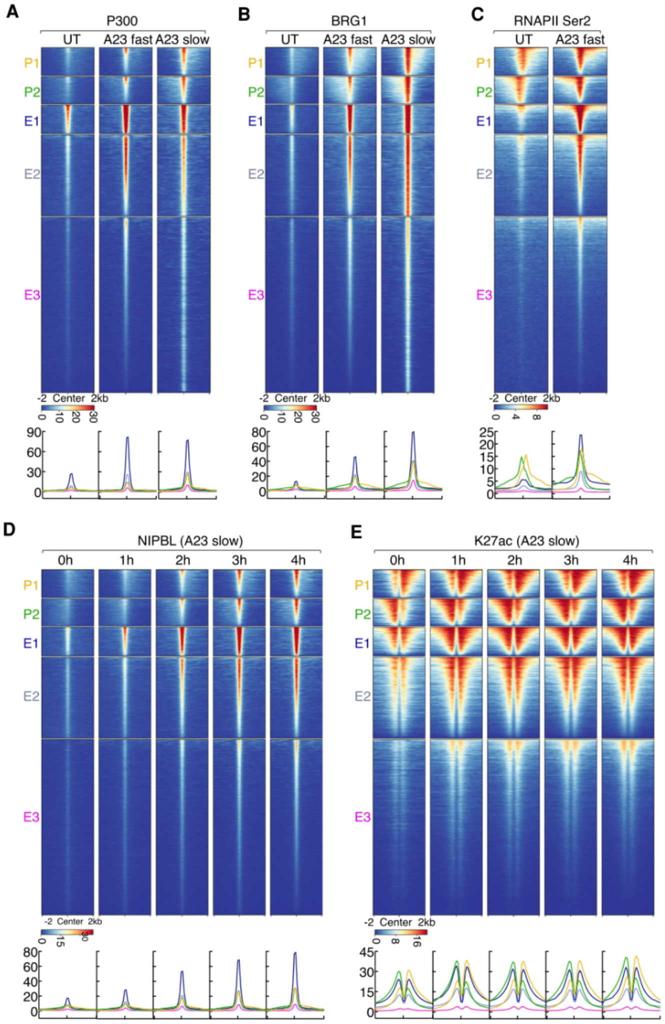


Figure 3. Calcium signaling rapidly elevates with distinctive kinetics P300, BRG1, and RNAPII co-occupancy at NIPBL-bound enhancers and promoters. (*A*) Heat map displaying P300 ChIP-seq signals plotted for five clusters of genomic regions defined in Figure 2A. Heat map was sorted by the maximum signal intensity of P300 in A23187 fast activation conditions. (*B*) Heat map displaying BRG1 ChIP-seq signals plotted for five clusters of genomic regions defined in Figure 2A is shown. Heat map was sorted by the maximum signal intensity of BRG1 in A23187 fast activation conditions. (*C*) Heat map displaying RNAPIISer2 ChIP-seq signals plotted for five clusters of genomic regions defined in Figure 2A. Heat map was sorted by the maximum signal intensity of RNAPII in A23187 fast activation conditions. (*D*) Kinetics of NIPBL recruitment is shown. Heat map of NIPBL ChIP-seq signals was plotted for five clusters of genomic regions defined in Figure 2A. Heat map was generated by centering on the maximum signal intensity of NIPBL occupancy in A23187 slow activation conditions (4 h). (*E*) Kinetics of H3K27ac distribution is shown. Heat map of H3K27ac ChIP-seq signals was plotted for five clusters of genomic regions defined in Figure 2A. Heat map was sorted by centering on the maximum signal intensity of H3K27ac in A23187 slow activation conditions (4 h). (*A–E*) *Below* the heat maps are indicated the corresponding average enrichment plots.

occupancy, we cultured neutrophils in the presence of Triptolide (TRP) and 5,6-dichlorobenzimidazole 1-β-D-ribofuranoside (DRB). Specifically, neutrophils were cultured for 2 h in the presence of inhibitors, followed by another 2 h in the presence of A23187. We found that in the presence of inhibitors calcium-induced NIPBL occupancy significantly accumulated at TSSs but was depleted across gene bodies (Fig. 4E–G; Zhou et al. 2012).

To determine how upon calcium entry NIPBL accumulation was affected by transcription amplitude, genes were ranked based on GRO-seq reads and compared with the strength of NIPBL occupancy at TSSs. We found that the strength of DRB-accumulated NIPBL occupancy at promoters positively correlated with nascent transcript abundance (Fig. 4H). Specifically, interference with transcriptional elongation slightly lowered NIPBL occupancy at modestly transcribed genes whereas genes that were transcribed at relatively high rates showed a substantial increase in NIPBL binding strength (Fig. 4H). These data indicate that transcriptional elongation facilitates NIPBL occupancy and promotes NIPBL translocation as previously reported for cohesin (Busslinger et al. 2017). Taken together, these data indicate that upon calcium entry NIPBL is sequestered at active promoters

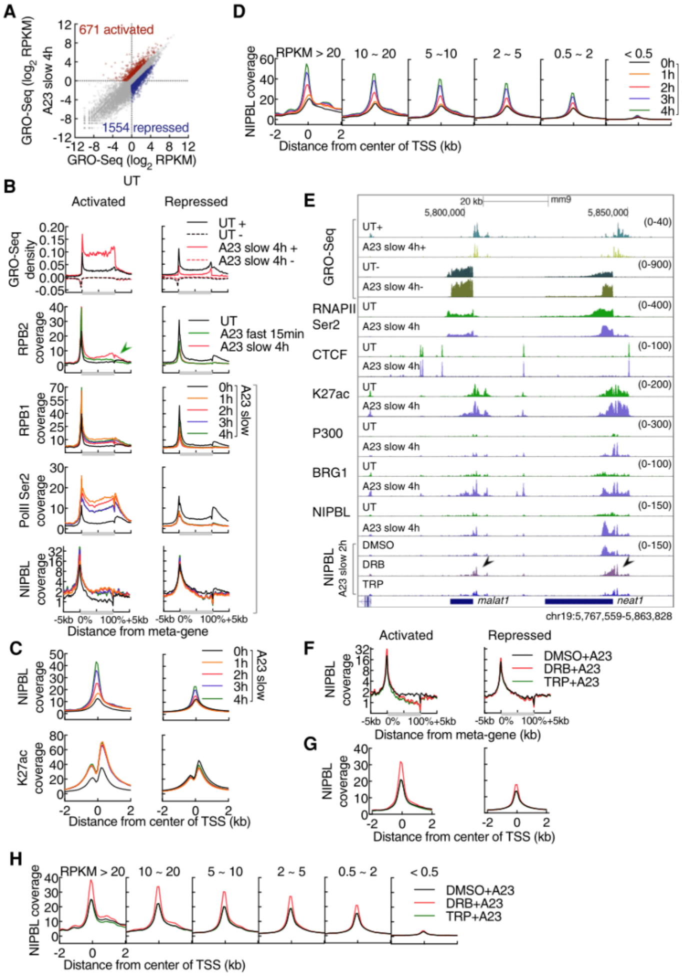


Figure 4. Transcriptional elongation modulates NIPBL occupancy at promoters. (*A*) Scatter plot for GRO-seq read density of untreated versus activated neutrophils. Color-codes indicate a decrease (blue) or increase (red) in transcript abundance in activated versus naïve neutrophils (FC > 2, RPKM > 0.5). Numbers of genes are indicated in parenthesis. (*B*) Meta-gene profile of GRO-seq (tags per base pair per peak) and ChIP-seq signals (tags per base pair per peak) across genic regions in untreated and activated neutrophils, gated on differentially expressed genes as defined in *A*. Gene lengths (from TSS to TTS) were rescaled to the same pseudo-length as shown on the *X*-axis. Sense (+) and anti-sense (−) transcripts are plotted *above* and *below* the *X*-axis, respectively. Note that the green arrow shows an atypical distribution pattern for RNAPII at activated genes in response to fast activation. (*C*) Average enrichment plots for NIPBL and H3K27ac ChIP-seq signals in untreated and activated neutrophils, plotted as a function of genomic distance from the center of TSSs (±2 kb), gated on differentially expressed genes as defined in *A*. (*D*) Average enrichment plots for NIPBL ChIP-seq signals in untreated and activated neutrophils, plotted as a function of genomic distance from the center of TSSs (±2 kb). Genes are grouped according to expression levels in activated neutrophils defined by GRO-seq reads. (*E*) Genome browser tracks of GRO-seq and ChIP-seq signals in untreated and activated neutrophils, pretreated with or without transcription inhibitors for 2 h, for representative loci named *malat1* and *neat1* are shown. DRB refers to 5,6-dichlorobenzimidazole 1-β-D-ribofuranoside. TRP refers to triptolide (TRP). Black arrowheads indicate NIPBL bound sites affected by DRB. (*F*) Meta-gene profile associated with NIPBL ChIP-seq reads across genic regions pretreated with or without transcription inhibitors for 2 h, gated on differentially expressed genes as defined in *A*. Gene lengths were rescaled to the same size as shown for the *X*-axis. (*G*) Average enrichment plots for NIPBL ChIP-seq signals in activated neutrophils, pretreated with or without transcription inhibitors for 2 h, plotted as a function of genomic distance from the center of TSSs (±2 kb), gated on differentially expressed genes as defined in *A*. (*H*) Average enrichment plots for NIPBL ChIP-seq signals in activated neutrophils, pretreated with or without transcription inhibitors for 2 h, plotted as a function of genomic distance from the center of TSSs (±2 kb). Genes were grouped according to expression levels in activated neutrophils as defined by GRO-seq signals.

and translocated coordinately with elongating RNA polymerase.

Targeting histone acetyltransferases, chromatin remodelers, and components of the transcriptional and cohesin machinery for acute depletion

The data described above indicate that calcium signaling promotes P300, BRG1, and RNAPII occupancy at NIPBL-bound sites. To explore the immediate effects of these factors in orchestrating calcium-induced NIPBL occupancy we applied the FKBP12^{F36V} acute degradation system. FKBP12^{F36V} permits, in the presence of a heterobifunctional molecule named dTAG, rapid, specific, and near-complete degradation of targeted proteins (Nabet et al. 2018). The acute degradation system was established in ECOMG progenitor cells using CRISPR–Cas9 genome editing. Specifically, proteins of interest were tagged in frame at the N-terminal or C- terminal region with FKBP12^{F36V} (Supplemental Fig. S5A). To monitor expression levels, tagged proteins were also marked with eYFP or mScarlet as well as an HA epitope (Supplemental Fig. S5A). As predicted, incubation with dTAG-13 rapidly lowered FKBP12^{F36V} fusion protein fluorescence (Supplemental Fig. S5B). Western Blot validated these findings and showed near-complete loss of expression in response to incubation with dTAG-13 (Supplemental Fig. S5C,D). Incubation of wild-type ECOMG cells or neutrophils differentiated from ECOMG cells with dTAG-13 did not perturb cell viability, growth, or differentiation for up to 5–7 d of treatment (data not shown). Fusion with FKBP12^{F36V} of targeted factors did not lead to constitutive, basal levels of degradation (data not shown). Basal protein abundance was equivalent in FKBP12^{F36V} targeted cells when compared with parental lines cultured in the absence of dTAG-13 (data not shown). Additionally, targeted cells divided at similar rates and appeared indistinguishable from the parental cell lines cultured in the absence of dTAG-13. Thus, FKBP12^{F36V} tagging does not abrogate essential neutrophil functions but permits, upon stimulation with dTAG-13, rapid and near complete depletion of targeted proteins of interest.

Histone acetyltransferases as well as transcription elongation are essential to maintain global H3K27ac abundance

The transcription coactivators P300 and CBP are highly homologous acetyltransferases that establish the H3K27ac and H3K18ac landscape (Jin et al. 2011; Raisner et al. 2018). As a first approach, to determine whether P300 and CBP directly regulate H3K27ac abundance, ECOMG cells harboring P300-FKBP12^{F36V} or CBP-FKBP12^{F36V} were cultured in the presence of dTAG-13. We found that neither P300 nor CBP were essential for viability, proliferation, and differentiation of ECOMG cells as well as maintenance of H3K27ac abundance (Supplemental Table S6; data not shown). These data raised the possibility that P300 and CBP act redundantly to maintain H3K27ac abundance. To examine this possibility, we generated ECOMG cells that harbored the FKBP12^{F36V} degron in both the P300 and CBP loci. The resulting cell line is referred to as "P300&CBP degron." FACS analysis showed that P300 protein levels remained equivalent whereas CBP protein abundance decreased during neutrophil differentiation (Supplemental Fig. S5B). Simultaneous degradation of P300 and CBP was tolerated for 6 h without obvious effects on cell viability (Supplemental Table S6). Culturing P300&CBP targeted cells for longer time periods (18 h), however, resulted in extensive cell death (Supplemental Table S6). We found that in resting and activated neutrophils within 2-h depletion of both P300 and CBP resulted in a global loss of H3K27ac abundance, while H3K27me3 levels remained unaffected (Fig. 5A). Thus, P300 and CBP maintain H3K27ac abundance across the enhancer landscape.

The chromatin remodeler BAF complexes are composed of a core ATPase, BRG1 or BRM, and ~15 associated subunits (Alfert et al. 2019). BRG1 and BRM are >75% identical and display similar activities in vitro (Phelan et al. 1999; Gatchalian et al. 2020). To explore whether BRG1 and BRM regulate H3K27ac abundance and NIPBL-binding directly, we began by degrading them individually. We found that BRG1 and BRM were dispensable for cell viability, proliferation, differentiation, and H3K27ac maintenance (Supplemental Table S6; data not shown). To determine whether they act redundantly we generated ECOMG cells that harbored the FKBP12^{F36V}degron in both the BRG1 and BRM loci. FACS analysis showed that during neutrophil differentiation BRM protein levels remained unaffected whereas BRG1 protein levels decreased (Supplemental Fig. S5B). Depletion of both BRG1 and BRM protein levels was tolerated for 24 h without obvious cell death, but we found that when incubated for 36 h the majority of cells had died (Supplemental Table S6). Near complete degradation of BRG1 and BRM was achieved by 6 h (Fig. 5B). Only prolonged depletion of both BRG1 and BRM modestly decreased H3K27ac levels, immediately prior to massive cell death, indicating that BRG1 and BRM do not directly modulate H3K27ac abundance (Fig. 5B).

To explore whether nascent transcription, plausibly involving lncRNAs or eRNAs, regulates H3K27ac abundance, we tagged the largest subunit of RNAPII, RPB1, as well as SPT5, an essential part of the RNAPII elongation complex with the FKBP12^{F36V}-degron (Zhou et al. 2012; Shetty et al. 2017). We found that during neutrophil differentiation RPB1 and SPT5 protein levels substantially decreased (Supplemental Fig. S5B). Degradation of RPB1 or SPT5 was tolerated for 8 h without obvious cell death whereas by 18 h the majority of cells had died (Supplemental Table S6). Near-complete degradation of RBP1 and SPT5 was achieved by 4 h. We found that depletion of either RPB1 or SPT5 resulted in global loss of H3K27ac abundance indicating that transcriptional elongation is essential to maintain H3K27ac abundance (Fig. 5D,E).

To determine whether NIPBL and MAU2 expression is essential to maintain H3K27ac abundance, differentiated ECOMG cells expressing NIPBL-FKBP12^{F36V}-EYFP or MAU2-FKBP12^{F36V}-mScarlet were cultured in the

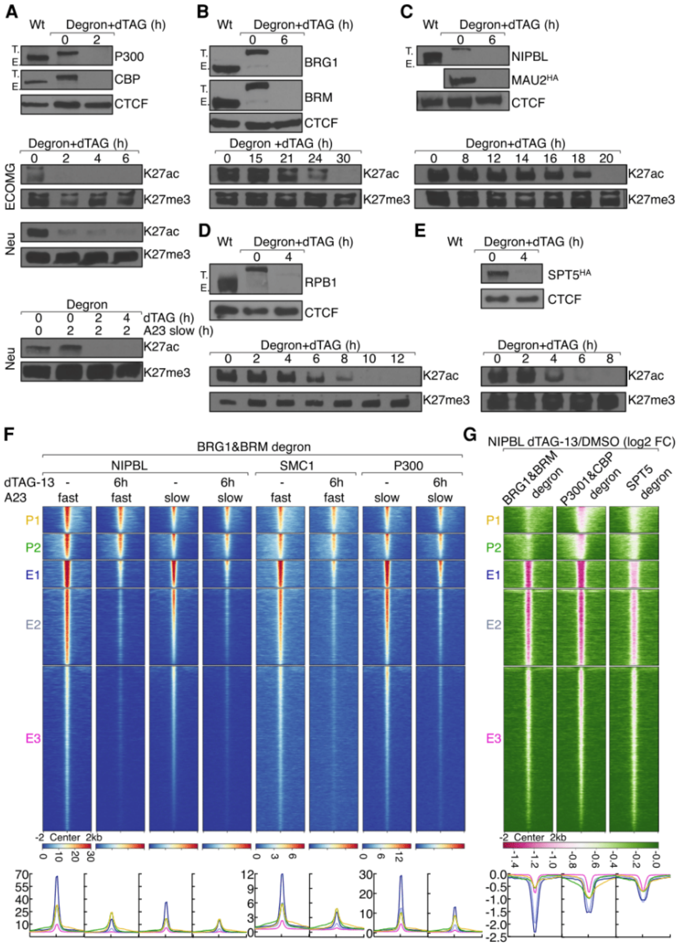


Figure 5. Factors that directly orchestrate H3K27ac abundance and NIPBL occupancy. (*A*) The *top* panel shows Western blot validating dTAG-13 mediated depletion of P300 and CBP abundance in neutrophils expressing "P300&CBP degrons." CTCF was probed to serve as a loading control. Note that successful biallelic tagging is shown by the higher-molecular-weight band (T.) and the lack of native-sized protein (E.) in CRISPR-modified cells. The *bottom* panels indicate Western blots showing changes in H3K27ac abundance upon dTAG-13-mediated P300 and CBP degradation. H3K27me3 was probed as a loading control. Whole-cell lysates were extracted from ECOMG cells, resting neutrophils, and activated neutrophils, respectively. (*B*) The *top* panel shows Western blot validating dTAG-13 mediated depletion of BRG1 and BRM abundance in neutrophils expressing "BRG1&BRM degrons." Antibodies are indicated to the *right*. CTCF was probed to serve as a loading control. The *bottom* panel indicates Western blot showing showing changes in H3K27ac abundance upon dTAG-13-mediated P300 and CBP degradation. H3K27me3 was probed to serve as a loading control. Whole cell lysates were extracted from resting neutrophils. (*C*) As in *B* for NIPBL and MAU2 in "NIPBL&MAU2 degron" neutrophils. Note that MAU2 was blotted with anti-HA antibody. (*D*) As in *B* for RPB1 in "RPB1 degron" neutrophils. (*E*) As in *B* for SPT5 in "SPT5 degron" neutrophils. Note that SPT5 was blotted with anti-HA antibody. (*F*) Heat map of NIPBL, SMC1, and P300 ChIP-seq signals in activated neutrophils in the presence or absence of dTAG-13, plotted for five clusters of genomic regions defined in Figure 2A. Heat map was generated by gating on the maximum signal intensity of NIPBL occupancy using A23187 slow activation conditions. *Below* the heat map corresponding average enrichment plots are shown. Near-complete degradation was achieved by 6 h. Therefore, 6 h of dTAG-13 treatment was chosen to study the immediate consequences for loss of BRG1&BRM expression while minimizing indirect effects on transcription and genome organization caused by cell death. (*G*) Heat map of $\log_2$ FC of NIPBL ChIP-seq signals in dTAG-13 versus DMSO-treated activated neutrophils plotted for five clusters of regions defined in Figure 2A. Heat map was sorted by the minimum signal intensity of $\log_2$ FC of NIPBL in "P300&CBP degron" cells. Pink color indicates signal loss upon dTAG-13 treatment. *Below* the heat map is indicated the corresponding average enrichment plots with *Y*-axis values indicated as log2FC. *Below* the heat map are corresponding average enrichment plots. "BRG1&BRM degron," "P300&CBP degron," and "SPT5 degron" neutrophils were pretreated with dTAG-13 for 6, 2, and 4 h before activation, respectively.

presence of dTAG-13. We found that neither NIBPL nor MAU2 expression alone was required for ECOMG cell and neutrophil viability, proliferation, differentiation, and H3K27ac maintenance (Supplemental Table S6). To determine whether they act redundantly, we generated ECOMG cells that express both NIPBL-FKBP12^{F36V} and MAU2-FKBP12^{F36V}. Degradation of both NIPBL and MAU2 was tolerated for 18 h without obvious cell death, but by 24 h the majority of cells had died (Supplemental Table S6). Near-complete depletion of NIPBL and MAU2 abundance was achieved by 6 h of culturing ECOMG cells in dTAG-13. Only prolonged NIPBL and MAU2 depletion (18–20 h) showed decreased H3K27ac levels indicating that as expected NIPBL and MAU2 do not modulate H3K27ac abundance directly (Fig. 5C). Taken together, these data indicate that P300 and CBP as well as transcriptional elongation maintain H3K27ac abundance.

NIPBL occupancy at enhancers vs. promoters is differentially regulated by chromatin remodelers, histone acetyltransferases, and transcriptional elongation

The data described above indicate that calcium-induced recruitment of BRG1 and NIPBL at promoters occurs with relatively slow kinetics (3–4 h) when compared with enhancers. To determine whether the differences in NIPBL occupancy at enhancers reflect differences in occupancy of chromatin remodelers "BRG1&BRM degron" neutrophils were activated and cultured in the absence or presence of dTAG-13 (Fig. 5F). We found that NIPBL occupancy at enhancers was severely perturbed upon depleting BRG1 and BRM expression (Fig. 5F; Supplemental Fig. S6A). In contrast, NIPBL occupancy at promoter regions was only marginally affected upon eliminating BRG1 and BRM expression (Fig. 5F). Likewise, while depleting BRG1 and BRM levels greatly reduced SMC1 occupancy at active enhancers SMC1 occupancy at promoters was only marginally affected (Fig. 5F; Supplemental Fig. S6A). To determine whether P300 occupancy is regulated by BRG1 and BRM we examined "BRG1&BRM degron" neutrophils for P300 occupancy. We found that activated neutrophils depleted for BRG1 and BRM expression showed significant lower levels of P300 occupancy at enhancers but not at promoters (Fig. 5F; Supplemental Fig. S6A).

To determine whether P300 and CBP expression facilitates calcium-induced NIPBL occupancy, activated "P300&CBP degron" neutrophils were cultured in the absence or presence of dTAG-13 and examined for NIPBL occupancy (Fig. 5G). We found that calcium-induced NIPBL occupancy was substantially lowered at active enhancers in cells depleted for P300 and CBP (Fig. 5G; Supplemental Fig. S6A). Conversely, NIPBL occupancy at promoters was only marginally affected by depleting P300 and CBP expression (Fig. 5G; Supplemental Fig. S6A). To determine whether transcriptional elongation instructs calcium-induced NIPBL occupancy, activated neutrophils depleted for SPT5 expression were examined for NIPBL occupancy. We found that calcium-induced NIPBL occupancy was substantially lowered at active enhancers in cells depleted for SPT5 while NIPBL occupancy at promoters was only slightly affected (Fig. 5G; Supplemental Fig. S6A). Taken together, these observations indicate that chromatin remodelers, histone acetyltransferases and transcriptional elongation, facilitate NIPBL and SMC1 occupancy at enhancers rather than at promoters.

Neutrophil differentiation is closely associated with a decline in loop extrusion

Our previous observations as well as the data described here indicate that neutrophil differentiation is accompanied by large-scale changes in nuclear architecture (Zhu et al. 2017; Denholtz et al. 2020). These findings raised the question as to whether these alterations in genome topology are established by changes in the expression of genes encoding for NIPBL, MAU2, CTCF, and RAD21. To address this question, we probed the expression of NIPBL, MAU2, CTCF, and RAD21 using ECOMG cell lines that carry fluorescently labeled degrons. Briefly, ECOMG cells expressing NIPBL-FKBP12^{F36V}-EYFP, MAU2-FKBP12^{F36V}-mScarlet, CTCF-FKBP12^{F36V}-EYFP, and RAD21-FKBP12-F36V-EYFP were differentiated and analyzed for EYFP or mScarlet expression. Interestingly, we found that NIPBL, CTCF, MAU2, and RAD21 expression all substantially declined in differentiating neutrophils (Supplemental Fig. S5B). Taken together, these data suggest that the large-scale alterations in nuclear architecture associated with neutrophil differentiation is dictated by alterations in the expression of genes encoding for factors that instruct loop extrusion.

Calcium-induced transcription is instructed by RAD21, chromatin remodelers, and histone acetyltransferases

To determine whether CTCF, RAD21, and WAPL, instruct calcium-induced NIPBL occupancy, activated neutrophils carrying degrons for these factors were cultured in the absence or presence of dTAG-13 and examined for NIPBL occupancy. As predicted, calcium-induced NIPBL occupancy was not lowered at active enhancers and promoters in cells depleted for CTCF, RAD21, and WAPL (unpublished observations). To determine whether and how depletion of RAD21, BRG1, BRM, P300, and CBP modulate calcium-induced transcription, activated neutrophils depleted for these factors were analyzed for RPB1 occupancy. We found that upon calcium entry depletion of RAD21, P300, and CBP as well as BRG1 and BRM lowered RBP1 occupancy (Supplemental Fig. S6B). Taken together, these data indicate that in activated neutrophils calcium-induced transcription is instructed by histone acetyltransferases, chromatin remodelers and the cohesin machinery.

Calcium signaling activates NFAT to promote NIPBL occupancy at enhancers

To identify potential TFs that promote calcium-signaling induced NIPBL occupancy, we examined induced NIPBL-binding sites for enriched DNA sequences. A progressive enrichment of TF motifs across the activation time course was identified. At 1 h after activation, induced NIPBL-binding sites were enriched for *cis* elements associated with CEBP and NFAT occupancy (Fig. 6A; Supplemental Fig. S7). Calcium signaling for a duration of 2 h doubled the number of NIPBL-binding sites that were enriched for DNA sequences associated with NFAT, CEBP, ETS, AP-1, and RUNX occupancy (Fig. 6A; Supplemental Fig. S7). The number of NIPBL-bound sites increased 3 h after activation and were enriched for *cis* elements associated with bHLH, SP/KLF, NRF, and NFY occupancy (Fig. 6A;

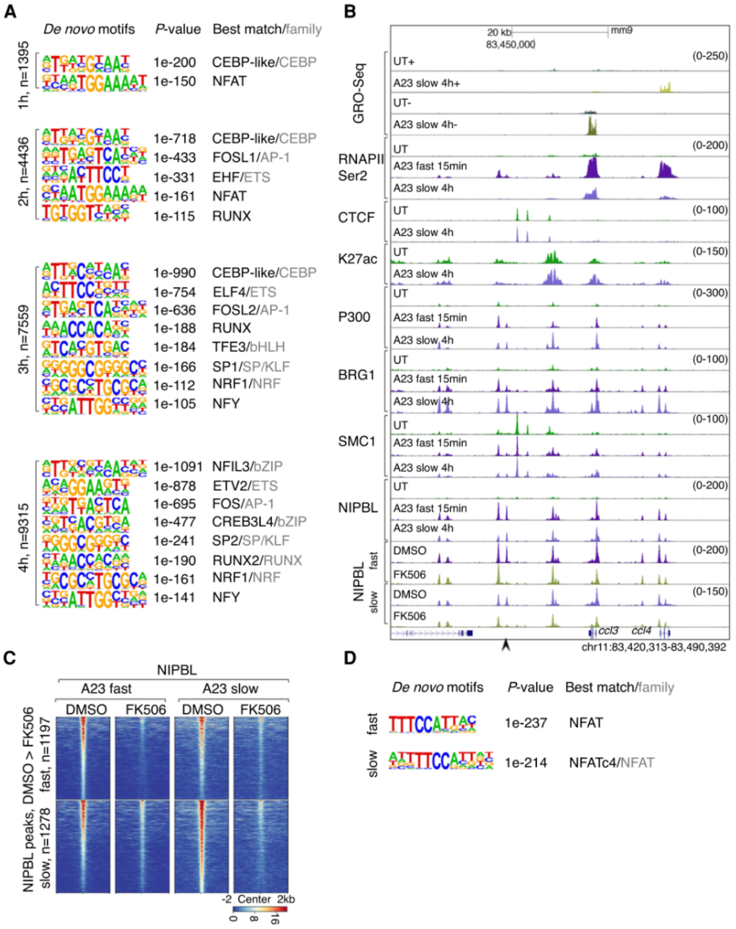


Figure 6. Calcium signaling activates NFAT to promote NIPBL occupancy. (*A*) Motifs enriched in the vicinity of NIPBL-binding sites (± 200 bp) that are up-regulated in activated neutrophils at respective time points using GC-matched genomic backgrounds. Note that all members of a TF family tended to show similar enrichment due to motif similarity. *P*-value for motif enrichment, the best-matched TF and TF family if applicable, are indicated at the *right* of the sequence logos (cutoff: $P < 1 \times 10^{-100}$). (*B*) Genome browser tracks depicting GRO-seq and ChIP-seq reads in untreated and activated neutrophils, cultured in the presence or absence of FK506. A representative locus *mmp8* is shown. Black arrowheads indicate NIPBL-bound sites that were down-regulated in FK506-treated neutrophils. (*C*) Heat maps of NIPBL occupancy in activated neutrophils cultured in the absence or presence of FK506, using a window of ± 2 kb centered on NIPBL bound sites. Heat maps were gated on down-regulated NIPBL-bound sites identified in FK506-treated versus DMSO-treated activated neutrophils. Corresponding average enrichment plots are shown *below*. (*D*) Motifs enriched in the vicinity of NIPBL-bound sites (±200 bp) that were depleted in FK506-treated versus DMSO-treated activated neutrophils. A GC-matched genomic background was used for the analysis. *P*-value for motif enrichment, the best-matched TF and TF family if applicable are indicated at the *right* of the sequence logos (cutoff: $P < 1 \times 10^{-100}$).

Supplemental Fig. S7). The number of NIPBL-bound sites continued to increase 4 h after activation, but the composition of TF binding motifs essentially remained the same (Fig. 6A; Supplemental Fig. S7). Notably, however, *cis* elements associated with calcium-induced NIPBL-occupancy at promoters were distinct from those associated with enhancers. Specifically, we found that calcium-induced NIPBL occupancy at promoters was enriched for GC-rich DNA sequences (SP/KLF) and appeared with different kinetics following calcium entry when compared with enhancers (3-4 h) (Supplemental Fig. S7).

To validate these findings, we tested whether NFAT binding sites were associated with calcium-induced NIPBL occupancy. Specifically, activated neutrophils were cultured in the absence or presence of FK506, a well-characterized NFAT inhibitor and examined for NIPBL binding. As predicted, inhibition of NFAT activity interfered with calcium-induced NIPBL occupancy (Fig. 6B,C). Specifically, pretreatment with FK506 interfered with calcium-induced NIPBL occupancy at 1197 sites (fast activation) and at 1278 sites (slow activation) (Fig. 6C). Consistent with these results, de novo motif analysis revealed that calcium-responsive FK506-sensitive NIPBL-binding sites were predominantly enriched for NFAT motifs (Fig. 6D). Taken together, these data indicate that upon calcium influx NIPBL is initially recruited at enhancers that are enriched for NFAT, PU.1, CEBP, and AP-1 binding sites while at later stages NIPBL is sequestered at promoters characterized by GC-rich DNA sequences.

PU.1 modulates calcium-induced NIPBL occupancy across the chromatin landscape

It is well established that PU.1 is a pioneering factor that orchestrates neutrophil development and activation (Anderson et al. 1998; Iwasaki et al. 2005; Fischer et al. 2019; Minderjahn et al. 2020). The data described above raised the possibility that PU.1 as a pioneering factor acts to promote NIPBL occupancy. To determine whether, indeed, PU.1 expression is essential to instruct NIPBL occupancy we generated PU.1-knockout (KO) ECOMG-derived neutrophils. ECOMG-derived PU.1 wild-type (WT) and knock-out (KO) neutrophils were stimulated with A23187 and examined for calcium-induced NIPBL occupancy. We found that depletion of PU.1 globally modulated NIPBL binding (Supplemental Fig. S8A,B). Specifically, PU.1 depletion resulted in a loss of 4780 NIPBL-bound sites and a gain of 11434 NIPBL-bound sites (Supplemental Fig. S8B). Likewise, we found that PU.1 depletion coordinately changed SMC1 occupancy (Supplemental Fig. S8B). To determine whether upon depletion of PU.1 expression loss or gain of NIPBL occupancy were differentially associated with *cis* elements we performed a de novo motif analysis. We found that upon depleting PU.1 abundance, calcium-induced loss of NIPBL binding sites were enriched for DNA sequences associated with PU.1 binding, while calcium-induced gain of NIPBL binding sites were enriched for DNA sequences associated with CEBP binding, plausibly involving a competitive displacement mechanism (Cloutier et al. 2009; Hosokawa et al. 2018). Taken together, these data imply that PU.1 uses distinct mechanisms to direct NIPBL occupancy across the chromatin landscape.

Calcium signaling instructs lineage-specific NIPBL occupancy in primary immune cells

The data described above indicate that calcium signaling induces NIPBL recruitment to active enhancers and promoters in neutrophils. However, these experiments were performed using in vitro differentiated neutrophils that were derived from ECOMG cells. To validate these findings for primary cells, neutrophils were isolated from the bone marrow and activated with A23187. We found that calcium influx in primary bone marrow-derived neutrophils (BMDN) rapidly modulated NIPBL-binding with the majority of binding sites showing increased NIPBL occupancy (Fig. 7A). We next examined NIPBL-binding in A23187-activated primary T and B cells isolated from the spleen. We found that calcium influx in primary T cells revealed equal ratios of elevated and declining levels of NIPBL occupancy (Fig. 7A). Calcium signaling in primary B cells showed a reverse pattern with a substantial fraction of NIPBL-bound sites being associated with decreased occupancy (Fig. 7A). Calcium-induced NIPBL bound sites showed substantial overlap between activated B and T lineage cells (Fig. 7B). In contrast, we found that only a minor fraction of calcium-induced NIPBL-binding sites overlapped between BMDN versus T and B cells (Fig. 7B). Calcium-induced NIPBL-binding sites in BMDN cells were enriched for a wide spectrum of *ci*s elements, resembling that of ECOMG-derived neutrophils, whereas T and B cells were predominantly enriched for NFAT binding sites (Fig. 7C). In line with these results, we found FK506 severely impaired enrichment of NIPBL occupancy for a cluster of genes encoding for chemokines as well as a genomic region spanning the c-myc and pvt1 loci (Fig. 7D). To determine whether activation of NFAT promotes chromatin remodeling at calcium-induced NIPBL-bound sites we performed ATAC-seq. We found that activated-B cells were associated with altered chromatin accessibility at calcium-induced NIPBL bound sites as well as calcium-repressed NIPBL-bound sites (Fig. 7E). Notably, we found that NFAT inhibition by FK506 pretreatment abolished the increase in chromatin accessibility but only modestly reversed the decrease in chromatin accessibility (Fig. 7E). Finally, de novo motif analysis revealed that FK506-downregulated ATAC-seq reads were enriched for NFAT cis elements suggesting that NFAT binds these sites directly to promote chromatin accessibility and to facilitate NIPBL occupancy (Fig. 7F). Collectively, these data indicate that calcium signaling orchestrates a lineage-specific pattern of NIPBL occupancy in primary immune cells and that NFAT plays a key role in modulating chromatin accessibility at NIPBL bound sites.

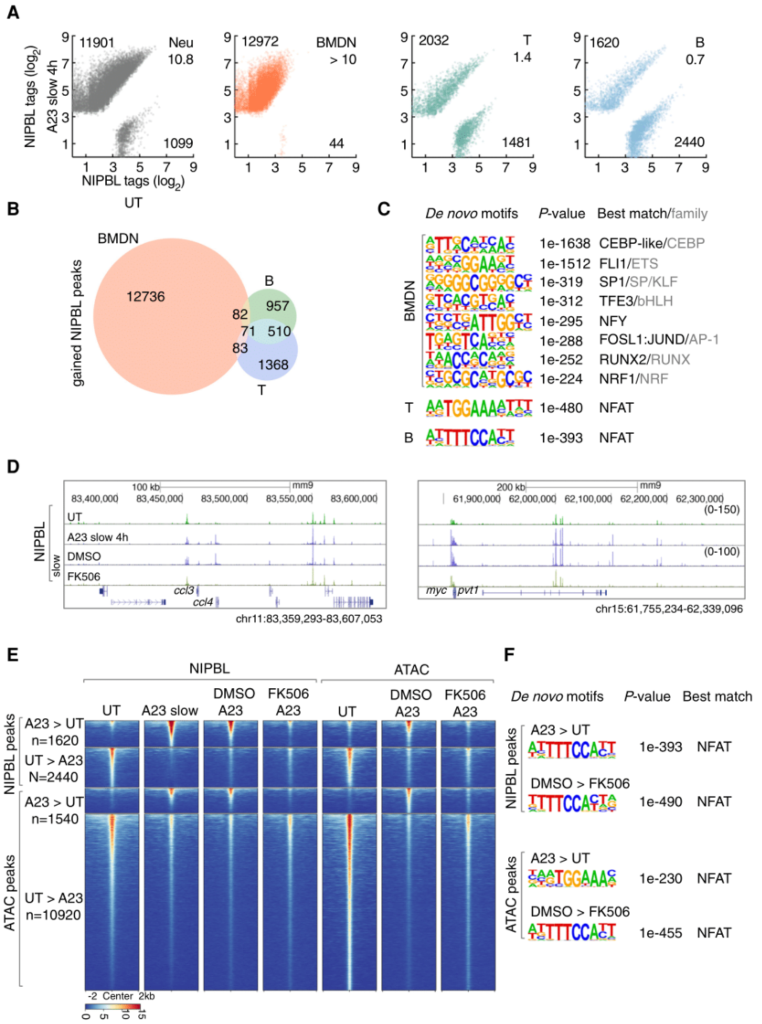


Figure 7. Calcium influx in primary immune cells modulates NIPBL occupancy. (*A*) Scatter plot for $\log_2$ tag counts for NIPBL peaks in untreated versus activated ECOMG-derived neutrophils, bone marrow-derived neutrophils (BMDN), primary T and B cells, respectively, assessed for a 200-bp region centered on differential NIPBL-bound sites. Numbers of gained and lost NIPBL-bound sites (FC > 4) and ratios between them are indicated. (*B*) Venn diagram showing overlap and specificity of gained NIPBL-binding sites in activated BMDN, primary T cells and B cells. Peaks were considered overlapping if the distance separating peak centers were <100 bp. Numbers of specific and overlapping bound sites are indicated for each group. (*C*) Motifs enriched in the vicinity of NIPBL-binding sites (± 200 bp) that were gained in activated BMDN, primary T cells and B cells using GC-matched genomic backgrounds, respectively. *P*-values for motif enrichment, the best-matched TF and TF family if applicable are indicated at the *right* of the sequence logos (cutoff: $P < 1\times10^{-100}$). (*D*) Genome browser tracks of NIPBL-bound sites derived from untreated and activated primary B cells cultured in the presence or absence of FK506 at representative loci, *ccl4* and *myc*. (*E*) Heat map of NIPBL-bound sites and ATAC-seq signals in untreated and activated primary B cells cultured in the presence or absence of FK506, across genomic regions that were centered (±2 kb) on NIPBL-bound sites. (*F*) The *top* panels indicate motifs enriched within the vicinity of NIPBL-bound sites (±200 bp) that are gained in activated primary B cells using GC-matched genomic backgrounds. The *bottom* panels show motifs enriched in the vicinity of accessible chromatin regions defined by ATAC-seq (±200 bp) that were lost in activated primary B cells cultured in the presence of FK506 using GC-matched genomic backgrounds. *P*-values for motif enrichment, the best-matched TF and TF family if applicable, are indicated to the right of the sequence logos (cutoff: $P < 1 \times 10^{-100}$).

Discussion

Previous studies demonstrated that the transition from progenitors to differentiated neutrophils, in mice and humans, is closely associated with large-scale changes in genome topology (Zhu et al. 2017; Denholtz et al. 2020). These changes involve a distinct pattern of remote genomic interactions that span entire chromosomes and a uniquely fragmented compartment structure (Zhu et al. 2017; Denholtz et al. 2020). During neutrophil differentiation how are these alterations in long-range genomic interactions and compartmentalization established? We suggest that they are instructed by a decrease in loop extrusion. This notion is supported by our observation that RAD21, NIPBL, and MAU2 expression levels readily decline during the transition from progenitors to terminally differentiated neutrophils. These findings raise the question of whether and how the decrease in NIPBL, MAU2, and RAD21 expression in differentiating neutrophils is linked to the alterations in nuclear architecture. The decline in NIPBL, MAU2, and RAD21 expression during neutrophil differentiation instructs hyper-fragmented compartmentalization that mimics the increase in compartment segregation observed in NIPBL-deficient hepatocytes (Schwarzer et al. 2017). Thus, we suggest that the decline in NIPBL, MAU2, and RAD21 expression during neutrophil differentiation instructs a distinct chromatin folding pattern that characterizes, at least in part, the terminally differentiated hyper-fragmented neutrophil compartment structure.

While resting neutrophils are associated with fragmented compartmentalization, upon calcium-mediated activation the neutrophil genome adopts a nuclear architecture that resembles that of other immune cells, essentially reversing hyper-fragmented compartmentalization. We found that these changes in nuclear architecture are associated with a global and rapid increase in NIPBL and SMC1 occupancy. Thus, upon calcium entry elevated levels of NIPBL dictate a nuclear architecture that is akin to that of other immune cells in terms of compartmentalization. This then raises the question why hyper-fragmented compartments are established in terminally differentiated neutrophils? We suggest that prior to activation a highly fragmented compartment structure enforces efficient insulation of distant regulatory elements to prevent stochastic activation of a neutrophil specific gene program. Conversely, upon calcium entry loss of highly segregated compartmentalization may facilitate local genomic interactions to bring into close spatial proximity enhancer and promoter elements allowing the expression of genes associated with an activated neutrophil specific gene program. Consistent with such a scenario are our observations that depletion of RAD21 in differentiated neutrophils altered calcium-induced RNA polymerase II occupancy. We note that gene expression modulated by the cohesin machinery is not unique to activated neutrophils. Recent studies indicated that cohesin expression is required to induce the expression of an inflammatory gene program in macrophages (Cuartero et al. 2018). Thus, loop extrusion-induced alterations in nuclear architecture might be a shared mechanism that is used by macrophages and neutrophils to orchestrate inducible programs of gene expression.

How does calcium signaling recruits NIPBL occupancy with such great speed and precision across the chromatin landscape? Calcium-induced NIPBL occupancy at enhancers involves a wide spectrum of TFs. Prominent amongst these are TFs that regulate gene expression in response to a spectrum of stimuli, including NFAT and AP1 as well as TFs that orchestrate lineage-specific gene programs such as PU.1 and CEBP. The second step involves TF-dependent recruitment of the histone acetyltransferases P300 and CBP. Once recruited P300 and CBP establish a calcium-induced specific H3K27ac landscape. Likewise, the chromatin remodelers BRG1 and BRM are sequestered at NIPBL-bound sites. Whether the recruitment of P300 and CBP as well as BRG1 and BRM engage sequential or parallel pathways or pathways that involve complex feedback mechanisms, remains to be determined, but it is evident that they are essential to sequester NIPBL to its target sites. In addition to histone acetyltransferases and chromatin remodelers we found that transcriptional elongation plays a role as well in elevating H3K27 abundance and targeting NIPBL to enhancers. How may transcriptional elongation modulate H3K27ac levels and facilitate NIPBL occupancy? We suggest that noncoding transcripts, lncRNAs and/or eRNAs, function to modulate the catalytic activity of P300 and CBP. Finally, it is conceivable that noncoding transcription facilitates the assembly of hubs or condensates to facilitate NIPBL occupancy.

While depletion of BRG1 and BRM severely impaired calcium-induced NIPBL occupancy at enhancers, NIPBL binding at promoters was only modestly affected. How does recruitment of NIPBL to promoter regions differ from that to enhancers? We found that distinct *cis* elements are associated with NIPBL occupancy at promoters versus enhancers. Induced NIPBL-binding sites at promoters are enriched for DNA sequences associated with SP/KLF binding sites (GC-rich) rather than NFAT, AP-1, PU.1, and CEBP motifs that are associated with enhancers. The enrichment for SP/KLF *cis* elements is particularly intriguing since SP1 consensus DNA sequences have been shown to be associated with noncanonical secondary DNA structures such as G-quadruplexes (G4) (Raiber et al. 2012). We validated these observations using a bioinformatics approach, namely Pqsfinder. Notably, we found that indeed G4 DNA structures were enriched at induced NIPBL-binding sites (data not shown) (Hon et al. 2017). G4 structural motifs have also been linked to transcription and are enriched at promoters and 5′ UTRs associated with highly transcribed genes (Hänsel-Hertsch et al. 2016, 2018). While chromatin remodelers do not appear to promote NIPBL occupancy at promoters, the histone acetyltransferases P300 and CBP also modulate NIPBL occupancy at promoters, albeit to a much lesser degree. Thus, noncanonical DNA structures across the promoter landscape may permit recruitment of NIPBL in a pathway that is independent of chromatin remodelers but may involve the activities of histone acetyltransferases. Finally, we note that it is conceivable that in order to enable NIPBL occupancy at promoters, GC-rich noncanonical secondary DNA structures may enrich for chromatin accessible regions that resemble chromatin remodeling-induced nucleosome depleted regions at enhancers. In sum, while chromatin remodelers are key to facilitate NIPBL occupancy at enhancers, promoters use noncanonical GC-rich DNA structures to sequester NIPBL. While much remains to be learned, a cardinal conclusion reached from the studies described here is that transcription factors, noncoding transcription, the deposition of H3K27ac, and depletion of nucleosomes are instrumental in sequestering the cohesin machinery across the active enhancer repertoire.

Materials and methods

Cell culture

In vitro differentiated neutrophils from ECOMG cells were used as a model system to study nuclear architecture in response to calcium influx in neutrophils. ECOMG cells were cultured and differentiated as previously described (Zhu et al. 2017). For inhibitors, neutrophils were treated with 10 µM Triptolide (Cayman), 100 µM 5,6-dichloro-1-β-D-ribofuranosyl-1H-benzimidazole (Cayman) or 10 µM flavopiridol (Cayman) for 30 min, or 2 µM FK506 for 2 h before activation. Primary B cells were treated with 2 µM FK506 (Cayman) for 2 h before activation. Inhibitors were kept in culture during activation time course. Acute degradation was induced by treating the cells with 0.5 µM dTAG-13 (Tocris) before activation. dTAG-13 was kept in culture during activation time course. The time series degradation experiments were performed by inducing protein degradation at the beginning of the time course and harvesting the samples at different time points. RPMI-1640 (Gibco) contains 0.42 mM calcium. Additional $CaCl_2$ was supplied to the medium before activation to a final concentration of 1 mM. A23187 was purchased from Sigma and Cayman. Neutrophils were activated either using fast activation conditions (20 µM A23187 for 15 min), or using slow activation conditions (5 µM A23187 for 4 h). These two activation conditions were used throughout the manuscript unless otherwise mentioned.

Data availability

Data sets generated in this study are available as a series in the GEO database under accession number GSE154051.

Acknowledgments

Sequencing was performed at the IGM Genomics Center, University of California at San Diego, La Jolla, CA, supported by grant number P30CA023100. Imaging was performed at the microscopy core of the School of Medicine, University of California at San Diego, La Jolla, CA (CA23187100 and NS047101). This study was supported by grants from the National Institutes of Health to C.M. (AI00880, AI09599, and AI102853).

Author contributions: Y.Z., M.D., and C.M. conceived this project. Y.Z. designed and performed all experiments. H.L. cloned repair template targeting *cbp* gene locus, isolated primary T and B cells from mice and helped with FACS analysis. Y.Z. analyzed all data. H.L. analyzed compartment strength. Y.Z. and C.M. prepared the manuscript with input from other authors. C.M. supervised the study.

References

Alfert A, Moreno N, Kerl K. 2019. The BAF complex in development and disease. *Epigenetics and Chromatin* **12:** 19. doi:10.1186/s13072-019-0264-y

Anderson KL, Smith KA, Pio F, Torbett BE, Maki RA. 1998. Neutrophils deficient in PU.1 do not terminally differentiate or become functionally competent. *Blood* **92:** 1576–1585. doi:10.1182/blood.V92.5.1576

Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, Weinrauch Y, Zychlinsky A. 2004. Neutrophil extracellular traps kill bacteria. *Science* **303:** 1532–1535. doi:10.1126/science.1092385

Busslinger GA, Stocsits RR, Van Der Lelij P, Axelsson E, Tedeschi A, Galjart N, Peters JM. 2017. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544:** 503–507. doi:10.1038/nature22063

Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Shevchenko A, Nasmyth K. 2000. Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol Cell* **5:** 243–254. doi:10.1016/S1097-2765(00)80420-7

Cloutier A, Guindi C, Larivée P, Dubois CM, Amrani A, McDonald PP. 2009. Inflammatory cytokine production by human neutrophils involves C/EBP transcription factors. *J Immunol* **182:** 563–571. doi:10.4049/jimmunol.182.1.563

Cuartero S, Weiss FD, Dharmalingam G, Guo Y, Ing-Simmons E, Masella S, Robles-Rebollo I, Xiao X, Wang YF, Barozzi I, et al. 2018. Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. *Nat Immunol* **19:** 932–941. doi:10.1038/s41590-018-0184-1

Denholtz M, Zhu Y, He Z, Lu H, Isoda T, Döhrmann S, Nizet V, Murre C. 2020. Upon microbial challenge, human neutrophils undergo rapid changes in nuclear architecture and chromatin folding to orchestrate an immediate inflammatory gene program. *Genes Dev* **34:** 149–165. doi:10.1101/gad.333708.119

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485:** 376–380. doi:10.1038/nature11082

Fischer J, Walter C, Tönges A, Aleth H, Jordão MJC, Leddin M, Gröning V, Erdmann T, Lenz G, Roth J, et al. 2019. Safeguard function of PU.1 shapes the inflammatory epigenome of neutrophils. *Nat Immunol* **20:** 546–558. doi:10.1038/s41590-019-0343-z

Fric J, Zelante T, Wong AYW, Mertes A, Yu HB, Ricciardi-Castagnoli P. 2012. NFAT control of innate immunity. *Blood* **120:** 1380–1389. doi:10.1182/blood-2012-02-404475

Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. 2016. Formation of chromosomal domains by loop extrusion. *Cell Rep* **15:** 2038–2049. doi:10.1016/j.celrep.2016.04.085

Gatchalian J, Liao J, Maxwell MB, Hargreaves DC. 2020. Control of stimulus-dependent responses in macrophages by SWI/SNF chromatin remodeling complexes. *Trends Immunol* **41:** 126–140. doi:10.1016/j.it.2019.12.002

Gibcus JH, Dekker J. 2013. The hierarchy of the 3D genome. *Mol Cell* **49:** 773–782. doi:10.1016/j.molcel.2013.02.011

Haarhuis JH, Rowland BD. 2017. Cohesin: building loops, but not compartments. *EMBO J* **36:** 3549–3551. doi:10.15252/embj.201798654

Haarhuis JHI, van der Weide RH, Blomen VA, Yáñez-Cuna JO, Amendola M, van Ruiten MS, Krijger PHL, Teunissen H, Medema RH, van Steensel B, et al. 2017. The cohesin release factor WAPL restricts chromatin loop extension. *Cell* **169:** 693–707.e14. doi:10.1016/j.cell.2017.04.013

Hänsel-Hertsch R, Beraldi D, Lensing SV, Marsico G, Zyner K, Parry A, Di Antonio M, Pike J, Kimura H, Narita M, et al. 2016. G-quadruplex structures mark human regulatory chromatin. *Nat Genet* **48:** 1267–1272. doi:10.1038/ng.3662

Hänsel-Hertsch R, Spiegel J, Marsico G, Tannahill D, Balasubramanian S. 2018. Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing. *Nat Protoc* **13:** 551–564. doi:10.1038/nprot.2017.150

Hogan PG, Chen L, Nardone J, Rao A. 2003. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev* **17:** 2205–2232. doi:10.1101/gad.1102703

Hogan PG, Lewis RS, Rao A. 2010. Molecular basis of calcium signaling in lymphocytes: STIM and ORAI. *Annu Rev Immunol* **28:** 491–533. doi:10.1146/annurev.immunol.021908.132550

Hon J, Martínek T, Zendulka J, Lexa M. 2017. Pqsfinder: an exhaustive and imperfection-tolerant search tool for potential quadruplex-forming sequences in R. *Bioinformatics* **33:** 3373–3379. doi:10.1093/bioinformatics/btx413

Hosokawa H, Ungerbäck J, Wang X, Matsumoto M, Nakayama KI, Cohen SM, Tanaka T, Rothenberg EV. 2018. Transcription factor PU.1 represses and activates gene expression in early T cells by redirecting partner transcription factor binding. *Immunity* **48:** 1119–1134.e7. doi:10.1016/j.immuni.2018.04.024

Iwasaki H, Somoza C, Shigematsu H, Duprez EA, Iwasaki-Arai J, Mizuno S, Arinobu Y, Geary K, Zhang P, Dayaram T, et al. 2005. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* **106:** 1590–1600. doi:10.1182/blood-2005-03-0860

Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Lee JE, Wang C, Brindle PK, Dent SYR, Ge K. 2011. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J* **30:** 249–262. doi:10.1038/emboj.2010.318

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, Van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467:** 430–435. doi:10.1038/nature09380

Kenny EF, Herzig A, Krüger R, Muth A, Mondal S, Thompson PR, Brinkmann V, von Bernuth H, Zychlinsky A. 2017. Diverse stimuli engage different neutrophil extracellular trap pathways. *Elife* **6:** 1–21. doi:10.7554/eLife.24437

Kueng S, Hegemann B, Peters BH, Lipp JJ, Schleiffer A, Mechtler K, Peters JM. 2006. Wapl controls the dynamic association of cohesin with chromatin. *Cell* **127:** 955–967. doi:10.1016/j.cell.2006.09.040

Ley K, Hoffman HM, Kubes P, Cassatella MA, Zychlinsky A, Hedrick CC, Catz SD. 2018. Neutrophils: new insights and open questions. *Sci Immunol* **3:** eaat4579. doi:10.1126/sciimmunol.aat4579

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326:** 289–293. doi:10.1126/science.1181369

Minderjahn J, Schmidt A, Fuchs A, Schill R, Raithel J, Babina M, Schmidl C, Gebhard C, Schmidhofer S, Mendes K, et al. 2020. Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. *Nat Commun* **11:** 402. doi:10.1038/s41467-019-13960-2

Murayama Y, Uhlmann F. 2014. Biochemical reconstitution of topological DNA binding by the cohesin ring. *Nature* **505:** 367–371. doi:10.1038/nature12867

Nabet B, Roberts JM, Buckley DL, Paulk J, Dastjerdi S, Yang A, Leggett AL, Erb MA, Lawlor MA, Souza A, et al. 2018. The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol* **14:** 431–441. doi:10.1038/s41589-018-0021-8

Nasmyth K, Haering CH. 2009. Cohesin: its roles and mechanisms. *Annu Rev Genet* **43:** 525–558. doi:10.1146/annurev-genet-102108-134233

Nuebler J, Fudenberg G, Imakaev M, Abdennur N, Mirny LA. 2018. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc Natl Acad Sci* **115:** E6697–E6706. doi:10.1073/pnas.1717730115

Parenti I, Diab F, Gil SR, Mulugeta E, Casa V, Berutti R, Brouwer RWW, Dupé V, Eckhold J, Graf E, et al. 2020. MAU2 and NIPBL variants impair the heterodimerization of the cohesin loader subunits and cause Cornelia de Lange syndrome. *Cell Rep* **31:** 107647. doi:10.1016/j.celrep.2020.107647

Petela NJ, Gligoris TG, Metson J, Lee BG, Voulgaris M, Hu B, Kikuchi S, Chapard C, Chen W, Rajendra E, et al. 2018. Scc2 Is a potent activator of cohesin's ATPase that promotes loading by binding Scc1 without Pds5. *Mol Cell* **70:** 1134–1148.e7. doi:10.1016/j.molcel.2018.05.022

Phelan ML, Sif S, Narlikar GJ, Kingston RE. 1999. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol Cell* **3:** 247–253. doi:10.1016/S1097-2765(00)80315-9

Raiber EA, Kranaster R, Lam E, Nikan M, Balasubramanian S. 2012. A non-canonical DNA structure is a binding motif for the transcription factor SP1 in vitro. *Nucleic Acids Res* **40:** 1499–1508. doi:10.1093/nar/gkr882

Raisner R, Kharbanda S, Jin L, Jeng E, Chan E, Merchant M, Haverty PM, Bainer R, Cheung T, Arnott D, et al. 2018. Enhancer activity requires CBP/P300 bromodomain-dependent histone H3K27 acetylation. *Cell Rep* **24:** 1722–1729. doi:10.1016/j.celrep.2018.07.041

Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159:** 1665–1680. doi:10.1016/j.cell.2014.11.021

Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, et al. 2017. Cohesin loss eliminates all loop domains. *Cell* **171:** 305–320.e24. doi:10.1016/j.cell.2017.09.026

Rowley MJ, Corces VG. 2018. Organizational principles of 3D genome architecture. *Nat Rev Genet* **19:** 789–800. doi:10.1038/s41576-018-0060-8

Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J, et al. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci* **112:** E6456–E6465. doi:10.1073/pnas.1518552112

Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, Fonseca NA, Huber W, Haering CH, Mirny L, et al. 2017. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551:** 51–56. doi:10.1038/nature24281

Shetty A, Kallgren SP, Demel C, Maier KC, Spatt D, Alver BH, Cramer P, Park PJ, Winston F. 2017. Spt5 plays vital roles in the control of sense and antisense transcription elongation. *Mol Cell* **66:** 77–88.e5. doi:10.1016/j.molcel.2017.02.023

Tecchio C, Micheletti A, Cassatella MA. 2014. Neutrophil-derived cytokines: facts beyond expression. *Front Immunol* **5:** 1–7. doi:10.3389/fimmu.2014.00508

Vaeth M, Feske S, Macian F. 2018. NFAT control of immune function: new frontiers for an abiding trooper. *F1000Res* **7:** 260. doi:10.12688/f1000research.13426.1

Vian L, Pękowska A, Rao SSP, Kieffer-Kwon KR, Jung S, Baranello L, Huang SC, El Khattabi L, Dose M, Pruett N, et al. 2018. The energetics and physiological impact of cohesin extrusion. *Cell* **173:** 1165–1178.e20. doi:10.1016/j.cell.2018.03.072

Wutz G, Várnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, Schoenfelder S, Jessberger G, Muhar M, Hossain MJ, et al. 2017. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J* **36:** 3573–3599. doi:10.15252/embj.201798004

Zhou Q, Li T, Price DH. 2012. RNA polymerase II elongation control. *Annu Rev Biochem* **81:** 119–143. doi:10.1146/annurev-biochem-052610-095910

Zhu Y, Gong K, Denholtz M, Chandra V, Kamps MP, Alber F, Murre C. 2017. Comprehensive characterization of neutrophil genome topology. *Genes Dev* **31:** 141–153. doi:10.1101/gad.293910.116

Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization

Yina Zhu, Matthew Denholtz, Hanbin Lu, et al.

Genes Dev. published online December 17, 2020
Access the most recent version at doi:10.1101/gad.343475.120

Supplemental Material http://genesdev.cshlp.org/content/suppl/2020/12/15/gad.343475.120.DC1

4.3 Acknowledgements

Chapter 4 reprinted two published paper:

1. Zhu, Y.*, Denholtz, M.*, Lu, H., & Murre, C. (2021). Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization. Genes & Development, 35(65–81).
2. Denholtz, M.*, Zhu, Y.*, He, Z., Lu, H., Isoda, T., Döhrmann, S., Nizet, V., & Murre, C. (2020). Upon microbial challenge, human neutrophils undergo rapid changes in nuclear architecture and chromatin folding to orchestrate an immediate inflammatory gene program. Genes & Development, 34(3-4).

CHAPTER 5

Transcription Factor Network Instructs the Extraembryonic Endoderm Lineages

5.1 Introduction

The mouse blastocyst forming during preimplantation development comprises stem cells of three lineages: the trophectoderm (TE), the epiblast (Epi), and the primitive endoderm (PrE) (Watson, 1992). By the time the blastocyst about to implantation PrE differentiate into the parietal endoderm (PE) and visceral endoderm (VE), two cell fates of the extraembryonic endoderm (ExEn) lineage, localized in anatomically distinct areas in the embryo (Yamanaka et al., 2006). PE and VE remarkably differ in their morphologies and functions as revealed by the studies on postimplantation embryos (Takaoka and Hamada, 2011). Specifically, around embryo day 5.5 (E5.5) after implantation, extraembryonic VE (ExVE) and embryonic VE (EmVE) cells form the epithelial layer which encapsulates the ‘egg cylinder’ entailing the TE-derived extraembryonic ectoderm (ExE) and the Epi cells (Arnold and Robertson, 2009). VE cells function in embryo inductive processes (e.g., communication with Epi cells to setup the embryo axis) as well as nonautonomous cell maintenance (e.g., nutritional support and waste product removal) (Arnold and Robertson, 2009). In contrast, mesenchymal PE cells, scattered along the epithelial TE, secret extracellular matrix (ECM) components that contribute to Reichert’s membrane (RM). RM serves as a barrier to protect the egg cylinder and maintains its structural integrity (Hogan et al., 1980). While the functions of PE and VE were documented in embryo development, the involved pathways are still poorly characterized. Even less is known about how PrE differentiates into PE and VE at E4.5 prior to implantation (Filimonow and Fuente, 2021).

Although genetic knockout studies in mouse models had linked key factors to ExEn development, it was challenged to attribute clear functions due to compounded phenotypes resulting from complex interdependencies among embryo cell types. For instance, both Gata6 and Sox17 knockouts cause PrE defects which carry to PE and VE in mouse embryos (Schrode et al., 2014). However, multiple lines of evidence from in vitro models for PrE differentiation pointed out they are PE determinants. Therefore, the gene regulatory network (GRN) that determines PE versus VE cell fates remains to be elucidated.

Another layer of the complexity of the ExEn GRN is that PE and VE demonstrate extraordinary plasticity. For example, VE cells from postimplantation embryos can be converted to PE when cultured with differentiating ExE cells. Likewise, BMP signaling suppresses the PE cell fate in favor of the VE cell fate for the PE-like XEN cells (Hogan and Tilly, 1981; Paca et al., 2012). Such plasticity indicates nodes, especially transcription factors (TFs) exist in the GRN to connect the transcriptional programs of PE and VE.

Here we identified TFs and biological pathways associated with PE or VE via systematic analysis on the single cell transcriptomes of PrE, PE and VE in E4.5 embryos. Unexpectedly, we found PrE co-expresses PE- and VE-associated TFs. A subset of VE enhancers is poised and co-bound by core PE- and VE-associated TFs in the PE-like chemical-induced XEN (cXEN) cells.

5.2 The Transcriptional Program of PE and VE Cells in E4.5 Embryos

To investigate the transcriptional programs of the early PE and VE cells, we integrated published scRNA-seq datasets of E4.5 embryos to increase single cell numbers that in turn improve cell state detection. We were able to detect TE, Epi, and PrE (ExEn) lineages (Figure 5.1A). Furthermore, the integrated data revealed three subpopulations within the ExEn lineages (Figure 5.1A). By examining the expression of Oct4, Lama1, and Ttr (markers of PrE, PE, and VE), we found the three subpopulations were associated with $\mathrm{Oct4^{hi}Lama1^{med}Ttr^{med}}$, $\mathrm{Oct4^{med}Lama1^{hi}Ttr^{lo}}$, and $\mathrm{Oct4^{lo}Lama1^{lo}Ttr^{hi}}$, indicative of PrE, PE, and VE cell states respectively (Figure 5.1B, 5.2A, and 5.2B). Indeed, cell trajectory analyses reconstructed from E4.5 and E5.5 (postimplantation) embryos confirmed that the PrE subpopulation is the root of the ExEn lineages, whereas the PE and VE subpopulations transits to the bifurcated branches corresponding to E5.5 PE and E5.5 VE (Figure 5.2C).

To identify gene modules that distinguish PE and VE cells at E4.5, we performed high-dimensional weighted gene correlation network analysis (hdWGCNA) to compare the single cell transcriptomes of PE and VE cells. We found two gene modules whose transcriptional profiles were correlated to PE or VE cells (Figure 5.1C). The expression biases of PE and VE modules (538 versus 1124 genes) were sustained in the E5.5 embryo (Figure 5.1C). Gene ontology (GO) analysis confirmed that the PE and VE modules were related to the terms of PE and VE tissue in the mouse anatomical ontology database (Figure 5.2E). Accordingly, the enriched KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways that were activated in the early PE and VE cells agreed with the documented PE and VE biological functions (Figure 5.1E and 5.1F). By constructing the molecular interaction networks of PE and VE modules, we corroborated the VE/PE-associated genes participating in these pathways (Figure 5.1E and 5.1F). Notably, we found PE cells were compatible with self-renewal as they upregulate Myc (a key factor related to pluripotency) and PI3K-Akt signaling pathways (see details below; Figure 5.1E). Taken together, we systematically characterized the distinct transcriptional programs of the early PE versus VE cells in E4.5 embryos.

Figure 5.1 Single cell transcriptomes of the ExEn lineage in E4.5 embryo. (A) UMAP embedding for the integrated scRNA-seq data of E4.5 embryo. (B) Expression profile of the marker genes, Lama1, Ttr, and Pou5f1 for PrE, PE, and VE respectively. (C and D) Violin plot displaying the signature scores of PE (C) and VE (D) gene modules. (E and F) Network connecting the genes and KEGG enriched pathways for PE (E) and VE (F) gene modules.

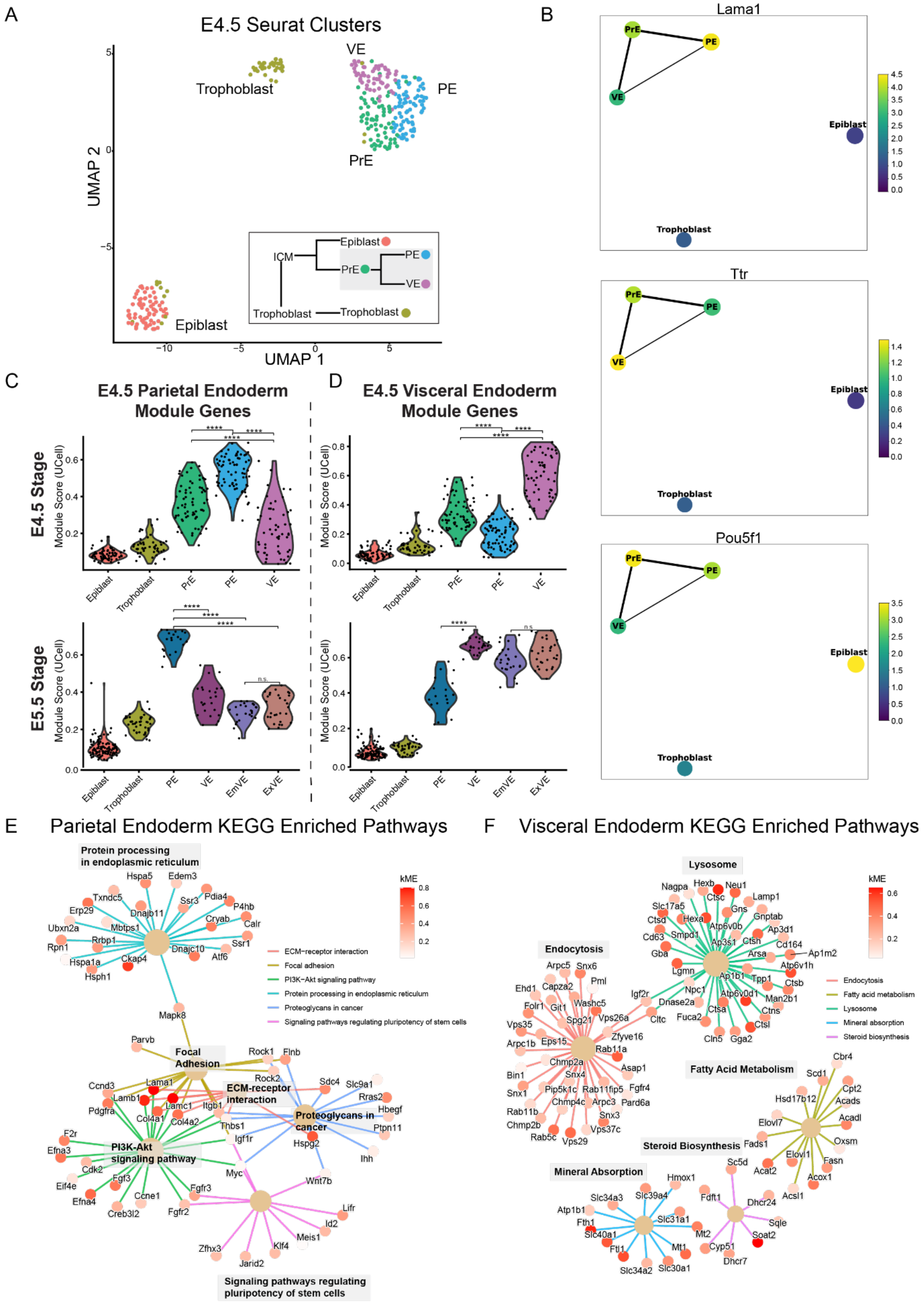
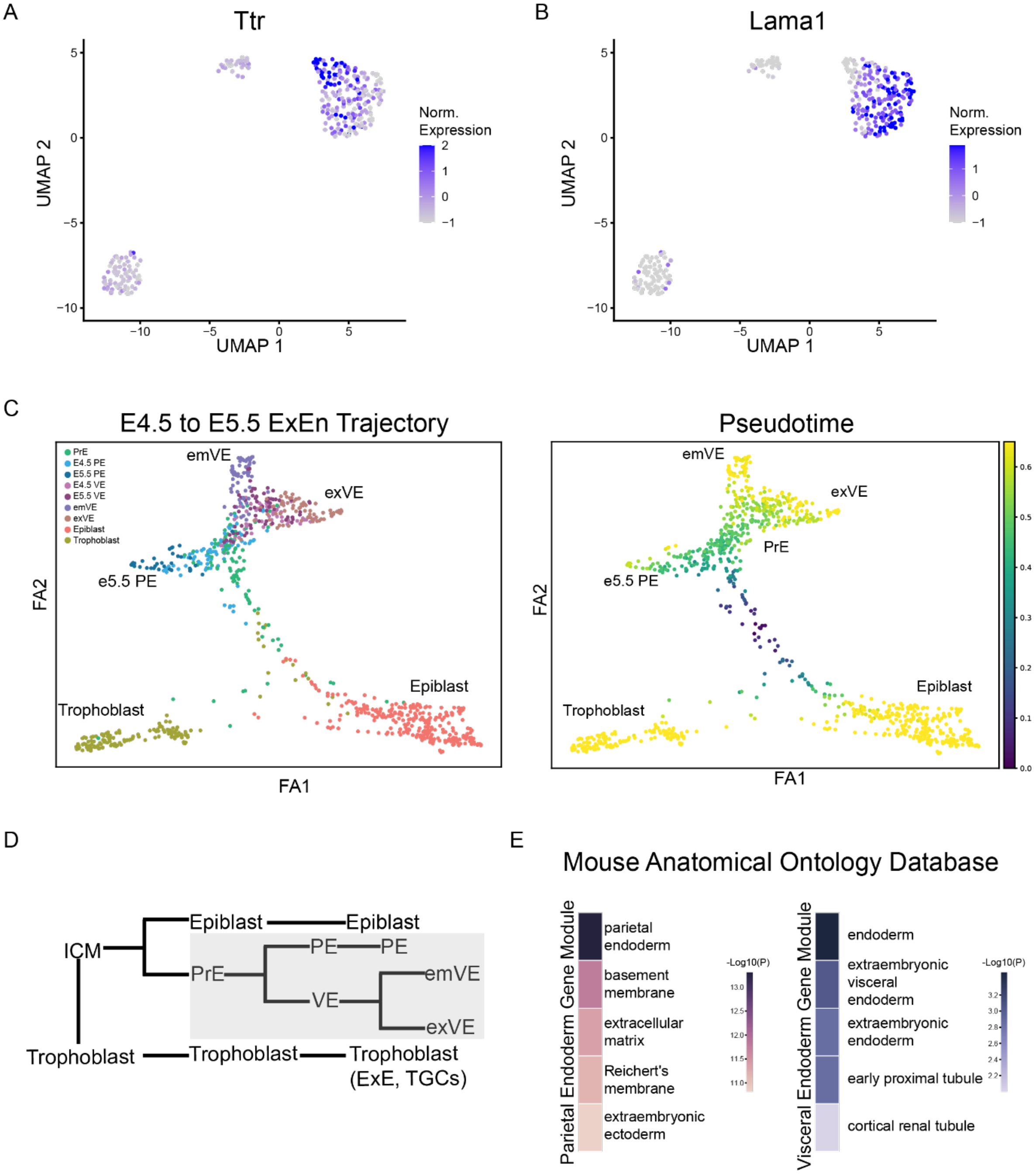

A
E4.5 Seurat Clusters
Trophoblast
VE
PE
PrE
Epiblast
UMAP 2
UMAP 1
ICM
Epiblast
PrE
PE
VE
Trophoblast
Trophoblast
B
Lama1
Ttr
Pou5f1
PrE
PE
VE
Epiblast
Trophoblast
C
E4.5 Parietal Endoderm Module Genes
D
E4.5 Visceral Endoderm Module Genes
E4.5 Stage
E5.5 Stage
Module Score (UCell)
Epiblast
Trophoblast
PrE
PE
VE
EmVE
ExVE
****
n.s
E
Parietal Endoderm KEGG Enriched Pathways
Protein processing in endoplasmic reticulum
Focal Adhesion
ECM-receptor interaction
Proteoglycans in cancer
PI3K-Akt signaling pathway
Signaling pathways regulating pluripotency of stem cells
kME
ECM−receptor interaction
Focal adhesion
PI3K−Akt signaling pathway
Protein processing in endoplasmic reticulum
Proteoglycans in cancer
Signaling pathways regulating pluripotency of stem cells
F
Visceral Endoderm KEGG Enriched Pathways
Lysosome
Endocytosis
Fatty Acid Metabolism
Steroid Biosynthesis
Mineral Absorption
kME
Endocytosis
Fatty acid metabolism
Lysosome
Mineral absorption
Steroid biosynthesis

Figure 5.2 Characterization of cell types of the ExEn lineage in E4.5 embryo. (A-B) Expression of Ttr (A) or Lama1 (B) for the single cells embedded in UMAP space (Figure 5.1A). (C) Force-directed graph embedding E4.5 and E5.5 single cells. Pseudotime was calculated using SCANPY. (D) Diagram displaying the embryo lineage development from E3.0 to E5.5. (E) Enriched anatomical terms of PE and VE gene module in the mouse anatomical ontology database.


A
Ttr
UMAP 2
UMAP 1
5
0
−5
−10
−10
−5
0
5
Norm.
Expression
2
1
0
−1
B
Lama1
UMAP 2
UMAP 1
Norm.
Expression
1
0
−1
C
E4.5 to E5.5 ExEn Trajectory
● PrE
● E4.5 PE
● E5.5 PE
● E4.5 VE
● E5.5 VE
● emVE
● exVE
● Epiblast
● Trophoblast
emVE
exVE
e5.5 PE
Epiblast
Trophoblast
FA2
FA1
Pseudotime
emVE
exVE
PrE
e5.5 PE
Epiblast
Trophoblast
FA2
FA1
0.6
0.5
0.4
0.3
0.2
0.1
0.0
D
ICM
Epiblast
Epiblast
PrE
PE
PE
VE
emVE
exVE
Trophoblast
Trophoblast
Trophoblast
(ExE, TGCs)
E
Mouse Anatomical Ontology Database
Parietal Endoderm Gene Module
parietal endoderm
basement membrane
extracellular matrix
Reichert's membrane
extraembryonic ectoderm
-Log10(P)
13.0
12.5
12.0
11.5
11.0
Visceral Endoderm Gene Module
endoderm
extraembryonic visceral endoderm
extraembryonic endoderm
early proximal tubule
cortical renal tubule
-Log10(P)
3.2
3.0
2.8
2.6
2.4
2.2

5.3 PrE Coexpresses the PE and VE Transcriptional Programs

To investigate the transcriptional state of the PrE, we compared the transcriptional profiles of the PE and VE modules in the E4.5 ExEn cell population. Intriguingly, both gene modules were expressed moderately in PrE cells when compared to PE or VE cells (Figure 5.3A and 5.3B). We then examined the cell-type associated TFs, the potential cell fate determining factors, in PE and VE modules (Figure 5.4A). As expected, PE TFs, including known factors such as Gata6 and Sox17, were preferentially expressed in the PE versus the VE, and vice versa (Figure 5.3C and 5.3D). Strikingly, the majority of PE and VE TFs were already expressed in detectable levels in PrE cells. These findings indicate that PrE cells are in a plastic state in which both PE and VE transcriptional programs are activated.

Figure 5.3 PrE co-expresses the PE and VE transcriptional programs. (A) Scatter plot showing the signature scores of PE and VE modules for the single cells of the ExEn lineage in E4.5 embryo. Marginal distributions were also plotted next to the axis. (B) Heatmap showing the gene expression profiles for the single cells in the ExEn lineage. (C and D) Average gene expression of the top 20 transcription factors in the PE (C) or VE (D) modules for the five cell states in E4.5 embryo. (E) Tornado plot displaying the signal of chromatin accessibilities and depositions of H3K27ac at the combined H3K27ac enriched regions. (F) Motif enrichments at PE and VE enhancers.

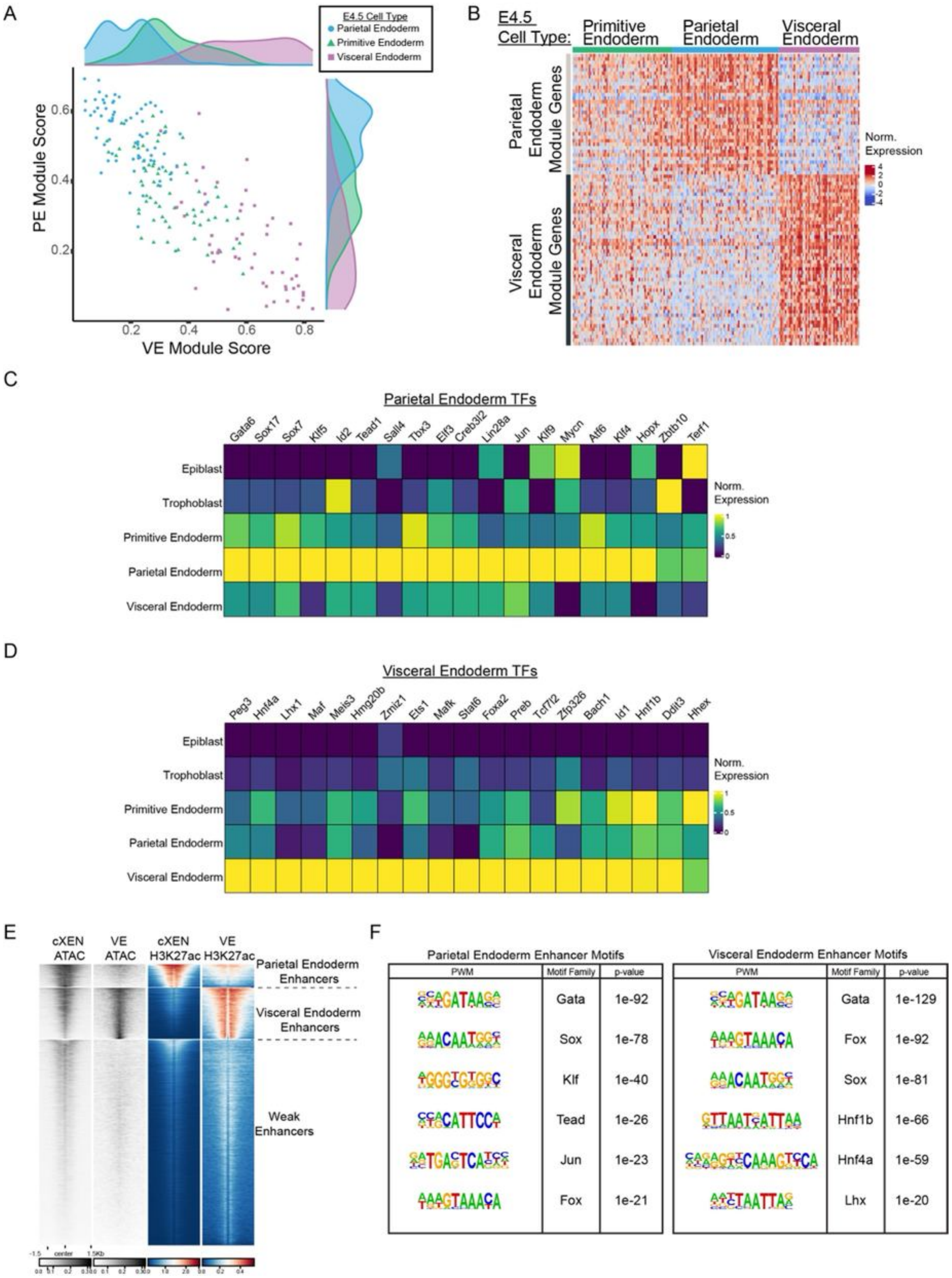
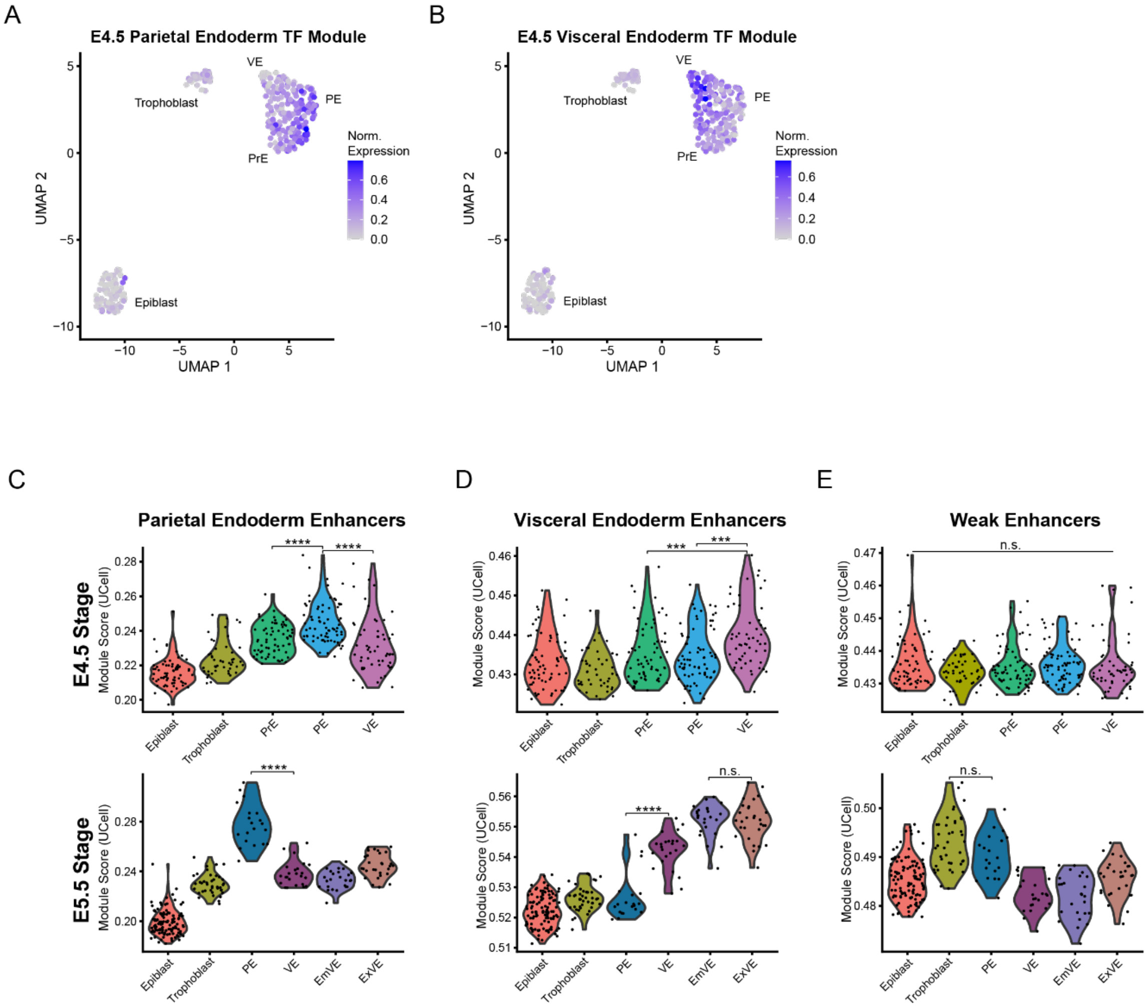
A
E4.5 Cell Type
Parietal Endoderm
Primitive Endoderm
Visceral Endoderm
PE Module Score
0.6
0.4
0.2
0.2
0.4
0.6
0.8
VE Module Score
B
E4.5
Cell Type:
Primitive Endoderm
Parietal Endoderm
Visceral Endoderm
Parietal Endoderm Module Genes
Visceral Endoderm Module Genes
Norm. Expression
4
2
0
-2
-4
C
Parietal Endoderm TFs
Gata6
Sox17
Sox7
Klf5
Id2
Tead1
Sall4
Tbx3
Elf3
Creb3l2
Lin28a
Jun
Klf9
Mycn
Atf6
Klf4
Hopx
Zbtb10
Terf1
Epiblast
Trophoblast
Primitive Endoderm
Parietal Endoderm
Visceral Endoderm
Norm. Expression
1
0.5
0
D
Visceral Endoderm TFs
Peg3
Hnf4a
Lhx1
Maf
Meis3
Hmg20b
Zmiz1
Ets1
Mafk
Stat6
Foxa2
Preb
Tcf7l2
Zfp326
Bach1
Id1
Hnf1b
Ddit3
Hhex
Epiblast
Trophoblast
Primitive Endoderm
Parietal Endoderm
Visceral Endoderm
Norm. Expression
1
0.5
0
E
cXEN ATAC
VE ATAC
cXEN H3K27ac
VE H3K27ac
Parietal Endoderm Enhancers
Visceral Endoderm Enhancers
Weak Enhancers
-1.5
center
1.5Kb
F
Parietal Endoderm Enhancer Motifs
PWM
Motif Family
p-value
Gata
1e-92
Sox
1e-78
Klf
1e-40
Tead
1e-26
Jun
1e-23
Fox
1e-21
Visceral Endoderm Enhancer Motifs
PWM
Motif Family
p-value
Gata
1e-129
Fox
1e-92
Sox
1e-81
Hnf1b
1e-66
Hnf4a
1e-59
Lhx
1e-20

Figure 5.4 Gene sets associated with the PE and VE enhancers. (A and B) Signature scores for the transcription factors in PE (A) or VE (B) modules. (C-E) Signature scores for the gene sets nearby PE (C), VE (D), or weak enhancers (E) as described in Figure 5.3E.


A
E4.5 Parietal Endoderm TF Module
VE
Trophoblast
PE
PrE
Epiblast
Norm. Expression
0.6
0.4
0.2
0.0
UMAP 2
UMAP 1
5
0
−5
−10
−10
−5
0
5
B
E4.5 Visceral Endoderm TF Module
VE
Trophoblast
PE
PrE
Epiblast
Norm. Expression
0.6
0.4
0.2
0.0
UMAP 2
UMAP 1
C
Parietal Endoderm Enhancers
E4.5 Stage
Module Score (UCell)
0.28
0.26
0.24
0.22
0.20
****
****
Epiblast
Trophoblast
PrE
PE
VE
E5.5 Stage
Module Score (UCell)
0.28
0.24
0.20
****
Epiblast
Trophoblast
PE
VE
EmVE
ExVE
D
Visceral Endoderm Enhancers
Module Score (UCell)
0.46
0.45
0.44
0.43
***
***
Epiblast
Trophoblast
PrE
PE
VE
Module Score (UCell)
0.56
0.55
0.54
0.53
0.52
0.51
****
n.s.
Epiblast
Trophoblast
PE
VE
EmVE
ExVE
E
Weak Enhancers
Module Score (UCell)
0.47
0.46
0.45
0.44
0.43
n.s.
Epiblast
Trophoblast
PrE
PE
VE
Module Score (UCell)
0.50
0.49
0.48
n.s.
Epiblast
Trophoblast
PE
VE
EmVE
ExVE

5.4 Coordination of ExEn TFs and Enhancers Orchestrates PE and VE Cell Fates

To understand how cell-type-specific factors instruct the PE or VE cell fate, we first examined the enhancer landscapes of PE-like cXEN cells (in vitro) and E6.5 VE cells as profiled by the chromatin accessibility and the active enhancer mark (H3K27ac). ATAC-seq and H3K27ac Cut&Run (Skene and Henikoff, 2017) were performed on cXEN cells generated in house using the protocol from Niakan et.al (Niakan et al., 2013). The epigenetic data of E6.5 VE were obtained from GSE125318. By clustering the combined enhancers, we segregated the active enhancers for PE and VE respectively (Figure 5.3E). We confirmed that the nearby genes were upregulated in the corresponding cell types in the E4.5 and E5.5 scRNA-seq datasets (Figure 5.4C and 5.4D). The PE enhancers were enriched for the motifs of Gata, Sox, Klf, Tead, and Jun, consistent with the TFs found in the PE module (Figure 5.3C and 5.3F). The VE enhancers were inactive, but accessible, in the PE-like cXEN cells (Figure 5.3E). Surprisingly, they were highly enriched for the PE-associated motifs of Gata and Sox families as well, followed by Fox, Hnf1, Hnf4, and Lhx families whose members (FoxA2, Hnf1b, Hnf4a, and Lhx1) were found in the VE module (Figure 5.3F). The observed motif composition indicates an intricate interplay between PE and VE TFs at the VE enhancers.

To characterize the VE enhancers in detail, we profiled the genome occupancies of Gata6, Sox17, and FoxA2 in PE-like cXEN cells by Cut&Run, since these are the known TFs, belonging to the top motif families in the VE enhancers, with reported functions in the ExEn lineages. We resolved the poised and de novo states in the VE enhancers (Figure 5.5A). Specifically, poised VE enhancers were occupied by the three TFs and had higher chromatin accessibilities compared to that of the de novo VE enhancers (Figure 5.5A, 5.6A, 5.6B and 5.6C). Additionally, depositions of H3K9me3, reported as the counterpart of H3K27me3 for the extraembryonic lineage, were detected in poised VE enhancers (Figure S3A). Furthermore, poised VE enhancers were associated with genes expressed in ExVE (Figure 5.5B) while de novo VE enhancers were associated with general VEs, suggesting that the ExVE cell fate is primed (Figure 5.5C). This observation orthogonally validates the prior study that proposed the ExVE is the default cell fate of VE cells while the EmVE is a specialized, subvariant.

To understand how the TFs modulate the activity of the poised VE enhancers, we exploited TOBIAS software (Bentsen et al., 2020) to detect differential TF footprints from ATAC-seq data of PE-like cXEN cells versus E6.5 VE cells (Figure 5.5D). The footprints of PE-associated motifs were enriched at the poised VE enhancers in cXEN, while the VE-associated Lhx motif has increased footprinting in the VE (Figure 5.5E and 5.6D). However, the major VE motifs, Fox, Hnf1, and Hnf4, were not differentially bound, possibly because the three VE TFs were expressed in PrE and PE, as evidenced in our transcriptomic analysis on the ExEn single cells (Figure 5.4D), cXEN cells (Figure 5.5F), and previous reports. Hence, the differential motif usages at the VE enhancers were correlated with the expression level of the TFs in their associated cell state (Figure 5.5F).

To further dissect the poised state of enhancers, we compared the cutting profiles of the transposase at the active PE enhancers versus the poised VE enhancers in PE-like cXEN cells. We found the cutting frequencies were high across the entire PE enhancers demonstrating that TFs are actively binding on the enhancer platform (Figure 5.6E). Conversely, cut sites were restricted at the center of the poised VE enhancers in the cXEN cells, suggestive of negative regulation at these enhancers.

Figure 5.5 Differential TF footprints at the poised VE enhancers. (A) Tornado plot displaying signals of chromatin accessibilities and chromatin bindings of FoxA2, Gata6, and Sox17 at the VE enhancers in Figure 5.3E. (B and C) Signature scores for the gene sets nearby the poised VE enhancers (B) and de novo VE enhancers (C). (D) Schematic illustrating differential footprinting analysis using TOBIAS. (E) Differences of footprinting score for the motifs when comparing PE versus VE ATAC-seq. (F) Gene expression profiles for TFs belonging to the TF families in (E).

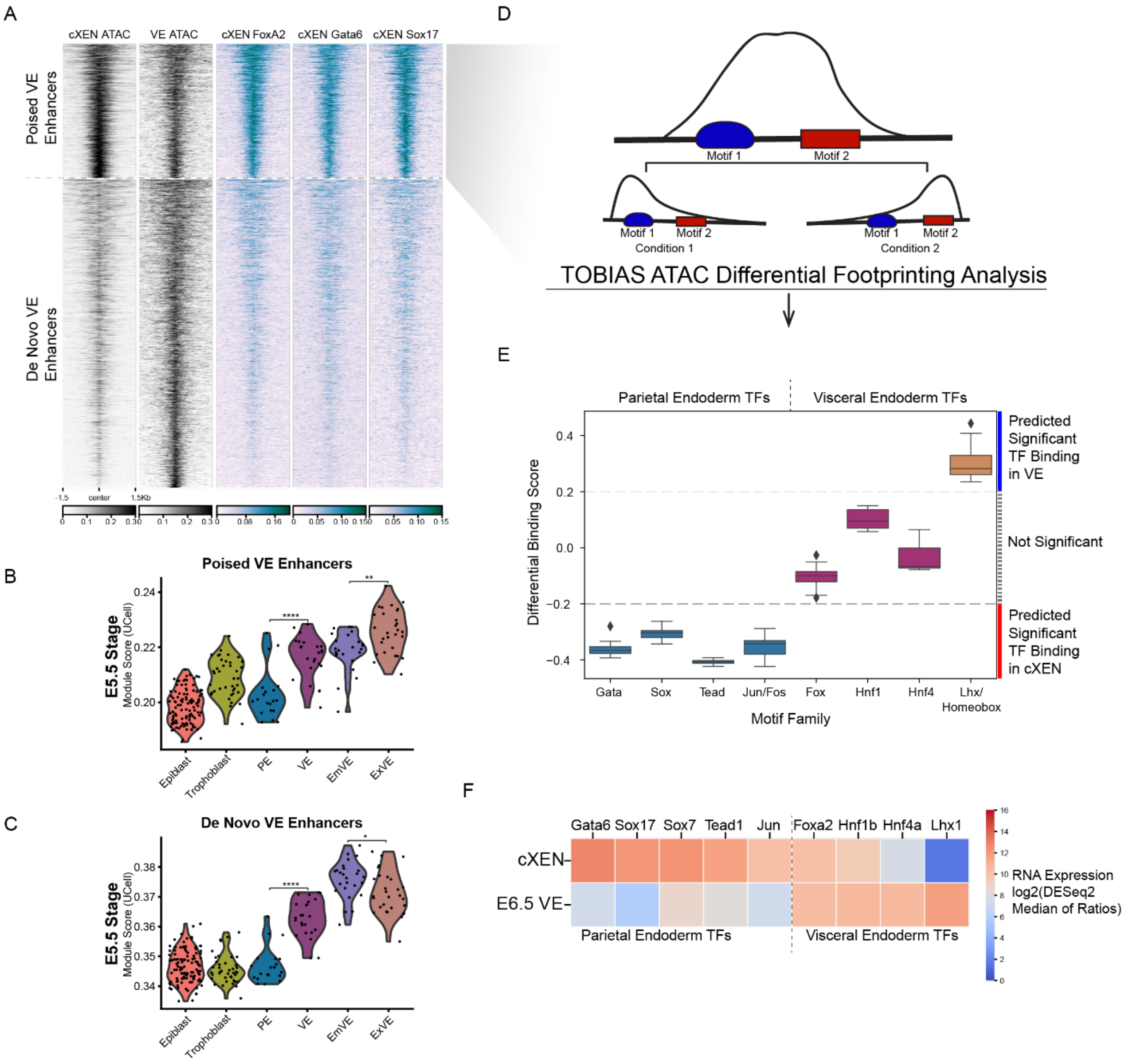
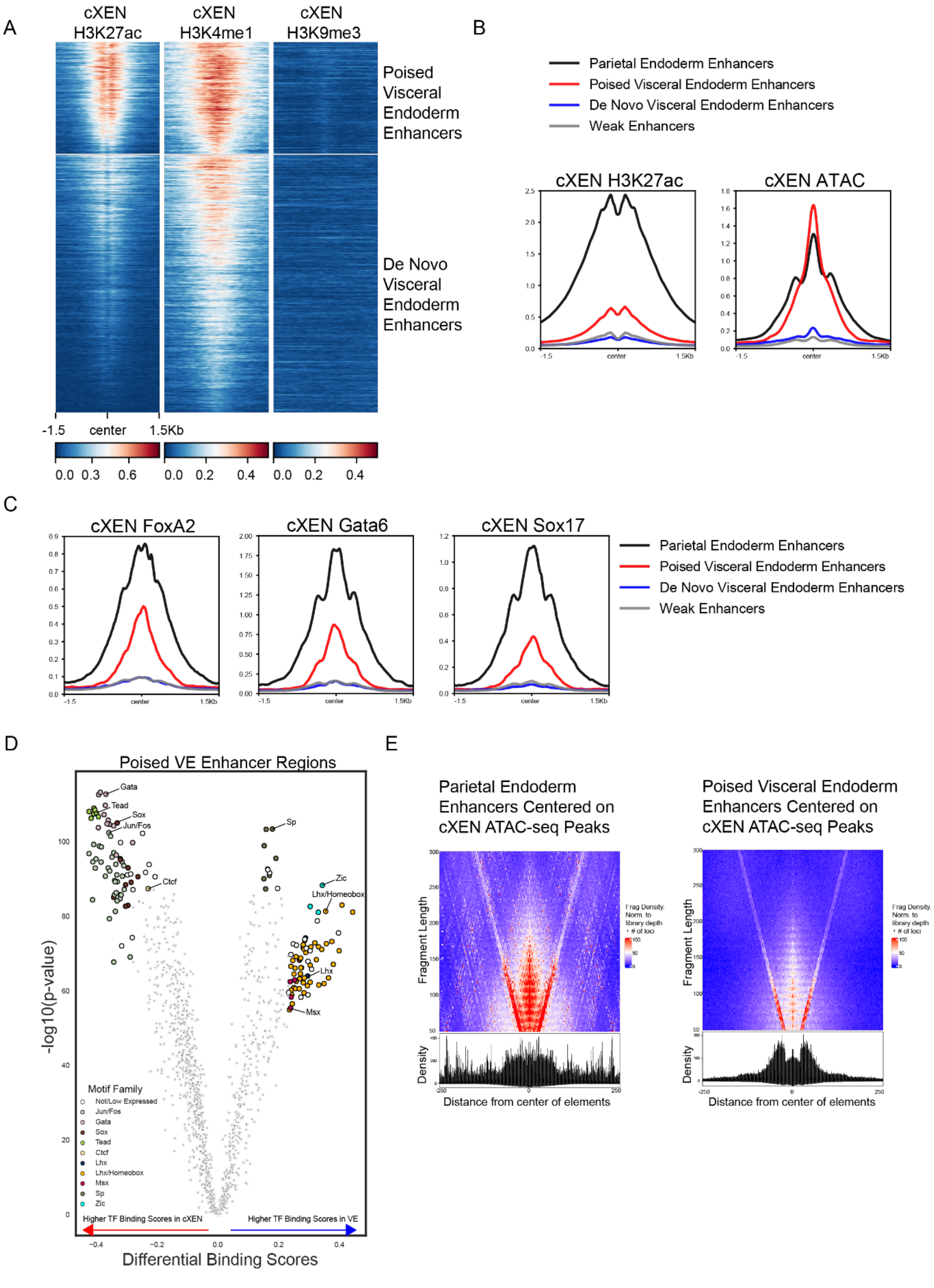

A
cXEN ATAC
VE ATAC
cXEN FoxA2
cXEN Gata6
cXEN Sox17
Poised VE Enhancers
De Novo VE Enhancers
-1.5
center
1.5Kb
0
0.1
0.2
0.3
0
0.1
0.2
0.3
0
0.08
0.16
0
0.05
0.10
0.15
0
0.05
0.10
0.15
B
Poised VE Enhancers
E5.5 Stage
Module Score (UCell)
0.24
0.22
0.20
****
**
Epiblast
Trophoblast
PE
VE
EmVE
ExVE
C
De Novo VE Enhancers
E5.5 Stage
Module Score (UCell)
0.38
0.37
0.36
0.35
0.34
****
*
Epiblast
Trophoblast
PE
VE
EmVE
ExVE
D
Motif 1
Motif 2
Motif 1
Motif 2
Condition 1
Motif 1
Motif 2
Condition 2
TOBIAS ATAC Differential Footprinting Analysis
E
Parietal Endoderm TFs
Visceral Endoderm TFs
Differential Binding Score
0.4
0.2
0.0
−0.2
−0.4
Predicted Significant TF Binding in VE
Not Significant
Predicted Significant TF Binding in cXEN
Gata
Sox
Tead
Jun/Fos
Fox
Hnf1
Hnf4
Lhx/ Homeobox
Motif Family
F
Gata6
Sox17
Sox7
Tead1
Jun
Foxa2
Hnf1b
Hnf4a
Lhx1
cXEN
E6.5 VE
Parietal Endoderm TFs
Visceral Endoderm TFs
16
14
12
10
8
6
4
2
0
RNA Expression log2(DESeq2 Median of Ratios)

Figure 5.6 Epigenetic states at the poised VE enhancers. (A) Tornado plots displaying signals of H3K27ac, H3K4me1, and H3K9me3 at the VE enhancers. (B) Meta profiles of the enrichments for H3K27ac and ATAC-seq in cXEN at PE enhancers, poised VE enhancers, de novo VE enhancers and weak enhancers. (C) Meta profiles of the enrichments for occupancies of FoxA2, Gata6 or Sox17 in cXEN at PE enhancers, poised VE enhancers, de novo VE enhancers and weak enhancers. (D) Differential TF footprinting comparing PE and VE ATAC-seq signal at the poised VE enhancers. (E) Fragment distribution for cXEN ATAC centering at the active PE enhancers or the poised VE enhancers. Column-wise summation were plotted at the bottom.


A
cXEN H3K27ac
cXEN H3K4me1
cXEN H3K9me3
Poised Visceral Endoderm Enhancers
De Novo Visceral Endoderm Enhancers
-1.5
center
1.5Kb
0.0 0.3 0.6
0.0 0.2 0.4
0.0 0.2 0.4
B
Parietal Endoderm Enhancers
Poised Visceral Endoderm Enhancers
De Novo Visceral Endoderm Enhancers
Weak Enhancers
cXEN H3K27ac
cXEN ATAC
C
cXEN FoxA2
cXEN Gata6
cXEN Sox17
Parietal Endoderm Enhancers
Poised Visceral Endoderm Enhancers
De Novo Visceral Endoderm Enhancers
Weak Enhancers
D
Poised VE Enhancer Regions
Gata
Tead
Sox
Jun/Fos
Ctcf
Sp
Zic
Lhx/Homeobox
Lhx
Msx
-log10(p-value)
Motif Family
Not/Low Expressed
Jun/Fos
Gata
Sox
Tead
Ctcf
Lhx
Lhx/Homeobox
Msx
Sp
Zic
Higher TF Binding Scores in cXEN
Higher TF Binding Scores in VE
Differential Binding Scores
E
Parietal Endoderm Enhancers Centered on cXEN ATAC-seq Peaks
Poised Visceral Endoderm Enhancers Centered on cXEN ATAC-seq Peaks
Fragment Length
Frag Density. Norm. to library depth + # of loci
Density
Distance from center of elements

5.5 Summary

In summary, we reported that PE and VE TFs were co-expressed in cXEN (and PrE) where they co-bind at a subset of VE enhancers related to a poised state. Therefore, we hypothesized that in the PrE or PE cells, the VE transcriptional program is primed by FoxA2, Hnf1b, and Hnf4a (VE TFs) via pioneering the VE enhancers but antagonized by Gata6 and Sox17 (PE TFs). We predicted that the downregulation of Gata6 and Sox17, as seen in the progression of VE cell fates, is necessary for the upregulation of the VE transcriptional program. To validate the model, we generated knockout cXEN lines targeting Gata6, Sox17, and FoxA2. Ongoing studies will be focused on integrating the phenotypes to corroborate the gene regulation network.

5.6 Material and Methods

cXEN cell line derivation

cXEN cell were derived from mESCs cells (E14TG2a) using a previously established chemical induction protocol (Niakan et al., 2013) in a cXEN derivation medium consisting of 0.010 M Retinoic Acid (Sigma, 50-185-8562), 10 ng/mL Activin A 24 ng/mL (R&D Systems, 338AC010), Fgf2 (R&D Systems, 23-3FB0-10), and 1 ug/mL Heparin (Sigma, H3393). mESCs were maintained in KnockOut DMEM (Thermo Fisher,10829018), 15% FBS (R&D Systems, S10250H), 1mM MEM Non-Essential Amino Acids (Thermo Fisher, 11140-050), 2mM Glutamax (Thermo Fisher, 35050061), 100 U/mL Pen/Strep (Thermo Fisher, 15140122), 0.1 mM 2(β)-ME (Sigma, M-3148), and 1000 U/mL LIF (Cell Guidance Systems, GFM200) prior to media change using cXEN derivation medium [citation]. The established XEN cell lines were maintained, and passaged in Advanced RPMI (Thermo Fisher,12633012) supplemented with 15% FBS (R&D Systems, S10250H), 2 mM Glutamax (Thermo Fisher, 35050061), 100 U/mL Pen/Strep (Thermo Fisher, 15140122), 0.1 mM 2(β)-ME (Sigma, M-3148).

CRISPR/Cas9 cell line generation

We cloned the FKBP12F36V-mNeonGreen-(3X)HA-tag-P2A-Blasticidin cassette and the flanked homology arms of the target proteins into pUC19 backbone. SgRNA were designed to cut around the stop codons of the endogenous loci of the targeted genes. The homology arms were designed to inframe insertion at the C-terminus of the targeted proteins. The homology arm sequences and sgRNAs were validated by sanger sequencing before use for further experiments. For generation of FKBP12F36V cell lines, cXEN cells were seeded the day before transfection on a 10 cm2 tissue culture plate. 12 ug of plasmid DNA, using equimolar amounts of sgRNA and homology repair template were transfected using 40 ul of Lipofectamine 3000 with 24 ul of P3000 reagent (Thermo Fisher, L3000008). Two days after transfection, the transfected cells were selected in 10 ug/mL of Blasticidin for a week (Cayman Chemical, NC1445974). Single cells of mNeonGreen positive population were then sorted using a BDFACSAria Fusion Sorter in 96-well round bottom plates with irradiated CF-1 MEFs (Thermo Fisher, A34180). The homozygous clones were genotyped by PCR with primers amplifying the knock-in insertion site for FKBP12F36V and verified using flow cytometry.

To generate FoxA2 knockout lines on the Gata6 or Sox17 FKBP12F36V knockin lines, two sgRNAs were designed to delete the sequence encoding for the DNA binding domain. cXEN knockin cell lines were seeded the day before transfection on a 6-well tissue culture plate. 2.5 ug of total plasmid DNA, using equimolar amounts of sgRNA were transfected with 5 ul of lipofectamine 2000 (Thermo Fisher 11668019) for 24h. The BFP positive cells were then sorted using a BDFACSAria in 96-well round bottom plates with irradiated CF-1 MEFs (Thermo Fisher, A34180). The homozygous clones were screened by genotyping.

RNA-seq

Total RNA was extracted with Trizol (Thermo Fisher, 15596026). Strand-specific RNA-seq libraries were generated with 1000 ng of total RNA using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (E7760S). Libraries were paired-end sequenced on the NovaSeq 6000, 100 cycles.

ATAC-seq

50,000 cXEN cells were dissociated using Accutase (Thermo Fisher, A1110501). ATAC-seq was performed using an ATAC-seq kit (Active Motif, 53150) which included components for tagmentation, tagmented DNA clean-up, and library PCR amplification which were all done according to manufacturer's instructions. Libraries were paired-end sequenced on the NovaSeq 6000, 100 cycles.

CUT&RUN

500,000 cXEN and its derivative cell lines were dissociated using Accutase (Thermo Fisher, A1110501). Then, CUT&RUN was performed using pAG-MNase (Epicypher, 15-1016) following the standard protocol published by the Henikoff lab [citation]. Half of the DNA eluted from the CUT&RUN reaction was used for library prep using the NEBNext Ultra II system based on Nan Liu's CUT&RUN library prep protocol but without size selection [citation]. Libraries were paired-end sequenced on the NovaSeq 6000, 100 cycles. Antibodies used in these experiments were as follows: anti-Gata6 (5851S, Cell Signaling Technology), anti-Sox17 (AF1924, R&D Systems), anti-Gata4 (sc-25310, Santa Cruz Biotechnology), anti-FoxA2 (07-633, EMD Millipore Sigma), anti-H3K27ac (MABE647, EMD Millipore Sigma), anti-H3K4me3 (39159, Active Motif), anti-H3K9me3 (ab8898, Abcam), and anti-H3K4me1 (ab8895, Abcam).

Data Analysis

Cut&Run Data Processing

Paired-end fragments were trimmed using trimmomatic as previously described. Paired-end reads were aligned using Bowtie2 with options: --local --very-sensitive --dovetail --phred33 -I 10 -X 700. For transcription factors, mapped fragment sizes were divided into <120bp and >120bp. For most of the analyses, unless otherwise indicated, the <120bp fraction was used which is likely to contain direct TF binding sites as previously characterized. The fragments were filtered for a MAPQ score of 30 prior to calling peaks. To call peaks, MACS2 was called using the default narrowPeak setting with the following parameters: -f BEDPE -q 0.01 -B --SPMR --keep-dup auto. For histone marks, mapped fragment sizes between 0bp and 700bp were used for downstream analyses. The fragments were filtered for a MAPQ score of 30 prior to calling peaks. To call peaks for H3K27ac and H3K4me3, MACS2 was called using the default narrowPeak setting with the following parameters: -f BEDPE -q 0.01 -B --SPMR --keep-dup auto. To call peaks for H3K4me1 and H3K9me3, MACS2 was called using default broadPeak settings with the following parameters: -f BEDPE -q 0.01 -B --SPMR --keep-dup auto.

scRNA-seq Data Analysis

scRNA-seq data of E4.5 and E5.5 embryos were integrated from Nowotschin et. al. (Nowotschin et al., 2019) and Qiu et. al (Qiu et al., 2022) using seurat V4 (Hao et al., 2021). Cell clustering and plots were generated by seurat V4 (Hao et al., 2021) and scanpy (Wolf et al., 2018). The signature scores were calculated by the UCell algorithm (Andreatta and Carmona, 2021). Weighted gene correlation network analysis of single cell transcriptomes were performed by hdWGCNA (Morabito et al., 2021).

5.7 Acknowledgements

For their contributions in chapter 5, thanks to: Paula Pham, Han Han, Jeff Jiajing Zhou, Wenqi Wang, Cornelis Murre and Ken Cho. The material in this chapter is currently prepared for submission for publication.

CHAPTER 6

Concluding Remarks

The classic paradigm of cell fate decision denoted that master transcription factors execute the developmental instructions encoded in enhancer sequences to activate cell-type-specific gene programs (Furlong and Levine, 2018; Jindal and Farley, 2021). Decades of research had made considerable progress in deciphering the enhancer codes in the context of the linear configuration of DNA motifs (Furlong and Levine, 2018). However, numerous evidences indicate that chromatin organization also play key role in facilitating the spatial and temporal control of gene expression. For example, we had previously reported that nuclei repositioning of Ebf1 versus Bcl11b coincides with their activations that determines B cell versus T cell commitment (Isoda et al., 2017; Lin et al., 2012). In chapter 4 of this thesis, we presented two published works that characterized the dramatic alterations in chromatin large-scale organization when mouse or human neutrophils are challenged by external stimuli (Denholtz et al., 2020; Zhu et al., 2020). Although these data highlighted the importance of nuclei compartment organization in gene regulation, we are still not clear about the detailed sequence of the events in the processes. It will be critical to identify the factors orchestrating the changes in compartment.

Enhancer-promoter (E-P) communication is long speculated to directly control gene expression (Furlong and Levine, 2018; Lim and Levine, 2021). Despite the revolution of chromosome conformation capture (3C) tools like in situ Hi-C (Rao et al., 2014), profiling E-P interactions is still challenged. In line with the demand, we developed a new 3C method variant, named MID Hi-C akin to in situ Hi-C, that is capable of efficient detection of E-P interactions. MID Hi-C will facilitate the field to scale up mapping fine-scale chromatin interactions in various developmental and engineered systems. As a proof of principle, we leveraged MID Hi-C to study the responses of E-P interactions upon BAF complex perturbation. We uncovered that enhancer sequences encoding for differential recruitment of the BAF complex had profound impact in the configuration of E-P interaction. We observed that E-P interaction is extremely robust so that abrogation requires near eradication of the proteins bound on the enhancer. This observation suggests that establishment of enhancer-promoter communication may not need the enhancer being activated. For instance, enhancer priming might be sufficient to trigger the interaction with cognate promoters. As indirect evidence, several recent studies independently reported that chromatin structures including E-P interactions were already formed prior to cell-type-specific enhancer activation as well as gene induction (Espinola et al., 2021; Ing-Simmons et al., 2021).

In addition, we observed enhancers can affect the chromatin interactions within the E-P interacting domain. We show that BAF-dependent enhancers permit paired genomic interactions beyond enhancer boundaries and do not dictate genomic interactions within enhancer-promoter loop domains. In contrast, BAF-independent enhancers interact with promoter regions within tightly insulated enhancer-promoter loop domains that are marked by promoter and enhancer boundary elements. Same as the other reports, we also observed evidence that both loop extrusion and phase separation seem contributing to the formation of E-P interaction. As a result, our data agree with the mixture model (Hsieh et al., 2021). As more MID Hi-C data accumulated, the complete picture of E-P interaction will be unveiled in the near future.

REFERENCES

Abdennur, N., and Mirny, L.A. (2019). Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics 36, 311–316. https://doi.org/10.1093/bioinformatics/btz540.

Alexander, J.M., Guan, J., Li, B., Maliskova, L., Song, M., Shen, Y., Huang, B., Lomvardas, S., and Weiner, O.D. (2019). Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. Elife 8, e41769. https://doi.org/10.7554/elife.41769.

Alver, B.H., Kim, K.H., Lu, P., Wang, X., Manchester, H.E., Wang, W., Haswell, J.R., Park, P.J., and Roberts, C.W.M. (2017). The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. Nat Commun 8, 14648. https://doi.org/10.1038/ncomms14648.

Andreatta, M., and Carmona, S.J. (2021). UCell: Robust and scalable single-cell gene signature scoring. Comput Struct Biotechnology J 19, 3796–3798. https://doi.org/10.1016/j.csbj.2021.06.043.

Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat Rev Mol Cell Bio 10, 91–103. https://doi.org/10.1038/nrm2618.

Ay, F., Bailey, T.L., and Noble, W.S. (2014). Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. Genome Res 24, 999–1011. https://doi.org/10.1101/gr.160374.113.

Ba, Z., Lou, J., Ye, A.Y., Dai, H.-Q., Dring, E.W., Lin, S.G., Jain, S., Kyritsis, N., Kieffer-Kwon, K.-R., Casellas, R., et al. (2020). CTCF orchestrates long-range cohesin-driven V(D)J recombinational scanning. Nature 586, 305–310. https://doi.org/10.1038/s41586-020-2578-0.

Baù, D., Sanyal, A., Lajoie, B.R., Capriotti, E., Byron, M., Lawrence, J.B., Dekker, J., and Marti-Renom, M.A. (2011). The three-dimensional folding of the α-globin gene domain reveals formation of chromatin globules. Nat Struct Mol Biol 18, 107–114. https://doi.org/10.1038/nsmb.1936.

Benabdallah, N.S., Williamson, I., Illingworth, R.S., Kane, L., Boyle, S., Sengupta, D., Grimes, G.R., Therizols, P., and Bickmore, W.A. (2019). Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. Mol Cell https://doi.org/10.1016/j.molcel.2019.07.038.

Bentsen, M., Goymann, P., Schultheis, H., Klee, K., Petrova, A., Wiegandt, R., Fust, A., Preussner, J., Kuenne, C., Braun, T., et al. (2020). ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. Nat Commun 11, 4267. https://doi.org/10.1038/s41467-020-18035-1.

Bickmore, W.A., and van Steensel, B. (2013). Genome Architecture: Domain Organization of Interphase Chromosomes. Cell 152, 1270–1284. https://doi.org/10.1016/j.cell.2013.02.001.

Boeynaems, S., Alberti, S., Fawzi, N.L., Mittag, T., Polymenidou, M., Rousseau, F., Schymkowitz, J., Shorter, J., Wolozin, B., Bosch, L.V.D., et al. (2018). Protein Phase Separation: A New Phase in Cell Biology. Trends Cell Biol 28, 420–435. https://doi.org/10.1016/j.tcb.2018.02.004.

Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell https://doi.org/10.1016/j.cell.2018.10.042.

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

Bonev, B., Cohen, N.M., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.-P., Tanay, A., et al. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell 171, 557-572.e24. https://doi.org/10.1016/j.cell.2017.09.043.

Busslinger, G.A., Stocsits, R.R., Lelij, P. van der, Axelsson, E., Tedeschi, A., Galjart, N., and Peters, J.-M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. Nature 544, 503–507. https://doi.org/10.1038/nature22063.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative Binding of Transcription Factors Orchestrates Reprogramming. Cell 168, 442-459.e20. https://doi.org/10.1016/j.cell.2016.12.016.

Crump, N.T., Ballabio, E., Godfrey, L., Thorne, R., Repapi, E., Kerry, J., Tapia, M., Hua, P., Lagerholm, C., Filippakopoulos, P., et al. (2021). BET inhibition disrupts transcription but retains enhancer-promoter contact. Nat Commun 12, 223. https://doi.org/10.1038/s41467-020-20400-z.

Cuartero, S., Weiss, F.D., Dharmalingam, G., Guo, Y., Ing-Simmons, E., Masella, S., Robles-Rebollo, I., Xiao, X., Wang, Y.-F., Barozzi, I., et al. (2018). Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. Nat Immunol 19, 932–941. https://doi.org/10.1038/s41590-018-0184-1.

Dai, H.-Q., Hu, H., Lou, J., Ye, A.Y., Ba, Z., Zhang, X., Zhang, Y., Zhao, L., Yoon, H.S., Chapdelaine-Williams, A.M., et al. (2021). Loop extrusion mediates physiological Igh locus contraction for RAG scanning. Nature 590, 338–343. https://doi.org/10.1038/s41586-020-03121-7.

Davies, J.O.J., Telenius, J.M., McGowan, S.J., Roberts, N.A., Taylor, S., Higgs, D.R., and Hughes, J.R. (2016). Multiplexed analysis of chromosome conformation at vastly improved sensitivity. Nat Methods 13, 74–80. https://doi.org/10.1038/nmeth.3664.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. Science 295, 1306–1311. https://doi.org/10.1126/science.1067799.

Dekker, J., Belmont, A.S., Guttman, M., Leshyk, V.O., Lis, J.T., Lomvardas, S., Mirny, L.A., O’Shea, C.C., Park, P.J., Ren, B., et al. (2017). The 4D nucleome project. Nature 549, 219–226. https://doi.org/10.1038/nature23884.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling Long-Range Genomic Interactions at a Native Locus by Targeted Tethering of a Looping Factor. Cell 149, 1233–1244. https://doi.org/10.1016/j.cell.2012.03.051.

Denholtz, M., Zhu, Y., He, Z., Lu, H., Isoda, T., Döhrmann, S., Nizet, V., and Murre, C. (2020). Upon microbial challenge, human neutrophils undergo rapid changes in nuclear architecture and chromatin folding to orchestrate an immediate inflammatory gene program. Gene Dev 34, 149–165. https://doi.org/10.1101/gad.333708.119.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380. https://doi.org/10.1038/nature11082.

Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. Mol Cell 62, 668–680. https://doi.org/10.1016/j.molcel.2016.05.018.

Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C., et al. (2006). Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. Genome Res 16, 1299–1309. https://doi.org/10.1101/gr.5571506.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S.P., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016a). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Syst 3, 95–98. https://doi.org/10.1016/j.cels.2016.07.002.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016b). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst 3, 99–101. https://doi.org/10.1016/j.cels.2015.07.012.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat Methods 9, 215–216. https://doi.org/10.1038/nmeth.1906.

Ernst, J., and Kellis, M. (2017). Chromatin-state discovery and genome annotation with ChromHMM. Nat Protoc 12, 2478–2492. https://doi.org/10.1038/nprot.2017.124.

Espinola, S.M., Götz, M., Bellec, M., Messina, O., Fiche, J.-B., Houbron, C., Dejean, M., Reim, I., Gizzi, A.M.C., Lagha, M., et al. (2021). Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development. Nat Genet 53, 477–486. https://doi.org/10.1038/s41588-021-00816-z.

Filimonow, K., and Fuente, R. de la (2021). Specification and role of extraembryonic endoderm lineages in the periimplantation mouse embryo. Theriogenology 180, 189–206. https://doi.org/10.1016/j.theriogenology.2021.12.021.

Franke, M., Ibrahim, D.M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., Kraft, K., Kempfer, R., Jerković, I., Chan, W.-L., et al. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. Nature 538, 265–269. https://doi.org/10.1038/nature19800.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell Reports 15, 2038–2049. https://doi.org/10.1016/j.celrep.2016.04.085.

Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer Control of Transcriptional Bursting. Cell 166, 358–368. https://doi.org/10.1016/j.cell.2016.05.025.

Furlong, E.E.M., and Levine, M. (2018). Developmental enhancers and chromosome topology. Science 361, 1341–1345. https://doi.org/10.1126/science.aau0320.

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C., et al. (2018). A pathway for mitotic chromosome formation. Science 359, eaao6135. https://doi.org/10.1126/science.aao6135.

Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell https://doi.org/10.1016/j.cell.2019.08.037.

Guo, C., Gerasimova, T., Hao, H., Ivanova, I., Chakraborty, T., Selimyan, R., Oltz, E.M., and Sen, R. (2011). Two Forms of Loops Generate the Chromatin Conformation of the Immunoglobulin Heavy-Chain Gene Locus. Cell 147, 332–343. https://doi.org/10.1016/j.cell.2011.08.049.

Haarhuis, J.H.I., Weide, R.H. van der, Blomen, V.A., Yáñez-Cuna, J.O., Amendola, M., Ruiten, M.S. van, Krijger, P.H.L., Teunissen, H., Medema, R.H., Steensel, B. van, et al. (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell 169, 693-707.e14. https://doi.org/10.1016/j.cell.2017.04.013.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell 184, 3573-3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol Cell 38, 576–589. https://doi.org/10.1016/j.molcel.2010.05.004.

Hill, L., Ebert, A., Jaritz, M., Wutz, G., Nagasaka, K., Tagoh, H., Kostanova-Poliakova, D., Schindler, K., Sun, Q., Bönelt, P., et al. (2020). Wapl repression by Pax5 promotes V gene recombination by Igh loop extrusion. Nature 1–6. https://doi.org/10.1038/s41586-020-2454-y.

Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.-L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A., et al. (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science 351, 1454–1458. https://doi.org/10.1126/science.aad9024.

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

Hogan, B.L.M., and Tilly, R. (1981). Cell interactions and endoderm differentiation in cultured mouse embryos. Development 62, 379–394. https://doi.org/10.1242/dev.62.1.379.

Hogan, B.L.M., Cooper, A.R., and Kurkinen, M. (1980). Incorporation into Reichert's membrane of laminin-like extracellular proteins synthesized by parietal endoderm cells of the mouse embryo. Dev Biol 80, 289–300. https://doi.org/10.1016/0012-1606(80)90405-4.

Hsieh, T.-H.S., Fudenberg, G., Goloborodko, A., and Rando, O.J. (2016). Micro-C XL: assaying chromosome conformation from the nucleosome to the entire genome. Nat Methods 13, 1009–1011. https://doi.org/10.1038/nmeth.4025.

Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R., and Darzacq, X. (2020). Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. Mol Cell https://doi.org/10.1016/j.molcel.2020.03.002.

Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Darzacq, X., and Tjian, R. (2021). Enhancer-promoter interactions and transcription are maintained upon acute loss of CTCF, cohesin, WAPL, and YY1. Biorxiv 2021.07.14.452365. https://doi.org/10.1101/2021.07.14.452365.

Ibn-Salem, J., Köhler, S., Love, M.I., Chung, H.-R., Huang, N., Hurles, M.E., Haendel, M., Washington, N.L., Smedley, D., Mungall, C.J., et al. (2014). Deletions of chromosomal regulatory boundaries are associated with congenital disease. Genome Biol 15, 423. https://doi.org/10.1186/s13059-014-0423-1.

Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods 9, 999. https://doi.org/10.1038/nmeth.2148.

Ing-Simmons, E., Vaid, R., Bing, X.Y., Levine, M., Mannervik, M., and Vaquerizas, J.M. (2021). Independence of chromatin conformation and gene regulation during Drosophila dorsoventral patterning. Nat Genet 53, 487–499. https://doi.org/10.1038/s41588-021-00799-x.

Isoda, T., Moore, A.J., He, Z., Chandra, V., Aida, M., Denholtz, M., Hamburg, J.P. van, Fisch, K.M., Chang, A.N., Fahl, S.P., et al. (2017). Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. Cell 171, 103-119.e18. https://doi.org/10.1016/j.cell.2017.09.001.

Jaeger, M.G., Schwalb, B., Mackowiak, S.D., Velychko, T., Hanzl, A., Imrichova, H., Brand, M., Agerer, B., Chorn, S., Nabet, B., et al. (2020). Selective Mediator dependence of cell-type-specifying transcription. Nat Genet 52, 719–727. https://doi.org/10.1038/s41588-020-0635-0.

Jain, S., Ba, Z., Zhang, Y., Dai, H.-Q., and Alt, F.W. (2018). CTCF-Binding Elements Mediate Accessibility of RAG Substrates During Chromatin Scanning. Cell 174, 102-116.e14.

https://doi.org/10.1016/j.cell.2018.04.035.

Jerkovic´, I., and Cavalli, G. (2021). Understanding 3D genome organization by multidisciplinary methods. Nat Rev Mol Cell Bio 22, 511–528. https://doi.org/10.1038/s41580-021-00362-w.

Jhunjhunwala, S., Zelm, M.C. van, Peak, M.M., Cutchin, S., Riblet, R., Dongen, J.J.M. van, Grosveld, F.G., Knoch, T.A., and Murre, C. (2008). The 3D Structure of the Immunoglobulin Heavy-Chain Locus: Implications for Long-Range Genomic Interactions. Cell 133, 265–279. https://doi.org/10.1016/j.cell.2008.03.024.

Jindal, G.A., and Farley, E.K. (2021). Enhancer grammar in development, evolution, and disease: dependencies and interplay. Dev Cell 56, 575–587. https://doi.org/10.1016/j.devcel.2021.02.016.

Khattabi, L.E., Zhao, H., Kalchschmidt, J., Young, N., Jung, S., Blerkom, P.V., Kieffer-Kwon, P., Kieffer-Kwon, K.-R., Park, S., Wang, X., et al. (2019). A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. Cell 178, 1145-1158.e20. https://doi.org/10.1016/j.cell.2019.07.011.

Kieffer-Kwon, K.-R., Tang, Z., Mathe, E., Qian, J., Sung, M.-H., Li, G., Resch, W., Baek, S., Pruett, N., Grøntved, L., et al. (2013). Interactome Maps of Mouse Gene Regulatory Domains Reveal Basic Principles of Transcriptional Regulation. Cell 155, 1507–1520. https://doi.org/10.1016/j.cell.2013.11.039.

Kieffer-Kwon, K.-R., Nimura, K., Rao, S.S.P., Xu, J., Jung, S., Pekowska, A., Dose, M., Stevens, E., Mathe, E., Dong, P., et al. (2017). Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation. Mol Cell 67, 566-578.e10. https://doi.org/10.1016/j.molcel.2017.07.013.

Kim, E., Kerssemakers, J., Shaltiel, I.A., Haering, C.H., and Dekker, C. (2020). DNA-loop extruding condensin complexes can traverse one another. Nature 579, 438–442. https://doi.org/10.1038/s41586-020-2067-5.

Krietenstein, N., Abraham, S., Venev, S.V., Abdennur, N., Gibcus, J., Hsieh, T.-H.S., Parsi, K.M., Yang, L., Maehr, R., Mirny, L.A., et al. (2020). Ultrastructural Details of Mammalian Chromosome Architecture. Mol Cell https://doi.org/10.1016/j.molcel.2020.03.003.

Laat, W. de, and Duboule, D. (2013). Topology of mammalian developmental enhancers and their regulatory landscapes. Nature 502, 499–506. https://doi.org/10.1038/nature12753.

Lajoie, B.R., Dekker, J., and Kaplan, N. (2015). The Hitchhiker's guide to Hi-C analysis: Practical guidelines. Methods 72, 65–75. https://doi.org/10.1016/j.ymeth.2014.10.031.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923.

Lareau, C.A., and Aryee, M.J. (2018). hichipper: a preprocessing pipeline for calling DNA loops from HiChIP data. Nat Methods 15, 155–156. https://doi.org/10.1038/nmeth.4583.

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

Lieberman-Aiden, E., Berkum, N.L. van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science 326, 289–293. https://doi.org/10.1126/science.1181369.

Lim, B., and Levine, M.S. (2021). Enhancer-promoter communication: hubs or loops? Curr Opin Genet Dev 67, 5–9. https://doi.org/10.1016/j.gde.2020.10.001.

Lin, Y.C., Benner, C., Mansson, R., Heinz, S., Miyazaki, K., Miyazaki, M., Chandra, V., Bossen, C., Glass, C.K., and Murre, C. (2012). Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. Nat Immunol 13, 1196. https://doi.org/10.1038/ni.2432.

Linares-Saldana, R., Kim, W., Bolar, N.A., Zhang, H., Koch-Bojalad, B.A., Yoon, S., Shah, P.P., Karnay, A., Park, D.S., Luppino, J.M., et al. (2021). BRD4 orchestrates genome folding to promote neural crest differentiation. Nat Genet 53, 1480–1492. https://doi.org/10.1038/s41588-021-00934-8.

Link, V.M., Duttke, S.H., Chun, H.B., Holtman, I.R., Westin, E., Hoeksema, M.A., Abe, Y., Skola, D., Romanoski, C.E., Tao, J., et al. (2018). Analysis of Genetically Diverse Macrophages Reveals Local and Domain-wide Mechanisms that Control Transcription Factor Binding and Function. Cell 173, 1796-1809.e17. https://doi.org/10.1016/j.cell.2018.04.018.

Liu, N.Q., Maresca, M., Brand, T. van den, Braccioli, L., Schijns, M.M.G.A., Teunissen, H., Bruneau, B.G., Nora, E.P., and Wit, E. de (2021). WAPL maintains a cohesin loading cycle to preserve cell-type-specific distal gene regulation. Nat Genet 53, 100–109. https://doi.org/10.1038/s41588-020-00744-4.

Lopez-Delisle, L., Rabbani, L., Wolff, J., Bhardwaj, V., Backofen, R., Grüning, B., Ramírez, F., and Manke, T. (2020). pyGenomeTracks: reproducible plots for multivariate genomic data sets. Bioinformatics 37, btaa692-. https://doi.org/10.1093/bioinformatics/btaa692.

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

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. Cell 161, 1012–1025. https://doi.org/10.1016/j.cell.2015.04.004.

Mifsud, B., Tavares-Cadete, F., Young, A.N., Sugar, R., Schoenfelder, S., Ferreira, L., Wingett, S.W., Andrews, S., Grey, W., Ewels, P.A., et al. (2015). Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet 47, 598–606. https://doi.org/10.1038/ng.3286.

Morabito, S., Miyoshi, E., Michael, N., Shahin, S., Martini, A.C., Head, E., Silva, J., Leavy, K., Perez-Rosendahl, M., and Swarup, V. (2021). Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. Nat Genet 53, 1143–1155. https://doi.org/10.1038/s41588-021-00894-z.

Morgan, S.L., Mariano, N.C., Bermudez, A., Arruda, N.L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.-F., et al. (2017). Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. Nat Commun 8, 15993. https://doi.org/10.1038/ncomms15993.

Mumbach, M.R., Rubin, A.J., Flynn, R.A., Dai, C., Khavari, P.A., Greenleaf, W.J., and Chang, H.Y. (2016). HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Methods 13, 919–922. https://doi.org/10.1038/nmeth.3999.

Münkel, C., and Langowski, J. (1998). Chromosome structure predicted by a polymer model. Phys Rev E 57, 5888–5896. https://doi.org/10.1103/physreve.57.5888.

Nabet, B., Roberts, J.M., Buckley, D.L., Paulk, J., Dastjerdi, S., Yang, A., Leggett, A.L., Erb, M.A., Lawlor, M.A., Souza, A., et al. (2018). The dTAG system for immediate and target-specific protein degradation. Nat Chem Biol 14, 431–441. https://doi.org/10.1038/s41589-018-0021-8.

Nakayama, R.T., Pulice, J.L., Valencia, A.M., McBride, M.J., McKenzie, Z.M., Gillespie, M.A., Ku, W.L., Teng, M., Cui, K., Williams, R.T., et al. (2017). SMARCB1 is required for widespread BAF complex–mediated activation of enhancers and bivalent promoters. Nat Genet 49, 1613–1623. https://doi.org/10.1038/ng.3958.

Niakan, K.K., Schrode, N., Cho, L.T.Y., and Hadjantonakis, A.-K. (2013). Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells. Nat Protoc 8, 1028–1041. https://doi.org/10.1038/nprot.2013.049.

Niu, L., Shen, W., Shi, Z., Tan, Y., He, N., Wan, J., Sun, J., Zhang, Y., Huang, Y., Wang, W., et al. (2021). Three-dimensional folding dynamics of the Xenopus tropicalis genome. Nat Genet 53, 1075–1087. https://doi.org/10.1038/s41588-021-00878-z.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., Berkum, N.L. van, Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485, 381–385. https://doi.org/10.1038/nature11049.

Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell 169, 930-944.e22. https://doi.org/10.1016/j.cell.2017.05.004.

Nora, E.P., Caccianini, L., Fudenberg, G., So, K., Kameswaran, V., Nagle, A., Uebersohn, A., Hajj, B., Saux, A.L., Coulon, A., et al. (2020). Molecular basis of CTCF binding polarity in genome folding. Nat Commun 11, 5612. https://doi.org/10.1038/s41467-020-19283-x.

Nowotschin, S., Setty, M., Kuo, Y.-Y., Liu, V., Garg, V., Sharma, R., Simon, C.S., Saiz, N., Gardner, R., Boutet, S.C., et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. Nature 569, 361–367. https://doi.org/10.1038/s41586-019-1127-1.

Oksuz, B.A., Yang, L., Abraham, S., Venev, S.V., Krietenstein, N., Parsi, K.M., Ozadam, H., Oomen, M.E., Nand, A., Mao, H., et al. (2021). Systematic evaluation of chromosome conformation capture assays. Nat Methods 1–10. https://doi.org/10.1038/s41592-021-01248-7.

Paca, A., Séguin, C.A., Clements, M., Ryczko, M., Rossant, J., Rodriguez, T.A., and Kunath, T. (2012). BMP signaling induces visceral endoderm differentiation of XEN cells and parietal endoderm. Dev Biol 361, 90–102. https://doi.org/10.1016/j.ydbio.2011.10.013.

Park, Y.-K., Lee, J.-E., Yan, Z., McKernan, K., O'Haren, T., Wang, W., Peng, W., and Ge, K. (2021). Interplay of BAF and MLL4 promotes cell type-specific enhancer activation. Nat Commun 12, 1630. https://doi.org/10.1038/s41467-021-21893-y.

Paulson, J.R., and Laemmli, U.K. (1977). The structure of histone-depleted metaphase chromosomes. Cell 12, 817–828. https://doi.org/10.1016/0092-8674(77)90280-x.

Phanstiel, D.H., Boyle, A.P., Heidari, N., and Snyder, M.P. (2015). Mango: a bias-correcting ChIA-PET analysis pipeline. Bioinformatics 31, 3092–3098. https://doi.org/10.1093/bioinformatics/btv336.

Phillips-Cremins, J.E., Sauria, M.E.G., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S.K., Ong, C.-T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment. Cell 153, 1281–1295. https://doi.org/10.1016/j.cell.2013.04.053.

Pradhan, B., Barth, R., Kim, E., Davidson, I.F., Bauer, B., Laar, T. van, Yang, W., Ryu, J.-K., Torre, J. van der, Peters, J.-M., et al. (2021). SMC complexes can traverse physical roadblocks bigger than their ring size. Biorxiv 2021.07.15.452501. https://doi.org/10.1101/2021.07.15.452501.

Qiu, C., Cao, J., Martin, B.K., Li, T., Welsh, I.C., Srivatsan, S., Huang, X., Calderon, D., Noble, W.S., Disteche, C.M., et al. (2022). Systematic reconstruction of cellular trajectories across mouse embryogenesis. Nat Genet 54, 328–341. https://doi.org/10.1038/s41588-022-01018-x.

Quinodoz, S.A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J.M., Detmar, E., Lai, M.M., Shishkin, A.A., Bhat, P., Takei, Y., et al. (2018). Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. Cell 174, 744-757.e24. https://doi.org/10.1016/j.cell.2018.05.024.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res 42, W187–W191.

https://doi.org/10.1093/nar/gku365.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159, 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021.

Rao, S.S.P., Huang, S.-C., Hilaire, B.G.S., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.-R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., et al. (2017). Cohesin Loss Eliminates All Loop Domains. Cell 171. https://doi.org/10.1016/j.cell.2017.09.026.

Rattner, J.B., and Lin, C.C. (1985). Radial loops and helical coils coexist in metaphase chromosomes. Cell 42, 291–296. https://doi.org/10.1016/s0092-8674(85)80124-0.

Ryu, J.-K., Rah, S.-H., Janissen, R., Kerssemakers, J.W.J., Bonato, A., Michieletto, D., and Dekker, C. (2021). Condensin extrudes DNA loops in steps up to hundreds of base pairs that are generated by ATP binding events. Nucleic Acids Res 50, 820–832. https://doi.org/10.1093/nar/gkab1268.

Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science 361, eaar3958. https://doi.org/10.1126/science.aar3958.

Sanborn, A.L., Rao, S.S.P., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc National Acad Sci 112, E6456–E6465. https://doi.org/10.1073/pnas.1518552112.

Schoenfelder, S., Furlan-Magaril, M., Mifsud, B., Tavares-Cadete, F., Sugar, R., Javierre, B.-M., Nagano, T., Katsman, Y., Sakthidevi, M., Wingett, S.W., et al. (2015). The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. Genome Res 25, 582–597. https://doi.org/10.1101/gr.185272.114.

Schrode, N., Saiz, N., Talia, S.D., and Hadjantonakis, A.-K. (2014). GATA6 Levels Modulate Primitive Endoderm Cell Fate Choice and Timing in the Mouse Blastocyst. Dev Cell 29, 454–467. https://doi.org/10.1016/j.devcel.2014.04.011.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C.H., Mirny, L., et al. (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature 551, 51. https://doi.org/10.1038/nature24281.

Sedat, J., and Manuelidis, L. (1978). A Direct Approach to the Structure of Eukaryotic Chromosomes. Cold Spring Harb Sym 42, 331–350. https://doi.org/10.1101/sqb.1978.042.01.035.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol 16, 259. https://doi.org/10.1186/s13059-015-0831-x.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. Cell 148, 458–472. https://doi.org/10.1016/j.cell.2012.01.010.

Shin, H., Shi, Y., Dai, C., Tjong, H., Gong, K., Alber, F., and Zhou, X.J. (2016). TopDom: an efficient and deterministic method for identifying topological domains in genomes. Nucleic Acids Res 44, e70–e70. https://doi.org/10.1093/nar/gkv1505.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., Wit, E. de, Steensel, B. van, and Laat, W. de (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). Nat Genet 38, 1348–1354. https://doi.org/10.1038/ng1896.

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

Stadhouders, R., Vidal, E., Serra, F., Stefano, B.D., Dily, F.L., Quilez, J., Gomez, A., Collombet, S., Berenguer, C., Cuartero, Y., et al. (2018). Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. Nat Genet 50, 238–249. https://doi.org/10.1038/s41588-017-0030-7.

Stovner, E.B., and Sætrom, P. (2019). PyRanges: efficient comparison of genomic intervals in Python. Bioinformatics 36, 918–919. https://doi.org/10.1093/bioinformatics/btz615.

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

Takaoka, K., and Hamada, H. (2011). Cell fate decisions and axis determination in the early mouse embryo. Development 139, 3–14. https://doi.org/10.1242/dev.060095.

Thiecke, M.J., Wutz, G., Muhar, M., Tang, W., Bevan, S., Malysheva, V., Stocsits, R., Neumann, T., Zuber, J., Fraser, P., et al. (2020). Cohesin-Dependent and -Independent Mechanisms Mediate Chromosomal Contacts between Promoters and Enhancers. Cell Reports 32, 107929. https://doi.org/10.1016/j.celrep.2020.107929.

Thompson, J.J., Lee, D.J., Mitra, A., Frail, S., Dale, R.K., and Rocha, P.P. (2022). Extensive co-binding and rapid redistribution of NANOG and GATA6 during emergence of divergent lineages. Nat Commun 13, 4257. https://doi.org/10.1038/s41467-022-31938-5.

Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F., and Laat, W. de (2002). Looping and Interaction between Hypersensitive Sites in the Active β-globin Locus. Mol Cell 10, 1453–1465. https://doi.org/10.1016/s1097-2765(02)00781-5.

Vian, L., Pękowska, A., Rao, S.S.P., Kieffer-Kwon, K.-R., Jung, S., Baranello, L., Huang, S.-C., Khattabi, L.E., Dose, M., Pruett, N., et al. (2018). The Energetics and Physiological Impact of Cohesin Extrusion. Cell https://doi.org/10.1016/j.cell.2018.03.072.

Wang, X., Lee, R.S., Alver, B.H., Haswell, J.R., Wang, S., Mieczkowski, J., Drier, Y., Gillespie, S.M., Archer, T.C., Wu, J.N., et al. (2017). SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. Nat Genet 49, 289–295. https://doi.org/10.1038/ng.3746.

Watson, A.J. (1992). The cell biology of blastocyst development. Mol Reprod Dev 33, 492–504. https://doi.org/10.1002/mrd.1080330417.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 19, 15. https://doi.org/10.1186/s13059-017-1382-0.

Yamanaka, Y., Ralston, A., Stephenson, R.O., and Rossant, J. (2006). Cell and molecular regulation of the mouse blastocyst. Dev Dynam 235, 2301–2314. https://doi.org/10.1002/dvdy.20844.

Yang, H., Luan, Y., Liu, T., Lee, H.J., Fang, L., Wang, Y., Wang, X., Zhang, B., Jin, Q., Ang, K.C., et al. (2020). A map of cis-regulatory elements and 3D genome structures in zebrafish. Nature 1–7. https://doi.org/10.1038/s41586-020-2962-9.

Zhang, L., Zhang, Y., Chen, Y., Gholamalamdari, O., Wang, Y., Ma, J., and Belmont, A.S. (2020). TSA-seq reveals a largely conserved genome organization relative to nuclear speckles with small position changes tightly correlated with gene expression changes. Genome Res 31, gr.266239.120. https://doi.org/10.1101/gr.266239.120.

Zhang, X., Zhang, Y., Ba, Z., Kyritsis, N., Casellas, R., and Alt, F.W. (2019a). Fundamental Roles of Chromatin Loop Extrusion in Antibody Class Switching. Nature 575, 385–389. https://doi.org/10.1038/s41586-019-1723-0.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biol 9, R137–R137. https://doi.org/10.1186/gb-2008-9-9-r137.

Zhang, Y., Zhang, X., Ba, Z., Liang, Z., Dring, E., Hu, H., Lou, J., Kyritsis, N., Zurita, J., Shamim, M.S., et al. (2019b). The Fundamental Role of Chromatin Loop Extrusion in Physiological V(D)J Recombination. Nature 573, 600–604. https://doi.org/10.1038/s41586-019-1547-y.

Zheng, H., and Xie, W. (2019). The role of 3D genome organization in development and cell differentiation. Nat Rev Mol Cell Bio 20, 535–550. https://doi.org/10.1038/s41580-019-0132-4.

Zheng, X., and Zheng, Y. (2017). CscoreTool: fast Hi-C compartment analysis at high resolution. Bioinformatics 34, 1568–1570. https://doi.org/10.1093/bioinformatics/btx802.

Zhou, G.-L., Xin, L., Song, W., Di, L.-J., Liu, G., Wu, X.-S., Liu, D.-P., and Liang, C.-C. (2006). Active Chromatin Hub of the Mouse α-Globin Locus Forms in a Transcription Factory of

Clustered Housekeeping Genes. Mol Cell Biol 26, 5096–5105. https://doi.org/10.1128/mcb.02454-05.

Zhu, Y., Denholtz, M., Lu, H., and Murre, C. (2020). Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization. Gene Dev 35, 65–81. https://doi.org/10.1101/gad.343475.120.