Exploring Mitotic Chromosomes; from epigenetic bookmarking to sister chromatid conformation

Marlies E. Oomen

Printed by Gildeprint (the Netherlands)
Cover design and layout by Marlies E. Oomen

The research presented in this thesis was performed at the Program in Systems Biology at University of Massachusetts Medical School, Worcester MA, United States of America.

Exploring Mitotic Chromosomes; from epigenetic bookmarking to sister chromatid conformation

---

Een expeditie naar mitotische chromosomen;
Van epigenetische boekmarkeringen naar
de conformatie van zusterchromatiden

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the rector magnificus

Prof.dr. F.A. van der Duijn Schouten

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on
Wednesday July 7th at 15.30hrs

by

Maria Elisabeth Oomen
born in Utrecht, Netherlands

Erasmus University Rotterdam

Doctoral Committee

Promotors	Prof. dr. Job Dekker Prof. dr. Bas van Steensel
Other members	Prof. dr. Wouter de Laat Dr. Kerstin Wendt Prof. dr. Niels Galjart

Voor Grootma en Oma Tok
Dedicated to my grandmothers

Table of content

Chapter 1	Introduction and scope of this thesis	7
Chapter 2	Epigenetic characteristics of the mitotic chromosome in 1D and 3D	15
Chapter 3	CTCF sites display cell cycle-dependent dynamics in factor binding and nucleosome positioning	49
Chapter 4	Differences in mitotic chromosome organization between cell types and species	93
Chapter 5	Detecting chromatin interactions between and along sister chromatids with SisterC	113
Chapter 6	SisterC 2.0: Optimizing SisterC for application in mammalian cells	151
Chapter 7	General Discussion	168
Addendum		179
	Summary	180
	Nederlandse samenvatting	183
	Abbreviations	186
	Curriculum vitae	187
	List of publications	188
	PhD portfolio	189
	Acknowledgements	191

Chapter 1

Introduction and scope of this thesis

Introduction

When high-school students learn about chromosomes and DNA in biology class, it is often visualized with the iconic images of condensed mitotic chromosomes. It is hard to imagine that all genomic information is captured inside those x-shaped structures. However, mitotic chromosomes do not only grasp the attention of teenagers, they have intrigued many generations of scientists over the past centuries[1–4]. Studying how chromosome are folded in general, and in mitosis in particular, is key to gain understanding on gene regulation, genome stability and inheritance of chromosome characteristics throughout cell cycles and even between generations[5–9]. Traditionally observed by microscopy, the rise of genomics techniques created many new possibilities to study chromosomes in general and mitotic chromosomes in particular. As technology development continuously brings new ways to study chromosomes, our insights are ever changing as well. For example, in this thesis research we use several techniques that were developed in the last decade; probing for chromatin accessibility using ATAC-seq[10], histone modifications and chromatin binding proteins using Cut&Run[11] and chromosome organization by Hi-C[12].

In interphase, chromosomes are characterized by cell type specific transcription patterns, epigenetic features, such as histone modifications and chromatin binding factors, and chromosome conformation in compartments and topologically associating domains (TADs)[13,14]. The mechanisms that regulate these chromatin characteristics are largely conserved across vertebrates and eukaryotes in general[15,16]. However, the exact genomic locations of these features can be specific to the species, the cell type and even the state of the cell[12–14,17–23]. Strikingly, as the cell progress into mitosis, these chromatin characteristics change dramatically; transcription is halted[24], histones and many other proteins become heavily phosphorylated[25], the majority of chromatin binding factors lose binding[26] and chromosome conformation is rearranged into helical loop arrays[27,28]. However, after cells complete mitosis and cells enter G1, chromosomes need to quickly reestablish the interphase chromatin characteristics[7,29,30]. To aid the memory of cell type identity during mitosis, certain chromatin characteristics are maintained throughout the cell cycle by leaving epigenetic bookmarks in the genome. These bookmarks can consist of remaining chromatin factors[31,32], histone modifications and variants[25], DNA methylation[33], chromatin accesiblity[34,35] and even non-coding RNAs[36]. We will discuss chromatin characteristics of mitotic cells and the different mechanisms of mitotic bookmarking in greater depth in **chapter 2**[37].

With many epigenomics consortium efforts focusing on interphase cells, we gained a tremendous amount of information about the epigenetic characteristics and the chromatin binding factors that regulate transcription and cell type identify in

interphase[13,14,17]. However, there are many open questions regarding chromosome characteristics in mitosis. For example, it is currently still unknown for several important interphase chromatin factors whether they maintain binding during mitosis. Furthermore, for factors and histone modifications that have been found to be maintained throughout the cell cycle, it is unclear whether their interphase function is maintained in mitosis. In **chapter 3**, we describe the cell cycle dynamics of chromosome architecture protein CTCF. Combining both genomics and imaging techniques, we find that CTCF binding is lost in human somatic cells during mitosis[38]. Strikingly, we find that histone modifications H3K4Me1 and H3K4Me3, and histone variant H2A.z are maintained around CTCF sites. This could possibly function as a mitotic bookmark, which enables rebinding of CTCF upon mitotic exit and the formation of CTCF mediated loops and TADs as cells enter G1.

Recently, there have been several contradicting reports on mitotic chromatin characteristics. Some of these results can be due to technical differences between studies. For example, it has been reported that cross-linking conditions can significantly influence observations using techniques such as immunofluorescence microscopy as well as ChIP-seq[39,40]. Perhaps more interestingly, these discrepancies could also be due to biological variance between cell types and species. We compare CTCF binding dynamics between different cell types and mitotic chromosome conformation in general between species in **chapter 4**. We find that, in contrast to human somatic cell lines, 30-50% of all interphase bound CTCF sites maintain binding during mitosis in mouse stem cells (mESCs). Although CTCF binding is partially maintained, we do not observe any CTCF-mediated chromosome structures, such as TADs and loops. This suggests that interphase function of CTCF is not maintained in mitosis. Additionally, we compare the loop arrays by which mitotic chromosomes are organized in chicken, mouse and human cell lines. We show that different species have different size loops in mitosis. The loop size is correlated with the average size of the q-arm in each species, where chicken chromosomes have both the smallest loop size and smallest q-arm on average and mouse shows both the largest loops and largest average q-arm length. This result could be interpreted that chromosomes with longer arms compact to greater extend by forming larger loops, in order to allow for proper separation of the sister chromatids in anaphase.

In addition to maintenance of cell type specific characteristics, chromosomes face a second major challenge the cell cycle. After DNA has been replicated in S-phase, the newly formed sister chromatids are severely entangled[41]. Although close proximity of sister chromatids to each other enables a pathway for DNA damage repair during G2-phase[42], it also creates a substantial topological problem when sister chromatids separate in anaphase[43]. In order to prevent potential catastrophic

events such as cell cycle arrest or even aneuploidy, the sister chromatids need to completely disentangle and condense as the cells progress through mitosis[44]. This requires that the machineries which align the sister chromatids orchestrate their actions with the machineries that form the loop array while shaping properly condensed and detangled mitotic chromosomes[45].

It remains however difficult to study the organization of sister chromatids by genomics techniques. By definition, sister chromatids have identical sequences, as they have been replicated from the same DNA template during S-phase. In recent years, a microscopy technique has been developed that allows visualization sister chromatids independently by using a two-color labelling system[46]. Additionally, the genomics technique Strand-seq allows for detection of sister chromatid exchange events[47]. It was however not yet possible to differentiate the interactions between and along sister chromatids by Hi-C. In **chapter 5**, we describe a novel Hi-C technique, SisterC, and apply this technique to study sister chromatid organization in mitotic budding yeast[48]. In yeast, where both intra-sister extruding loops and inter-sister connection are mediated by cohesin, we find that extruding cohesin and cohesive cohesin act independently. This allows for compaction of chromosomes during mitosis, while also maintaining a degree of alignment between the sister chromatids. Additionally, we propose strategies to optimize our SisterC technique and adapt the protocol for use in human cell line Hap1 in **chapter 6**, which will allow the study sister chromatid organization as cells progress from late S-phase up to anaphase.

Finally, in **chapter 7**, I will discuss our findings in relation to other recent publications and propose future directions. As every researcher knows, new scientific findings can bring answers to outstanding questions, but possibly more frequently, give rise to new research questions. In this thesis, I present my exploration expedition to mitotic chromosomes. During my PhD work, I was incredibly fortunate to be surrounded by an amazing group of 3C experts[12,49–51], test my hypotheses using brand new genomics techniques, such as ATAC-seq[10] and Cut&Run[11], collaborate with leaders in the field of cutting-edge imaging approaches[52] and go through the struggles of developing a new technique, SisterC[48,53]. By investigating mitotic chromosomes from chromosome folding on megabase scale to epigenetic bookmarks at nucleosome resolution, I hope this thesis brings new insights and raises several research questions that will instigate the next scientific expeditions to the characteristics of mitotic chromosomes.

Scope of this thesis

Chapter 2 gives a literature overview which introduces the field of epigenetic characteristics of mitotic chromosomes and their organization in 3D.

Chapter 3 presents a study to binding dynamics and epigenetic modifications at CTCF motifs throughout the cell cycle.

Chapter 4 describes CTCF binding dynamics and chromosome organization in different cell types and species in prometaphase

Chapter 5 introduces a novel Hi-C technique, SisterC, which allows for the detections of chromatin interactions between and along sister chromatids.

Chapter 6 describes a strategy to optimize the SisterC protocol and adapt the technique for use in mammalian cells.

Chapter 7 offers a general discussion of the research presented in this thesis and closes with current challenges and future research directions.

References

1. Flemming, W. Zur Kenntnis der Zelle und ihrer Teilung-Erscheinungen. *Schr. Nat. Wiss. Ver. Schlesw.-Holst.* **3**, 23–27 (1878).
2. Earnshaw, W. C. & Laemmli, U. K. Architecture of metaphase chromosomes and chromosome scaffolds. *J. Cell Biol.* **96**, 84–93 (1983).
3. Maeshima, K. & Laemmli, U. K. A Two-step scaffolding model for mitotic chromosome assembly. *Dev. Cell* **4**, 467–480 (2003).
4. Marsden, M. P. F. & Laemmli, U. K. Metaphase chromosome structure: Evidence for a radial loop model. *Cell* **17**, 849–858 (1979).
5. Collombet, S. *et al.* Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature* **580**, 142–146 (2020).
6. Flyamer, I. M. *et al.* Single-cell Hi-C reveals unique chromatin reorganization at oocyte-tozygote transition. *Nat. Publ. Gr.* 1–17 (2017). doi:10.1038/nature21711
7. Abramo, K. *et al.* A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. *Nat. Cell Biol.* **21**, 1393–1402 (2019).
8. Gibcus, J. H. & Dekker, J. The hierarchy of the 3D genome. *Mol. Cell* **49**, 773–82 (2013).
9. Sanyal, A., Lajoie, B. R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature* **489**, 109–13 (2012).
10. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–8 (2013).
11. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* 1–35 (2017). doi:10.1101/097188
12. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–93 (2009).
13. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
14. Consortium Roadmap Epigenomics *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
15. Lowdon, R. F., Jang, H. S. & Wang, T. Evolution of Epigenetic Regulation in Vertebrate Genomes. *Trends Genet.* **xx**, 1–15 (2016).
16. Hirano, T. Condensin-Based Chromosome Organization from Bacteria to Vertebrates. *Cell* **164**, 847–857 (2016).
17. Oksuz, B. A., Yang, L., Abraham, S., Venev, S. V & Krietenstein, N. Systematic evaluation of chromosome conformation capture assays. 0–42 (2020).
18. Rao, S. S. P. *et al.* A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* **159**, 1665–1680 (2014).
19. Hnisz, D. *et al.* Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
20. Smith, E. M., Lajoie, B. R., Jain, G. & Dekker, J. Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the CFTR Locus. *Am. J. Hum. Genet.* **98**, 185–201 (2016).
21. Deng, W. *et al.* Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **158**, 849–860 (2014).
22. Vernimmen, D., De Gobbi, M., Sloane-Stanley, J. a, Wood, W. G. & Higgs, D. R. Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* **26**, 2041–2051 (2007).
23. Lupiáñez, D. G., Spielmann, M. & Mundlos, S. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends Genet.* **32**, 225–237 (2016).
24. Prescott, D. M. & Bender, M. DNA synthesis and mitosis in cultures of human peripheral leukocytes. *Exp. Cell Res.* **27**, 221–9 (1962).

25. Wang, F. & Higgins, J. M. G. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol.* **23**, 175–84 (2013).
26. Martinez-Balbas, M. A., Dey, A., Rabindran, S. K., Ozato, K. & Wu, C. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* **83**, 29–38 (1995).
27. Gibcus, J. H. *et al.* A pathway for mitotic chromosome formation. *Science (80-.).* **359**, eaao6135 (2018).
28. Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342**, 948–53 (2013).
29. Zhang, H. *et al.* Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* **576**, 158–162 (2019).
30. Pelham-Webb, B. *et al.* Mitotic retention of H3K27 acetylation promotes rapid topological and transcriptional resetting of stem cell-related genes and enhancers upon G1 entry. *bioRxiv* (2020). doi:10.1101/2020.06.02.130104
31. Festuccia, N. *et al.* Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nat. Cell Biol.* **18**, (2016).
32. Owens, N. *et al.* CTCF confers local nucleosome resiliency after dna replication and during mitosis. *Elife* **8**, 1–26 (2019).
33. Wigler, M., Levy, D. & Perucho, M. The somatic replication of DNA methylation. *Cell* **24**, 33–40 (1981).
34. Hsiung, C. C. *et al.* Genome accessibility is widely preserved and locally modulated during mitosis. 1–29 (2015). doi:10.1101/gr.180646.114
35. Oomen, M. E., Hansen, A., Liu, Y., Darzacq, X. & Dekker, J. CTCF sites display cell cycle dependent dynamics in factor binding and nucleosome positioning. *bioRxiv* 365866 (2018). doi:10.1101/365866
36. Meng, Y. *et al.* The non-coding RNA composition of the mitotic chromosome by 5′-tag sequencing. *Nucleic Acids Res.* gkw195 (2016). doi:10.1093/nar/gkw195
37. Oomen, M. E. & Dekker, J. Epigenetic characteristics of the mitotic chromosome in 1D and 3D. *Crit. Rev. Biochem. Mol. Biol.* **0**, 1–20 (2017).
38. Oomen, M. E., Hansen, A. S., Liu, Y., Darzacq, X. & Dekker, J. CTCF sites display cell cycle–dependent dynamics in factor binding and nucleosome positioning. *Genome Res.* 1–14 (2019). doi:10.1101/gr.241547.118.
39. Festuccia, N. *et al.* Transcription factor activity and nucleosome organization in mitosis. *Genome Res.* **29**, 250–260 (2019).
40. Teves, S. S. *et al.* A Dynamic Mode of Mitotic Bookmarking by Transcription Factors. *Elife* 066464 (2016). doi:10.1101/066464
41. Peters, J.-M. & Nishiyama, T. Sister chromatid cohesion. *Cold Spring Harb. Perspect. Biol.* **4**, 1–18 (2012).
42. Johnson, R. D. & Jasin, M. Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *EMBO J.* **19**, 3398–3407 (2000).
43. Goloborodko, A., Imakaev, M. V., Marko, J. F. & Mirny, L. Compaction and segregation of sister chromatids via active loop extrusion. *Elife* **5**, 1–20 (2016).
44. Nasmyth, K. Separating sister chromatids. *Trends Biochem. Sci.* **24**, 98–104 (1999).
45. Yatskevich, S., Rhodes, J. & Nasmyth, K. Organization of Chromosomal DNA by SMC Complexes. 1–38 (2019).
46. Nagasaka, K., Hossain, M. J., Roberti, M. J., Ellenberg, J. & Hirota, T. Sister chromatid resolution is an intrinsic part of chromosome organization in prophase. *Nat. Cell Biol.* (2016). doi:10.1038/ncb3353
47. Falconer, E. *et al.* DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nat. Methods* **9**, (2012).
48. Oomen, M., Hedger, A., Watts, J. & Dekker, J. Detecting chromatin interactions along and between sister chromatids with SisterC. *Nat. Methods* (2020).

doi:10.1101/2020.03.10.986208
49. Belaghzal, H., Dekker, J. & Gibcus, J. H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* **123**, 56–65 (2017).
50. Dostie, J. & Dekker, J. Mapping networks of physical interactions between genomic elements using 5C technology. *Nat. Protoc.* **2**, 988–1002 (2007).
51. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–11 (2002).
52. Hansen, A. S. *et al.* Robust model-based analysis of single-particle tracking experiments with Spot-On. *Elife* **7**, 1–33 (2018).
53. Hedger, A. K. *et al.* Progress toward an amplifiable metabolic label for DNA: Conversion of 4-thiothymidine (4sT) to 5-methyl-2′-deoxycytidine and synthesis of a 4sT phosphorodiamidate prodrug. *Can. J. Chem.* **96**, 636–645 (2018).

Chapter 2

Epigenetic Characteristics of the Mitotic Chromosome in 1D and 3D

Marlies E. Oomen[1] & Job Dekker[1,2]

Published in *Critical Reviews in Biochemistry and Molecular Biology.* (2017) Volume 52 - Issue 2

1. Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 368 Plantation Street, Worcester, MA, 01605-0103, USA
2. Howard Hughes Medical Institute, Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 368 Plantation Street, Worcester, MA, 01605-0103, USA.

Abstract

While chromatin characteristics in interphase are widely studied, characteristics of mitotic chromatin and their inheritance through mitosis are still poorly understood. During mitosis chromatin undergoes dramatic changes: Transcription stalls, chromatin binding factors leave the chromatin, histone modifications change, and chromatin becomes highly condensed. Many key insights into mitotic chromosome state and conformation have come from extensive microscopy studies over the last century. Over the last decade the development of 3C-based techniques has enabled the study of higher order chromosome organization during mitosis in a genome-wide manner. During mitosis chromosomes lose their cell type specific and locus-dependent chromatin organization that characterizes interphase chromatin and fold into randomly positioned loop arrays. Upon exit of mitosis cells are capable of quickly rearranging the chromosome conformation to form the cell type specific interphase organization again. The information that enables this rearrangement after mitotic exit is thought to be encoded at least in part in mitotic bookmarks, e.g. histone modifications and variants, histone remodelers, chromatin factors and non-coding RNA. Here we give an overview of the chromosomal organization and epigenetic characteristics of the interphase and mitotic chromatin in vertebrates. Second, we describe different ways in which mitotic bookmarking enables epigenetic memory of the features of the interphase chromatin through mitosis. And third, we explore the role of epigenetic modifications and mitotic bookmarking in cell differentiation.

Introduction

A major question in cell biology is how cell type identity is maintained through mitosis. Imaging studies have been instrumental in studying mitotic chromosomes in live cells and after purification. These pioneering studies, mostly by the Laemmli group, led to fundamental insights into the architecture of mitotic chromosomes[1–3]. More recently, high-throughput genomic methods have been used to gain deeper and more detailed insights into the folding of chromatin inside mitotic chromosomes and the local characteristics of the chromatin fiber such the presence of open sites, patterns of histone modifications and the binding of other factors[4–6].

Decades of genetic and epigenetic studies have revealed many features of chromosome structure and how these could be involved in transcriptional control in the interphase cell. Over the last decade emerging high-throughput sequencing techniques like chromosome conformation capture (3C) based techniques, assays for transposase-accessible chromatin using sequencing (ATAC-seq), DamID and chromatin immunoprecipitation sequencing (ChIP-seq) enable the study of chromosome conformation, nuclear organization, chromatin state, its function and its regulators[7–11]. Large-scale consortia like the Encode project and NIH epigenome roadmap provide comprehensive overviews of cell type specific profiles of histone modifications, nuclear organization and DNA binding factors in non-synchronous, mostly interphase, cells[12,13]. These cell type specific features establish regulatory control of the genome and its effects on the phenotype of a cell. However, the characteristics of vertebrate chromatin change dramatically during mitosis. Chromosome conformation transforms from a cell type specific to a universal condensed organization, many chromatin factors and the transcription machinery are thought to no longer bind to the DNA, nuclear envelope and therefore lamina interactions disintegrate and new histone modifications specific for mitosis are deposited. After mitosis, chromatin returns to its uncondensed cell type specific shape, chromatin factors are bound again, the nuclear envelope and lamina interactions are restored and the histone modification pattern specific for interphase is reestablished[4,6,14,15]. However, for many of these changes in the vertebrate mitotic chromatin it is unknown how, with which function and in which order they occur and how the interphase chromatin state is re-formed upon mitotic exit (figure 1).

Mitosis has been an area of interest for over a century since condensation of chromosomes was first observed by microscopy. For many decades the main focus was the study of the mitotic chromatin through different microscopy techniques like FISH and immunofluorescence to localize chromatin proteins[16]. Because of the clear morphological features of mitotic cells, it is relatively easy to single out cells for study using microscopy. A downside of studies using these techniques is the

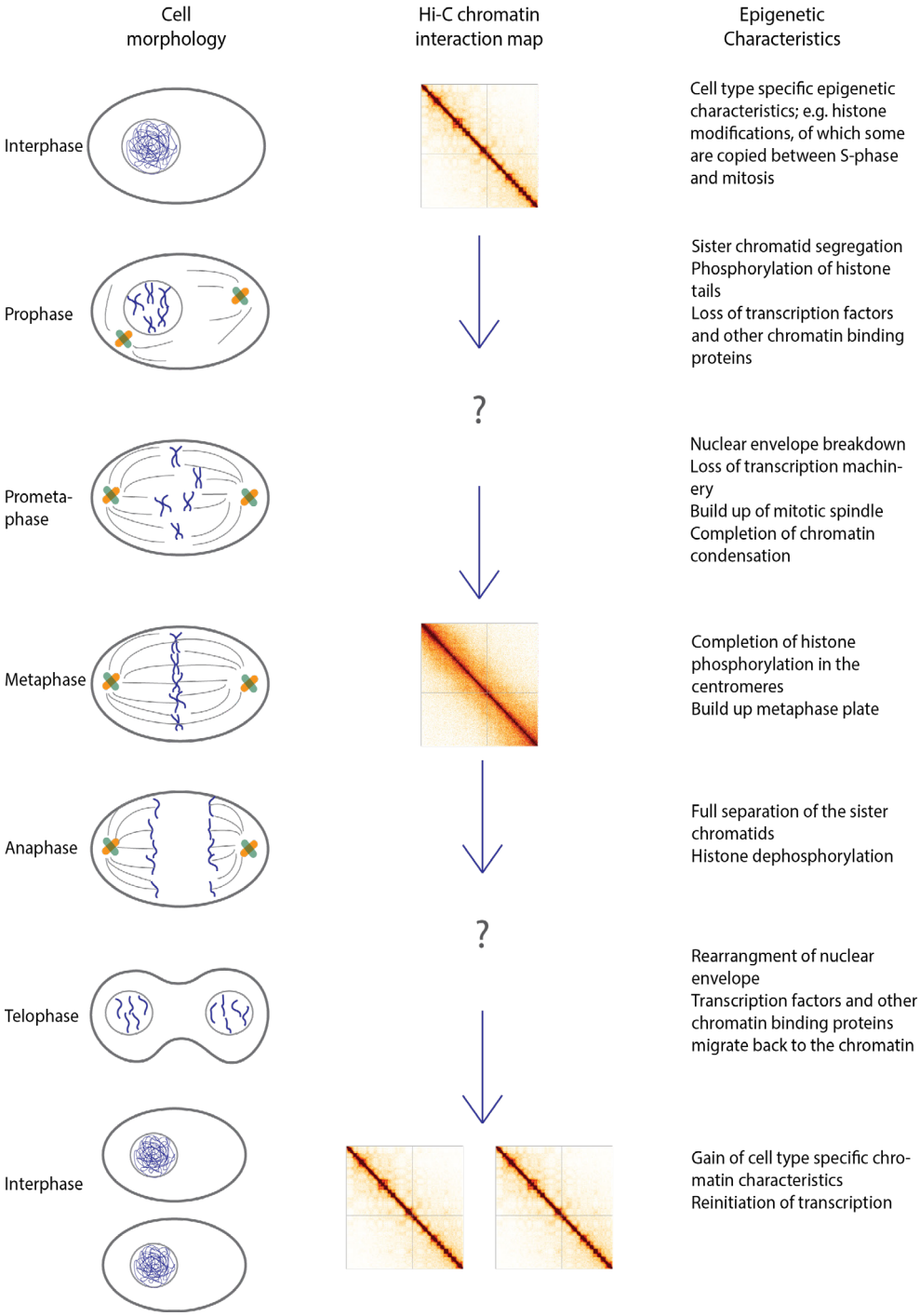


Figure 1. Overview of cell morphological changes, changes in the chromatin organization and changes known epigenetic characteristics during the different phases of mitosis. Hi-C data own is in HeLa cells and was previously published in Naumova et al.[6]

limitations of the scale of the experiment, since it is not possible to do genome-wide experiments (probing the position of all loci) using microscopy. The development of high-throughput sequencing techniques has opened new ways to study mitotic chromosomes. These methods enable genome-wide detection of chromatin state, and the mapping of chromatin structure to specific sequences. However, these methods have their own set of limitations. Most particularly, these methods do not analyze single cells, but determine population-averaged features. For this, they typically require large numbers of cells, which means that for cell cycle studies one has to carefully synchronize large cell cultures in the cell cycle phase of interest[17–20]. When doing such population-based studies, one needs to obtain samples of a homogenous population. Although synchronization protocols have been optimized over the years, it is good to keep in mind that it remains difficult to obtain a fully synchronized population and that heterogeneity in the population can be the cause of inconsistencies between different studies and contamination with unsynchronized cells can reduce the quality of the obtained data.

Here we review and discuss chromosome conformation in interphase and mitosis and explore how epigenetic information can be contained within the local and global organization of chromatin. While we focus on vertebrate chromosomes, there is wealth of data on these phenomena in plants as well. We refer the reader to several key publications for those studies[21–23].

Chromatin folding in Interphase and Metaphase

The fact that chromosomes do not simply consist of floating linear strands of DNA has been known since their discovery. In fact, chromosomes were first observed because of their dramatic condensation during mitosis, which allowed their visualization by microscopes of that time, described by Walther Flemming in the late 1800s[24,25]. For decades chromosomes and chromatin were studied by microscopy and techniques like X-ray crystallography. In the era of molecular biology and the development of sequencing, the research focus shifted towards unraveling the human genome by sequence and the concept of chromosome structure and conformation became less studied. However, it is clear other factors beyond DNA sequence contain instructions for the cell. The structural and physical organization of chromosomes inside the nucleus is an important carrier of information, which is important in many processes such as gene expression regulation and is in part specific for cell type identity[26–28].

Eukaryotic chromatin is organized on different levels which are represented in figure 2a as cartoons and as observations of these organization levels in Hi-C heatmaps of interphase HeLa cells represented in figure 2b (previously published data in a study by Naumova et al[6]). As interphase chromosomes are too large to

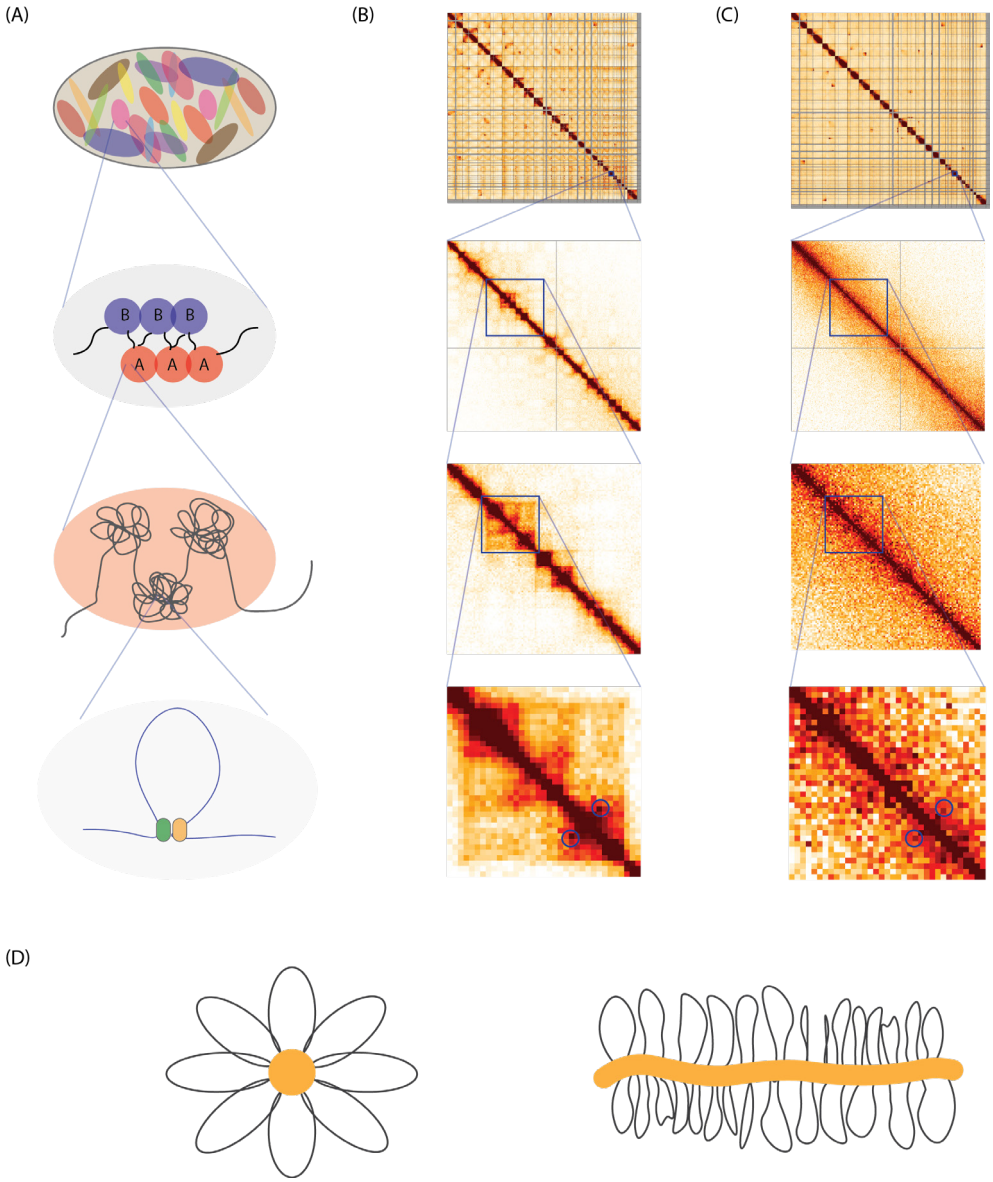


Figure 2. (**a**) Cartoon representation of chromosome territories (panel 1), compartments (panel 2), TADS (panel 3) and chromosome loops (panel 4). (**b**) Representation of the different chromosome organization levels in Hi-C heatmaps interphase of HeLa cell. Panel 1 shows chromosome territories in a heatmap showing all chromosomes. Panel 2 shows multiple compartments in a zoom in to the right arm of chromosome 18. Panel 3 shows a zoom in to 3.1-49.6 Mb of chromosome 18 representing multiple TADs. Panel 4 represents a possible looping interaction in a zoom in to chromosome 18 34.5-39.8 Mb. (**c**) Hi-C heatmaps of the same regions shown in (b) but for metaphase HeLa cells. Hi-C data shown was published in Naumova et al.[6] (**d**) Model of the bottle brush polymer conformation of the mitotic chromosome suggested by Naumova et al.[6] and Goloborodko et al.[31]

freely diffuse inside the nucleus, they occupy their own territories (figure 2a-b, first panel). These individual chromosome territories were already observed in the 1990s and have been confirmed with microscopy and chromosome conformation capture techniques[9,29]. Chromosomes can interact with neighboring chromosomal territories by looping part of one chromosome into another chromosome territory[30]. As a result of chromosome territories interchromosomal interactions are much less frequent than interactions between loci located on the same chromosome. The organization of chromosome territories within the nucleus is highly conserved between cell types and even across species[32].

The next layer of interphase chromosomal conformation is the organization in distinct sub-nuclear compartments where active and inactive regions, both from the same chromosome and occasionally from different chromosomes, cluster together (figure 2a-b, second panel)[9,33]. These regions are categorized as A or B compartments; A compartments are described as sub-nuclear neighborhoods of active and transcribed loci and B compartments mostly contain inactive regions. The compartmentalization correlates with other well-studied characteristics of the chromatin, such as gene expression, genome accessibility and histone marks: A-compartments are highly correlated with open, active, euchromatic regions and B compartments mostly contain loci of closed, inactive, heterochromatin regions. The organization in compartments leads to spatial separation of active loci from inactive loci, preventing interactions between active and inactive regions which may help prevent heterochromatin spreading. Compartmentalization is cell type specific[9,34], likely because gene expression and chromatin modifications are cell type specific. Some studies show that the organization in active and inactive compartments is correlated with an increased concentration of factors involved in the regulation of these regions. An example of such an enrichment can be found in so-called transcription factories; a concept introduced by Iborra et al[35]. Transcription factories are regions in the nucleus, where active genes and the transcription machinery are concentrated. This mechanism of enrichment would fit with an organization into A compartments. However, it is still to be determined whether compartmentalization in A and B compartments is cause or consequence of these proposed mechanisms.

The third layer of chromosomal organization in interphase is the organization of DNA into topologically associated domains, TADs (figure 2a and b, third panel)[36,37]. TADs are defined as contiguous genomic regions, typically several hundred kilobases in size that show elevated levels of self-interactions. TADs are separated from each other by boundaries that prevent interactions between loci located in adjacent TADs[38–40]. TAD boundaries are enriched for certain protein binding sites. It is still under debate which proteins are involved in defining TAD boundaries and

which genetic features define a TAD region, however some characteristics of TAD boundaries are well described, such as the CTCF binding motif[36]. These motifs are bound by structure-mediating proteins like CTCF and cohesin, which can act as an insulator between TADs[34,41–44]. Since CTCF binding motifs are genetically defined, TAD boundaries are regarded as universal features of all cell types[36]. Cell-type invariant TADs that can be located far apart on the same chromosome or on different chromosomes cluster together forming the cell type-specific A- and B-type compartments described above, depending on their chromatin and expression state in a given cell type. Therefore, TADs have been proposed to be basic building blocks of chromosome and nuclear organization[45].

At the scale of up to tens of Kb, chromatin interactions are organized in DNA loops (figure 2a-b fourth panel). Looping between two loci on a 10 to 100 kb scale, enables direct interaction between for example an enhancer and a promoter[34,42,46,47]. Many looping interactions appear to occur between loci located within the same TAD or insulated domain[48–50]. Even though TAD organization is often conserved among cell types, looping interactions can be cell type specific, but do rarely seem to cross TAD boundaries[50]. The presence of a looping interaction can alter gene expression in a cell type specific manner by bringing promoters in close proximity to promoters[47,51–53]. Looping interactions have been suggested to function as a fine-tuning mechanism of (transcriptional) regulation. It has been shown for several diseases, like cancer and polydactyly, how alteration in looping interactions is associated with the disease phenotype[48,54,55]. To summarize, interphase chromatin architecture is established by clustering of sets of TADs into cell type specific A and B compartments. The cell type specific regulation of a certain locus within a TAD is then enabled by forming loops of for example regulating loci with an effector locus.

This partly cell type specific conformation of interphase chromosomes changes dramatically during mitosis. Observations using microscopy techniques already proved the dramatic changes many decades ago. The distribution of the DNA in the nucleus changes from amorphous territories to elongated rode shaped structures, with the characteristic banding pattern upon staining[56,57]. It was shown that these mitotic bands along the chromosomes are the same between different cell types[58]. Both the dramatic change in shape and the cell type indifferent band patterns already suggested loss of the higher organization in compartments and TADs known for interphase chromosomes, as the rod shape and the individualization of the chromosomes do not allow for interaction within and between compartments. Using 5C and Hi-C-techniques Naumova et al. observed these dramatic changes in a genome wide manner, represented in figure 2c. Clearly, cells are capable of reestablishing the same chromosomal organization in early G1 phase as was

present before mitosis[6]. A recent study of Hsuing et al measured enhancer promoter interactions in interphase and mitosis using Capture-C, a technique to capture interactions anchored at hundreds of loci at the same time[5,59]. For the promoter enhancer pairs measured in this study, it was found that they are specific to interphase, and show to have largely reduced interaction frequencies in mitotic cells. This suggests that chromosomal organization in loops between functional elements is also impaired in mitosis, and that these interactions must be re-established in the next G1 phase.

Folding of the mitotic chromosome

Although the folding characteristics of interphase chromatin appear almost completely lost during mitosis, this does not mean mitotic chromatin has no higher order organization. The mitotic chromosome of vertebrates condenses 2-3 times in volume compared to interphase[60]. The prevailing model for mitotic chromosome architecture that is supported by pioneering microscopy studies by Laemmli and co-workers, more recent 5C and Hi-C analyses and polymer modeling is that chromosomes fold as longitudinally compressed arrays or stochastically positioned consecutive chromatin loops[1,6]. Furthermore, during the condensation process, sister chromatids are separated and individualized during pro- and prometaphase to accommodate proper division over the two new daughter cells[61–63].

The main machineries that drive mitotic chromosome morphogenesis are condensin complexes, including condensin I and II[64–66]. It was found that depletion of subunits of the condensin complexes delays the condensation process and progression of mitosis[66,67]. Although the exact working mechanism is not yet completely understood, the prevailing model is that condensin functions as a ring-like structure that keeps two strands of DNA from the same chromosome together[68]. Condensins are abundant chromosomal protein complexes that are located along the central axis of the mitotic chromosome, consistent with its proposed role to organize the chromosome as a long array of chromatin loops. Whereas condensin II is present in the nucleus throughout the cell cycle and mainly exerts its mitotic function in prophase, condensin I mostly gains access to the chromatin upon nuclear envelope breakdown in prometaphase[67].

A model introduced by Nasmyth proposes that protein complexes like condensins are capable of forming loops that progressively become larger until further progression is blocked by other chromatin bound factors (e.g. other condensins)[69]. This idea is further elaborated in loop extrusion models, proposed in earlier form by Arthur Riggs over 25 years ago[70], and recently developed to describe chromatin interaction data obtained with Hi-C[38,40,71,72]. Computational modeling and experimental

studies applied extrusion models for understanding formation of interphase chromatin domains[28,38,40,72,73]. Naumova et al. proposed that a loop extrusion process could lead to formation of the mitotic chromatin loop array[6]. One of the most recent studies directly tested the loop extrusion model for mitotic chromosome formation[31]. During the mitotic condensation process condensin complexes concentrate onto the chromatid axis and as interphase-specific boundary elements like CTCF dissociate from the mitotic chromatin, loop extrusion can occur unimpeded, which causes the DNA to progressively condense. Using this model, Goloborodko et al were able to computationally predict the condensation process into prophase-like chromosomes that show the same characteristics as was observed in experimental studies[1,6]. Loop-extrusion leads to arrays of stochastically positioned consecutive loops. General chromatin attraction leads to further loop condensation and stacking of loops on top of each other. The physical characteristics of DNA predict steric repulsion between loops which results in a bottlebrush-like organization of the DNA with loops arranged as rosettes around a more centrally located axis of loop bases, represented in figure 2d. This is a structure very reminiscent of the radial loop model for mitotic chromosomes proposed many years ago by Laemmli and co-workers[1,74]. This loop repulsion also mediates separation of the sister chromatids after disentanglement is established, which will be described later in this section. Although the loop extrusion model can explain many features of the condensing chromatin, it does not explain the higher order rod-like organization chromosomes in mitosis. The forces involved in the loop-extrusion model would cause a complete condensation and collapse of the DNA into a ball shape, caused by the chromosomes sticking to each other, if there would not be an external agent that functions as a surfactant coating the mitotic chromosomes to prevent multiple mitotic chromosomes sticking together[75]. A recent study suggested that the positively charged protein Ki-67 could act as such an agent to keep the mitotic chromosomes separated[76]. Ki-67 is associated with nucleoli in interphase cells, however it has been known for many years that Ki-67 is associated with the outside of the mitotic chromosome[77]. This makes Ki-67 a good candidate to be a key mediator in the individualization of mitotic chromosomes.

Another key player that regulates the organization of the mitotic chromosome is the cohesin complex, which holds sister chromatids together until late mitosis when the sister chromatids are separated and the DNA is divided between the new daughter cells[78–81]. Sister chromatids are tightly intertwined after DNA replication in S-phase up till prophase. This implies that the initial stages of condensation occur while the sister chromatids are entangled and interacting. However, sister chromatids need to be separated to enable proper segregation, equal distribution of chromosomes over the daughter cells and to prevent DNA breaks caused by the strong forces

on the DNA by the mitotic spindle. It has been known that cohesin removal from the mitotic chromosome has two distinct pathways[82]. Cohesin is first removed from the chromosome arms during prophase; however, localization of cohesin at the centromeres is maintained. This enables coupled progression of the condensation and sister separation of the chromosome arms, while they are still attached at the centromeres necessary for metaphase plate alignment. Cohesin complexes located at the centromeres are removed after completion of the condensation process and formation of the mitotic spindle in metaphase to anaphase transition[82]. However, the question how the topological entanglement of the sister chromatids is resolved after removal from cohesin in prophase, remains to be addressed. A recent study from Liang et al[83] showed how topoisomerase IIα functions to disentangle the sister chromatids. In addition to this, Goloborodko et al showed in their computational study how addition of topoisomerase II-activity to the loop extrusion model leads to separation of the sister chromatids[31]. The function of topoisomerase in sister segregation was later also confirmed using a labeling technique to visualize the chromatids separately by super resolution microscopy[63]. It was shown that sister chromatid separation is initiated in early prophase, however further condensation is halted when topoisomerase IIα is inhibited.

Chromatin condensation not only occurs at the level of chromatin loops and whole chromosomes, at the nucleosomal level changes are observed during mitosis as well. As many chromatin binding factors are thought to migrate off the chromatin during mitosis, the nucleosomes can rearrange their relative positioning. Although it is unknown how the activity and movement of ring-like complexes like condensin, as proposed for loop extrusion, are affected by the presence nucleosomes on the DNA, it has been shown that nucleosomes are evicted from the chromatin in order for condensin to load and start forming loops[84]. One can imagine that the forces on the chromatin fiber during the condensation process might cause the nucleosomes to redistribute along the chromatin. Using an electron microscopy technique called EM-assisted nucleosome interaction capture (EMANIC), Grigoryev et al were able to capture hierarchical looping of nucleosome chains that are ordered in a zig-zag like fashion[85]. The zigzag stacking of nucleosomes is also present in interphase chromatin, however the higher order organization of loops of nucleosomes appears to be unique for mitosis[86]. This emphasizes the importance of nucleosomes, their positioning and their modifications in mitosis, which will be more elaborately discussed below.

Chromatin accessibility in mitosis

Mitotic chromatin is highly condensed and appears to have universal organization

as detected by Hi-C. This does not mean that there are no differences in levels of local condensation and chromatin accessibility along the mitotic chromosome, as described for the interphase chromatin. A study by Hsuing et al observed variability in the accessibility of the mitotic chromatin using DNA hypersensitivity assay[4]. Hsuing and colleagues showed that certain elements like promoters are maintained accessible during mitosis, where other elements such as enhancers are only accessible in interphase but not during mitosis. This is interesting, because this suggests there are chromatin features present in the interphase chromatin that are maintaining these promoter sites accessible during mitosis. The accessibility of promoter during mitosis sites also implies that the motifs, although temporarily less stable bound by factors, maintain their open conformation and are accessible for these factors upon mitotic exit. This was surprising as the 3D organization of the mitotic chromatin does not seem to suggest differences across the mitotic chromatin[6]. Furthermore, again seemingly contradictory with the data on the 3D organization of the mitotic chromatin, this locus specific accessibility can be different between cell types and therefore enable epigenetic memory of chromatin accessibility after mitosis. The mechanism that enables the maintenance of chromatin accessibility in mitosis however still remains to be resolved and might be different for individual loci.

Mitotic Bookmarking and Epigenetic Memory

Mitotic chromosomes have their own, temporary, chromosomal organization with unique characteristics. Interestingly, this organization is highly similar between different cell types, whereas interphase chromosomal organization is on some levels highly cell type specific. Even though most transcription activity ceases and at least a subset of proteins dissociate from the chromosomes during mitosis, cells are capable of rearranging their chromosomes back into cell type specific conformations in early G1[6,14,15]. This suggest that the information to rearrange chromosomes into their interphase chromosomal organization after condensation is contained in the mitotic chromosomes, in the soluble fraction, or both even though it is temporary overruled by machineries in mitosis that enable chromosome condensation and segregation. One way cell type-specific information can be maintained could be through mitotic bookmarking of cell type specific gene regulatory elements by patterns of local histone modifications and other epigenetic marks. These mitotic bookmarks can be histone modifications and variants, DNA methylation, noncoding RNA and less frequently specific transcription factors and histone reader complexes that remain bound during mitosis[87]. There are two distinct functions of mitotic bookmarks. First, there are mitotic bookmarks that exert a function during mitosis, for example histone phosphorylation that is implicated in chromatin condensation[88]. In addition

to that, there are mitotic bookmarks that are important for epigenetic memory, which enables proper open or closed chromatin re-formation in early G1, like the phospho/methyl-switch H3K9me3-S10ph and the histone acetyl transferase Brd4[89,90]. The latter bookmarks may act at single sites, like enhancers and promoters, or at larger domains, for example larger heterochromatic lamina associated domains. Below we describe different forms of mitotic bookmarking, their function and mechanism and give some examples of bookmarks that are passed on through mitosis.

Histone modifications and variants

During S-phase the cell replicates not just the DNA but also the local chromatin state. The appropriate set of histone modifications has to be added onto the new histones assembled on the newly replicated DNA. Although new nucleosomes are incorporated directly after DNA duplication, the histone modifications are added in the time between S-phase up till early G1. A recent study by Alabert et al showed that not all histone modifications are added to the new nucleosomes at the same time[91,92].

Analysis of histone modifications as bookmarks in mitosis has been limited and sometimes contradictory because of the limitations of proper antibodies used to study the marks using immunofluorescence microscopy or ChIP-seq. Although mass spectrometry techniques have been used to study histone modifications in an antibody free and unbiased way, these techniques are not without limits either. Mass spectrometry can measure the mass change caused by the histone modifications. Furthermore, the co-localization of modifications on the same histone tail can be studied using mass spectrometry. However, a downside of mass spectrometry is that it cannot detect the location of the studied histones on the DNA, so the information of where these histone modifications reside on the mitotic chromatin is lost. In addition to that, like all population studies, mass spectrometry relies on the level of homogeneity of the synchronized population that is studied. A recently developed technique by the Bernstein lab combines the power of high-resolution imaging with antibody detection of histone modifications followed by sequencing[93]. Although this technique still relies on antibodies for detection, it can detect combinations of bivalent histone modifications at specific genomic locations.

Even though analysis of histone modifications can be challenging, it will be important to understand the pathways of mitotic bookmarks and the mechanisms by which modifications affect the dynamics and characteristics of mitotic chromosomes, and the transmission of gene regulatory programs that determine cell type identity. Although we will not be able to describe all studies, we will highlight recent and important studies of histone modifications and their modifiers in mitosis.

Histone Phosphorylation

One of the most pronounced histone modifications in mitosis is histone phosphorylation. Many serine and threonine residues on the histone 3 tail become phosphorylated during mitosis, which was already observed in the 1970s[94]. Since many of these histone phosphorylation residues exert their function solely during mitosis and are dephosphorylated upon mitotic exit, this is an example of a mitotic bookmark that has a function during mitosis and may be involved marking gene regulatory elements specifically during mitosis when factors that bind them in interphase are not bound, or in organizing the mitotic chromosome itself. One well-studied histone phosphorylation event is histone 3 serine 10 phosphorylation (H3S10ph). H3S10ph is important for chromosome condensation, most likely through recruitment of regulatory and structural proteins, but the precise mechanisms through which this modification affects chromosome conformation are not fully understood[88,95,96]. The kinase that is the main histone writer of this modification, Aurora B, is shown to be colocalized with other histone kinases like Haspin, which phosphorylates a second histone residues H3T3ph[97,98]. Although Haspin and Aurora B can act on individual chromosome arms, centromeric histone H3 phosphorylation is regulated by a positive feedback loop of these kinases. Recruitment of Aurora B and Haspin and the subsequent hyper phosphorylation of the centromere, enables the recruitment of the chromosomal passenger complex (CPC)[99,100]. The CPC is required for attachment of the mitotic spindle and kinetochores to the mitotic chromatin, which is necessary for proper sister chromatid segregation and completion of cytokinesis[98,101].

Phospho/Methyl Switches

H3S10ph and H3T3ph not only exert their function during mitosis to guide recruitment of the CPC to centromeres, but also have a function in maintaining epigenetic memory along the chromosomal arms during mitosis. Several residues that can be phosphorylated are located next to a lysine, that can be mono, di or tri-methylated[102]. For instance, H3S10 is located immediately adjacent to H3K9. Many of these lysines are modified and bound by regulating factors in interphase. However, when the neighboring serine or threonine becomes phosphorylated in mitosis, these regulating factors can no longer bind the modified lysine residue. The temporary phosphorylation of the neighboring residue switches the lysine off as a regulating histone modification, a so-called phospho/methyl switch[103]. An example of such a phospho/methyl switch is H3K9me3S10pho, where the function of tri-methylated H3 lysine 9 is affected by H3S10 phosphorylation in mitosis (figure 3a)[90,104]. The positioning of H3S10ph next to histone 3 lysine 9 di- and tri-methylation functions like a temporary shield from the chromatin binding factors. During interphase

H3K9me3 acts like a heterochromatic mark, which binds to the heterochromatin protein 1 (HP1). HP1 and H3K9me3 together enable heterochromatic spreading and compaction of the heterochromatin[105]. However, when the neighboring residue H3S10 is phosphorylated during mitosis, HP1 can no longer bind to H3K9me3. This enables the heterochromatin mark to be temporary overruled by the mitotic machinery, inactivating its role in interphase chromosome architecture and allows the chromatin to be condensed in a locus-independent mitosis-specific way. Then, upon mitotic exit, the H3S10ph is dephosphorylated by the PP1gamma complex and heterochromatin protein 1 (HP1) can bind H3K9me2/3 again which mediates proper reestablishment of the heterochromatin, and thus may contribute to the re-establishment of the interphase – specific spatial chromosome conformation[106].

There are also phospho/methyl switches at euchromatic marks that are temporarily affected by a neighboring phosphorylated residue. An example of this is H3T3ph/H3K4me3[107], which is represented in figure 3b. H3 lysine 4 tri-methylation is a mark known to be associated with active promoter sites[108]. H3K4me3 on the interphase chromatin has high affinity for the transcription factor TFIID, which is known to recruit the transcription machinery[109,110]. In mitotic chromatin however, the H3K4me3 neighboring residue H3T3 becomes phosphorylated by Haspin[111]. TFIID binding to H3K4me3 has been shown in a collaborative study by the labs of Jonathan Higgins and Marc Timmers to be severely reduced in mitosis as a result of H3T3 phosphorylation[107]. Upon mitotic exit, H3T3 becomes dephosphorylated by the phosphorylase complex RepoMan PP1 gamma[112,113]. Once the H3T3 residue is no longer phosphorylated, TFIID can bind again to H3K4me3 and RNA polII can bind to the promoter site to start transcription. The importance of maintenance of H3K4me3 as mitotic bookmark is emphasized by the fact that H3K4me3 is copied on to the new nucleosome early after DNA replication. This is in contrast with H3K4me1, a mark for enhancer regions, which is not shielded by H3T3ph and is shown to be added later in the cell cycle[107,114]. This result, combined with observations by Hsiung et al that showed that chromatin in promoter regions, typically marked by H3K4me2/3, remains accessible, whereas many enhancer regions close, or become less accessible, during mitosis[4], suggests that bookmarking sites are correlated with maintenance of accessibility at these sites through mitosis. Phospho/methyl-switches provide an elegant system to temporarily overrule the regulatory and structural effects of cis-elements by dissociating the trans factors that in interphase mediate their activity including long-range interactions with other elements, while maintaining their positional information so that in the next G1 the same set of regulatory elements can be re-activated. Although a lot is still unknown about the exact working mechanisms and whether the kinases and phosphorylases are specifically targeted to these sites,

the phospho/methyl switch is one of the best described and understood mechanisms of mitotic bookmarking.

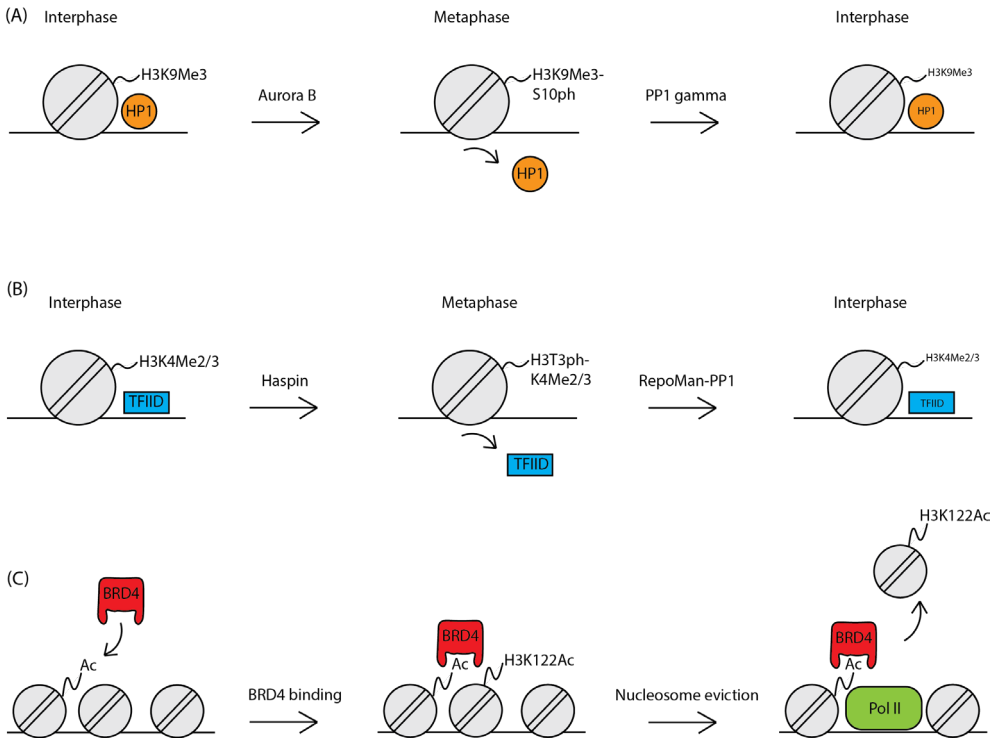


Figure 3. Examples of epigenetic bookmarks in mitosis and interphase. (A) The heterochromatic mark H3K9 trimethylation is shielded by H3S10 phosphorylation caused by the Aurora B kinase during mitosis, which causes heterochromatin protein 1 (HP1) to temporarily dissociated from the mitotic chromatin. When H3S10 is dephosphorylated by the PP1γ complex, HP1 binding is restored. (B) The histone mark H3K4 di and tri methylation becomes shielded during mitotis by the H3T3 phosphorylation mark regulated by the kinase Haspin. This causes the euchromatic regulatory protein TFIID to dissociate from the chromatin. Upon mitotic exit H3T3 is dephosphorylated by the RepoMan-PP1 γ complex which restores TFIID binding to the interphase chromatin. (C) The histone acetyl transferase BRD4 can bind to histone acetylation marks in interphase and mitosis. When bound to the chromatin BRD4 can then acetylate H3K122Ac, which results in nucleosome eviction. Local nucleosome eviction enables reorganization of the nucleosome distribution and binding of big complexes like the transcription machinery.

Histone Acetylation

Most histone acetyltransferases and histone deacetylases migrate off the mitotic chromatin[115]. This implies that many acetylation marks once added after S-phase and G2 are stable throughout mitosis. Although phosphorylation marks are not known to shield acetylation marks like they do for some methylation marks, there are pathways that may help the reestablishment of factors binding to these acetylation marks and

the associated chromatin after mitosis. An example is the chromatin binding protein bromodomain 4 protein (BRD4), that is known to bind either H3K14 or H4K5 or K12 acetylated residues[116,117]. Its working mechanism in interphase is represented in figure 3c. Brd4 binds acetylated histones, but also has histone acetyltransferase activity itself, which preferentially acetylates the H3 lysine 122 residue and several lysine residues on the tail of H3 and H4[89]. H3K122 is located in the core of the nucleosome where the nucleosome is bound to the DNA and acetylation of this residue causes nucleosome clearance of the chromatin[118]. The removal of nucleosomes results in local chromatin decompaction and exposes free DNA to which protein complexes like the transcription machinery can bind[89]. Although this process is only described in interphase and it is unknown if the histone acetyltransferase activity of Brd4 is present in mitosis, it is known that Brd4 is one of the few HATs that remain bound to mitotic chromatin[117,119]. Furthermore, the sites where it remains bound to the mitotic chromatin are associated to post-mitotic transcription and Brd4 overexpression was shown to accelerate the reactivation of transcription upon mitotic exit[117,120]. This suggests that Brd4 and its associated histone acetylation marks function as bookmarks during mitosis and enable reestablishment of post-mitotic transcription in a timely matter. One can imagine that loss of chromatin accessibility of certain regions as shown by Hsiung et al can be reestablished after mitosis by enzymes like Brd4[4].

Another histone acetylation mark associated with reactivation of post-mitotic transcription is H3K27Ac. In a recent paper by Hsiung et al. polII binding upon mitotic exit was measured using ChIP-seq to study prevalence, location and order of spikes in transcription after mitotic exit[5]. An interesting observation was that the histone modification H3K27Ac could best predict which sites were going to be transcriptionally active after mitosis. The regulatory pathways of H3K27Ac deposition and inheritance, and how this modification is involved in a spike in post-mitotic transcription still remain to be resolved. It will be very interesting to better understand the interplay of histone modification and re-initiation of transcription upon mitotic exit.

Histone Variants

Not much is known about the function of histone variants in mitosis and when histone variants are incorporated after DNA replication. During S-phase only the canonical histone variants H3.1 and H3.2 are incorporated in the newly synthesized DNA, which implies that histone variants that are functioning as mitotic bookmarks should be incorporated before the end of mitosis[121]. Which remodelers are regulating this replacement and when in the cell cycle this occurs is not known. The best studied

histone variant with a known function in mitosis is the centromeric variant H3, CENP-A in humans[122,123]. CENP-A is localized at centromeres and has been shown to help the localization of the mitotic kinetochores, which enables proper DNA segregation and cytokinesis. Defects in the CENP-A localization have been shown to affect proper DNA segregation in mitosis[124]. A recent study by Roulland et al showed the importance of the flexible tails of CENP-A in mitosis. When the flexible tail of CENP-A is switched with the more rigid tail of the canonical H3, several kinetochore proteins dislocate and severe mitotic and cytokinetic defects like aneuploidy are observed[125]. Although DNA sequences in centromeres have been shown not to be conserved, the centromeric H3 histone variant shows high levels of similarity across species[126]. This suggests that DNA sequence does not determine sites of CENP-A incorporation, but is likely epigenetically determined, e.g. by the presence of old CENP-A. CENP-A incorporation is uncoupled from DNA-replication and is mediated by a dedicated histone chaperone HJURP[127–130].

Histone variant H3.3 is another histone 3 variant known to be inherited through mitosis. H3.3 can function as a mark for promoters of transcriptionally active genes[131]. H3.3-containing nucleosomes appear less stable than the canonical H3 containing nucleosomes, which enables nucleosome clearance and remodeling, and therefore initiates transcription[132]. Furthermore, H3.3 is associated with many active histone modifications, such as histone acetylation and methylation (e.g. H3K4me3), which attracts histone modifiers that can spread the histone modifications to neighboring histones, both H3.3 and canonical H3 histones. In contrast to other H3 variants, H3.3 can be incorporated to the chromatin independent of DNA replication, as its remodeler is expressed in G1 and G2 as well as S-phase[121,131,133,134].

Linker Histone H1

Hyper phosphorylation of the histone variant H1 is also a hallmark of mitosis[135,136]. Histone H1 accumulates phosphorylation marks during the cell cycle, starting at no or low levels of phosphorylation in G1 to the highest levels of phosphorylation in M-phase[137,138]. Some histone H1 variants have been shown to become more phosphorylated than other variants, however all H1 variants gain phosphorylation marks in mitosis[139,140]. It has been suggested that H1 and its phosphorylated forms are key mediators of chromatin structure and chromosome condensation[141–143]. Maresca et al immuno-depleted H1 in Xenopus laevis egg extracts[144] and found chromatin does not condense properly and chromosomes have elongated arms. Furthermore, the chromatids showed misalignment on the metaphase plate, which lead to defects in sister segregation. Interestingly, positioning of the kinetochore proteins was unaffected in the H1 depleted chromatin. It has been suggested that

the centromeric H3 variant CENP-A might not need histone H1 to interact with the linker DNA, as CENP-A contains a conserved domain that shows highly similarity with motifs present on the H1 tails. The results of the recent study by Roulland et al confirms these suggestions as they show a low binding affinity between CENP-A and H1 as a result of the flexible histone tails of CENP-A[125].

DNA methylation

DNA methylation is a layer of epigenetic regulation that is closest to the genetic information in the DNA and was one of the first epigenetic marks to be discovered[145,146]. Methylated cytosines can function as a chromatin silencing mark and enable imprinting of gene silencing[147]. In contrast to histone modifications and histone variants, DNA methylation of the newly synthesized strand is established immediately after the replication fork has passed, using the old strands as template[148]. This enables correct copying of cytosine methylation and prevents loss during multiple cell divisions. This makes DNA methylation unique among the other mitotic bookmarks, since the copying of the other bookmarks are delayed and spread out over G2 and M-phase. DNA binding of factors can be both positively and negatively affected by DNA methylation[149–151], which can influence the chromatin accessibility state and long-range chromatin interactions. How these phenomena are altered or modulated in mitotic chromosomes to facilitate the folding of the chromosomes as linear loop arrays is not known yet. It will be interesting to study the interplay between DNA methylation and other mitotic bookmarks that mark repressive regions in the chromatin like histone 3 lysine 9 tri-methylation.

Non-coding RNA

During recent years roles for non-coding RNAs (ncRNAs) in epigenetic regulation of the genome have been uncovered. Most of these studies are done in non-synchronously cycling cells and little is known about roles of ncRNA and formation of mitotic chromosomes. It was generally assumed that RNAs do not play a big role in the mitotic chromatin, since the transcription machinery stalls and migrates off the mitotic chromatin[15]. This implies that there is no production of new RNAs during mitosis and RNA molecules that are retained on the mitotic chromatin have to be produced before mitosis initiation. A recent study however showed the presence of large group of ncRNAs during mitosis, which were coined mitotic chromatin associated RNAs, mCARs[152]. These authors used a novel technique called 5'-tag sequencing to detect a large set of new ncRNAs detected that either bind mitotic chromatin (mCARs) or interphase chromatin (iCARs). A large fraction of the detected mCARs consisted of small nucleolar RNA molecules (snoRNAs) and long non-coding RNAs (lncRNAs).

Furthermore, it was shown that the mCARs are highly conserved across species in a wide range of vertebrate species (Meng et al., 2016), which suggests a conserved function of these mCARs in mitosis. In an attempt to study the localization of these mCARs on mitotic chromatin, it was observed that there are two classes of mCARs which located either on the condensed chromosomal exterior or on the interior of the condensed chromosome. However, it remains to be resolved what the role of mCARs is during mitosis, whether their identity and location are cell type-specific, and whether they can function as mitotic bookmarks.

Transcription Factors

It was suggested for many years that most transcription factors and other interphase chromatin binding factors migrate from the chromatin upon mitotic entry and would remain in the cytosol until after DNA decondensation[153]. However, several studies have identified chromatin binding factors that remain on the mitotic chromatin, like the previously described histone acetyl transferase BRD4 and transcription factors like FOXA1, MYC and RUNX2, which have been shown to be important in cell fate determination[89,154–156]. FOXA1 was shown to specifically remain bound to highly transcribed regions in mitosis and promotes initiation of transcription of these regions after mitosis[154]. Although it is not clear why some transcription factors remain on the mitotic chromosome and others migrate away from the chromatin, it is clear that some transcription factors play an important role in mitotic bookmarking and transcriptional reactivation after mitosis[157]. A recent paper by Teves et al. even suggests that many transcription factors remain associated to mitotic chromosomes, although in a more dynamic fashion compared to interphase[158]. Teves et al. show that earlier studies showing transcription factors migrating of mitotic chromatin might be caused by fixation artefacts. Future studies will have to explore which transcription factors indeed remain bound to the mitotic chromosome and which function these factors exert during mitosis.

Lamina interactions

One of the major morphological changes of the vertebrate cell that occurs during mitosis is the breakdown of the nuclear envelope. During interphase the proteins that coat the inside of the nuclear membrane, the nuclear lamina, play roles in providing support and directing higher order chromosome folding and nuclear organization[159,160]. Using an elegant system to track individual lamina interactions in single cells over time, it was shown that lamina interacting regions are not necessarily the same after mitotic exit and that the lamina interacting regions reassemble over the nuclear envelope. However, the regions that interact with the nuclear lamina

after mitosis are always associated with other regions that are associated with transcription repression and regions assigned as LADs (lamina interacting domains) in population studies[161]. This implies that information regarding lamina interactions is not maintained by a mitotic bookmark throughout mitosis, however the epigenetic marks that label these regions as repressed and candidates for lamina interactions are maintained throughout mitosis. One of the histone modifications that could act as such a mark is H3K9me2, as sites of hyper methylated nucleosomes are typically found in regions containing LADs[161].

Meiotic bookmarking

Although not elaborately described in this review, it is known that epigenetic marks are not only inherited through mitotic cell divisions but can also be transmitted through meiosis and passed on to next generations. Recent studies in Drosophila showed how epigenetic marks in the parental epigenome are inherited into the following generation[162]. This was surprising considering that spermatocytes do not have nucleosomes (or only a small number at specific sites), but use other DNA binding proteins to enable compact folding and protection against DNA damage causing factors to which the spermatozoa are exposed[163,164]. Even more so, it has been shown that upon fertilization, the gametes undergo epigenome reprogramming[165]. However, studies like those performed by Ost et al showed that specific marks are able to escape this epigenetic reprogramming and can be passed on to epigenome of offspring. Further research is needed to explore the mechanisms by which meiotic epigenomic bookmarking is established and maintained.

Mitotic bookmarking in differentiation

As described in the paragraphs above, there are many different mechanisms that can establish mitotic bookmarking and epigenetic inheritance. However, one can imagine that the epigenetic marks need to change when cells undergo differentiation. During differentiation the chromatin characteristics of cells changes in many ways. Chromatin organization is changed at the scale of positioning and modification of individual nucleosomes at cell type specific gene regulatory elements[12,166], at the level of long-range interactions between such elements and the formation chromatin domains such as A- and B- compartments[47,167,168]. Correspondingly, gene expression programs are affected[169]. It is unlikely that large-scale changes in chromosome conformation are made during G1 as chromosomes are too big to move around in the interphase nucleus, which would be necessary to change the structures on the level of compartments[28,170]. During mitotic exit and early G1 chromosomes decondense and the nucleus reforms, providing the cell a window of opportunity to

spatially rearrange its genome in accordance with changes in cell type. This concept is interesting in the context of cell type differentiation. During differentiation there are two scenarios, one where the cell divides into two equal daughter cells with a distinct cell type from the original cell. The second scenario is one where the two daughter cells have different states, e.g. one daughter cell is initiated to become a different cell type and the other daughter cell will remain the same type as the mother cell, a stem cell. In both scenarios epigenetic features change before, during or after mitosis, suggesting roles for bookmarking processes. In addition, in the second scenario, where the daughter cells acquire each a different state, epigenetic changes in bookmarks are specific to each of the sister chromatids. This adds another layer of complexity to the mechanisms of epigenetic inheritance and modification: how do cells change their mitotic bookmarks in order to initiate correct differentiation and how do cells initiate and control differences in epigenetic states of otherwise identical sister chromatids?

Several histone modifications have been described to change during cell type differentiation and repression of enzymes that deposit these modifications represses the differentiation process and the activation of differentiation-dependent genes. Examples of these are H3K4me3, which has been shown to change during differentiation of human embryonic stem cells, and H3K36me3, which was shown to be required for differentiation of mouse embryonic stem cells to endoderm[172,173].

Several techniques have been developed to study how sister chromatids separate from each other. One of these techniques is Strand-seq[174,175]. Using BrdU labelling the sister chromatids can be separately observed and followed over stem cell differentiation. A hypothesis called the Immortal Strand Hypothesis has been introduced many years ago that proposes that stem cells will try to maintain the original copy of the chromosome to prevent accumulation of DNA replication mistakes[176]. Although some studies have contradicted this hypothesis, it is clear that at least some cells are retaining certain chromosomes in the daughter stem cells specifically[177–181]. Using Strand-seq, the distribution of strands was followed over multiple cell cycles after labelling with BrdU for 1 cycle. The immortal strand hypothesis suggests that for some chromosomes the labeled copy of the chromosome will be unequally distributed into on daughter cell, the stem cell (represented as a flowchart in figure 4a). In order to understand how histones and their modifications can segregate to the differentiating daughter cell, it is important to understand that the H3/H4 tetramer of nucleosomes are always incorporated together in the nucleosomes as the parental H3/H4 tetramer stays together during DNA replication. This implies that nucleosomes can be assigned as new and old nucleosomes after DNA replication[182]. A paper by Tran et al in 2013 already showed that new

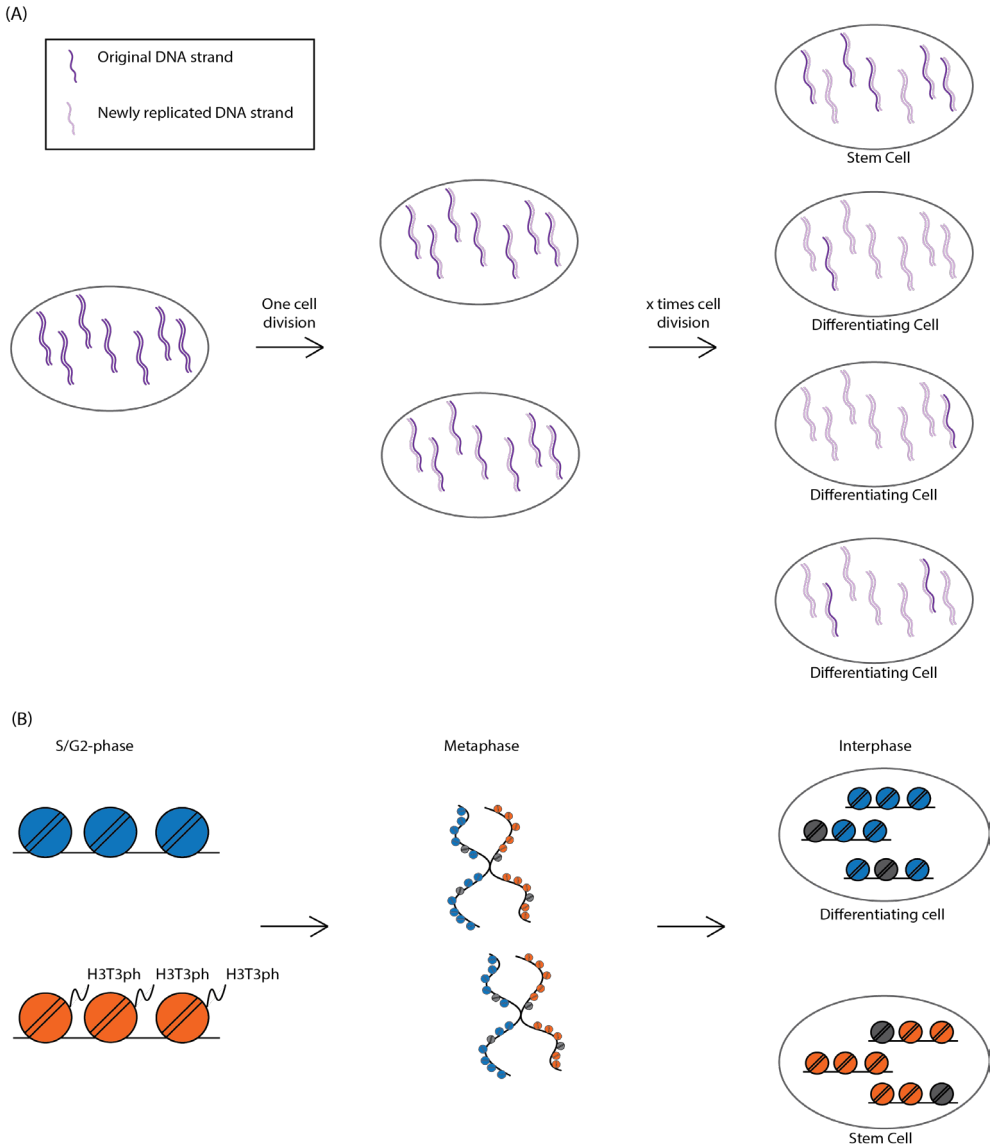


Figure 4. Mitotic bookmarks in cell differentiation. (A) Representation of the silent sister hypothesis as suggested by Falconer et al[171]. During stem cell differentiation it is believed that the daughter cells destined to be the stem cell will retain the original DNA strands of some chromosomes, where the differentiation daughter cells will mainly contain the newly replicated DNA strands. (B) Nucleosomes on one sister chromatid are specifically labelled with the H3T3 phosphorylation mark. This enables the cell to retain the sister chromatid containing to the stem cell and the other sister chromatid will be passed on to the differentiation daughter cell.

and old nucleosomes are not always equally represented on each sister chromatid, as they show that in stem cell differentiation of the male germline of Drosophila melanogaster preexisting nucleosomes preferentially segregate to the germline stem cell, whereas the new nucleosomes segregate to the differentiating daughter cell[183]. Furthermore, recent studies in Drosophila and mammalian cell lines observed differences in histone modifications between these old and new nucleosomes[91,92,184]. It was shown that H3T3 phosphorylation is preferentially more represented on newly synthesized nucleosomes in the male Drosophila germline, represented in figure 4b, and when H3T3ph is impaired, proper differentiation of the germline was affected[185]. It will be very interesting to explore the pathways and mechanisms that are used to initiate epigenetic changes in cellular phenotype, how differences between sister chromatids are established and proper sister segregation is controlled.

Concluding remarks

We are only starting to understand the mechanisms by which epigenetic information contained within the vertebrate chromatin is transmitted through mitosis and how this occurs in the context of a mitotic chromosome conformation that is dramatically different from interphase. One important question that remains unanswered is how molecular details of epigenetic bookmarks are read in early G1 and enable re-establishment of cell type specific chromatin organization. Insights into these processes promise not only to lead to mechanistic understanding of mitotic inheritance of cell type specific chromatin state, but they will also reveal how the spatial organization of interphase chromosomes is determined in general by the action of cis-acting elements along the chromatin fiber. This will also lead to a better understanding of what epigenetic mechanisms underlie processes in which cell type identity is changed, for example in stem cell differentiation or in diseases that result in cancer development and aging.

Declaration of Interest

Work in our laboratory is supported by the National Human Genome Research Institute (R01 HG003143, U54 HG007010, U01 HG007910), the National Cancer Institute (U54 CA193419), the NIH Common Fund (U54 DK107980, U01 DA 040588), the National Institute of General Medical Sciences (R01 GM 112720), and the National Institute of Allergy and Infectious Diseases (U01 R01 AI 117839). The authors declare that they have no competing interests. J.D. is an investigator of the Howard Hughes Medical Institute.

References

1. Earnshaw, W. C. & Laemmli, U. K. Architecture of metaphase chromosomes and chromosome scaffolds. *J. Cell Biol.* **96**, 84–93 (1983).
2. Maeshima, K. & Laemmli, U. K. A Two-step scaffolding model for mitotic chromosome assembly. *Dev. Cell* **4**, 467–480 (2003).
3. Marsden, M. P. F. & Laemmli, U. K. Metaphase chromosome structure: Evidence for a radial loop model. *Cell* **17**, 849–858 (1979).
4. Hsiung, C. C. *et al.* Genome accessibility is widely preserved and locally modulated during mitosis. 1–29 (2015). doi:10.1101/gr.180646.114
5. Hsiung, C. C.-S. & Blobel, G. A. A new bookmark of the mitotic genome in embryonic stem cells. *Nat. Cell Biol.* **18**, 1124–1125 (2016).
6. Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342**, 948–53 (2013).
7. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–8 (2013).
8. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–11 (2002).
9. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–93 (2009).
10. Orlando, V. Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem. Sci.* **25**, 99–104 (2000).
11. van Steensel, B., Delrow, J. & Henikoff, S. Chromatin profiling using targeted DNA adenine methyltransferase. *Nat. Genet.* **27**, 304–308 (2001).
12. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
13. Consortium Roadmap Epigenomics *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
14. Martinez-Balbas, M. A., Dey, A., Rabindran, S. K., Ozato, K. & Wu, C. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* **83**, 29–38 (1995).
15. Prescott, D. M. & Bender, M. DNA synthesis and mitosis in cultures of human peripheral leukocytes. *Exp. Cell Res.* **27**, 221–9 (1962).
16. Levsky, J. M. & Singer, R. H. Fluorescence in situ hybridization: past, present and future. *J. Cell Sci.* **116**, 2833–8 (2003).
17. Banfalvi, G. Overview of Cell Synchronization. in 1–23 (2011). doi:10.1007/978-1-61779-182-6_1
18. Vassilev, L. T. Cell cycle synchronization at the G2/M phase border by reversible inhibition of CDK1. *Cell Cycle* **5**, 2555–2556 (2006).
19. Vassilev, L. T. *et al.* Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10660–10665 (2006).
20. Xeros, N. Deoxyriboside control and synchronization of mitosis. *Nature* **194**, 682–683 (1962).
21. Eichten, S. R., Schmitz, R. J. & Springer, N. M. Epigenetics: Beyond Chromatin Modifications and Complex Genetic Regulation. *Plant Physiol.* **165**, 933–947 (2014).
22. Saze, H. Epigenetic memory transmission through mitosis and meiosis in plants. *Semin. Cell Dev. Biol.* **19**, 527–536 (2008).
23. Weimer, A. K., Demidov, D., Lermontova, I., Beeckman, T. & Van Damme, D. Aurora Kinases Throughout Plant Development. *Trends Plant Sci.* **21**, 69–79 (2016).
24. Flemming, W. Zur Kenntnis der Zelle und ihrer Teilung-Erscheinungen. *Schr. Nat. Wiss. Ver. Schlesw.-Holst.* **3**, 23–27 (1878).

25. Paweletz, N. Walther Flemming: pioneer of mitosis research. *Nat. Rev. Mol. Cell Biol.* **2**, 72–75 (2001).
26. Bickmore, W. A. & Van Steensel, B. Genome architecture: Domain organization of interphase chromosomes. *Cell* **152**, 1270–1284 (2013).
27. Dekker, J., Marti-Renom, M. a & Mirny, L. a. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat. Rev. Genet.* **14**, 390–403 (2013).
28. Dekker, J. & Mirny, L. The 3D Genome as Moderator of Chromosomal Communication. *Cell* **164**, 1110–1121 (2016).
29. Cremer, T. *et al.* Role of chromosome territories in the functional compartmentalization of the cell nucleus. *Cold Spring Harb. Symp. Quant. Biol.* **58**, 777–792 (1993).
30. Branco, M. R. & Pombo, A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.* **4**, 780–788 (2006).
31. Goloborodko, A., Imakaev, M. V., Marko, J. F. & Mirny, L. Compaction and segregation of sister chromatids via active loop extrusion. *Elife* **5**, 1–20 (2016).
32. Tanabe, H. *et al.* Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 4424–9 (2002).
33. Wang, S. *et al.* Spatial organization of chromatin domains and compartments in single chromosomes. **8084**, (2016).
34. Rao, S. S. P. *et al.* A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* **159**, 1665–1680 (2014).
35. Iborra, F. J., Pombo, a, Jackson, D. a & Cook, P. R. Active RNA polymerases are localized within discrete transcription 'factories'' in human nuclei.' *J. Cell Sci.* **109 (Pt 6**, 1427–1436 (1996).
36. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–80 (2012).
37. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–5 (2012).
38. Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 1–12 (2016). doi:10.1101/024620
39. Giorgetti, L. *et al.* Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* **157**, 950–63 (2014).
40. Sanborn, A. L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci.* (2015). doi:10.1073/pnas.1518552112
41. Phillips-Cremins, J. E. & Corces, V. G. Chromatin insulators: linking genome organization to cellular function. *Mol. Cell* **50**, 461–74 (2013).
42. Tang, Z. *et al.* CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell* **163**, 1–17 (2015).
43. Wit, E. De *et al.* CTCF Binding Polarity Determines Chromatin Looping. *Mol. Cell* **60**, 1–9 (2015).
44. Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 996–1001 (2014).
45. Gibcus, J. H. & Dekker, J. The hierarchy of the 3D genome. *Mol. Cell* **49**, 773–82 (2013).
46. Doyle, B., Fudenberg, G., Imakaev, M. & Mirny, L. a. Chromatin Loops as Allosteric Modulators of Enhancer-Promoter Interactions. *PLoS Comput. Biol.* **10**, e1003867 (2014).
47. Sanyal, A., Lajoie, B. R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature* **489**, 109–13 (2012).
48. Hnisz, D. *et al.* Activation of proto-oncogenes by disruption of chromosome

neighborhoods. *Science* **351**, 1454–1458 (2016).
49. Jin, W. *et al.* Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. *Nature* **528**, 142–146 (2015).
50. Smith, E. M., Lajoie, B. R., Jain, G. & Dekker, J. Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the CFTR Locus. *Am. J. Hum. Genet.* **98**, 185–201 (2016).
51. Deng, W. *et al.* Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **158**, 849–860 (2014).
52. Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. & De Laat, W. Looping and interaction between hypersensitive sites in the active ??-globin locus. *Mol. Cell* **10**, 1453–1465 (2002).
53. Vernimmen, D., De Gobbi, M., Sloane-Stanley, J. a, Wood, W. G. & Higgs, D. R. Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* **26**, 2041–2051 (2007).
54. Lupianez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
55. Zhao, J. *et al.* HnRNP U mediates the long-range regulation of Shh expression during limb development. *Hum. Mol. Genet.* **18**, 3090–3097 (2009).
56. Speicher, M. R. & Carter, N. P. The new cytogenetics: blurring the boundaries with molecular biology. *Nat. Rev. Genet.* **6**, 782–92 (2005).
57. Uchida, I. A. & Lin, C.-C. Fluorescent staining ofhuman chromosomes: Identification ofsome common aberrations. *C.M.A. J.* **105**, 479–482 (1971).
58. Craig, J. M. & Bickmore, W. A. Chromosome bands--flavours to savour. *Bioessays* **15**, 349–54 (1993).
59. Hughes, J. R. *et al.* Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* **46**, 205–12 (2014).
60. Vagnarelli, P. Mitotic chromosome condensation in vertebrates. *Exp. Cell Res.* **318**, 1435–41 (2012).
61. Liang, K. *et al.* Mitotic Transcriptional Activation : Clearance of Actively Engaged Pol II via Transcriptional Elongation Control in Mitosis Article Mitotic Transcriptional Activation : Clearance of Actively Engaged Pol II via Transcriptional Elongation Control in Mitosis. *Mol. Cell* **60**, 1–11 (2015).
62. Llères, D., James, J., Swift, S., Norman, D. G. & Lamond, A. I. Quantitative analysis of chromatin compaction in living cells using FLIM-FRET. *J. Cell Biol.* **187**, 481–496 (2009).
63. Nagasaka, K., Hossain, M. J., Roberti, M. J., Ellenberg, J. & Hirota, T. Sister chromatid resolution is an intrinsic part of chromosome organization in prophase. *Nat. Cell Biol.* (2016). doi:10.1038/ncb3353
64. Hirano, T. & Mitchison, T. J. A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. *Cell* **79**, 449–458 (1994).
65. Kimura, K. & Hirano, T. ATP-dependent positive supercoiling of DNA by 13S condensin: A biochemical implication for chromosome condensation. *Cell* **90**, 625–634 (1997).
66. Ono, T. *et al.* Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. *Cell* **115**, 109–121 (2003).
67. Hirota, T., Gerlich, D., Koch, B., Ellenberg, J. & Peters, J. M. Distinct functions of condensin I and II in mitotic chromosome assembly. *J Cell Sci* **117**, 6435–6445 (2004).
68. Cuylen, S., Metz, J., Hruby, A. & Haering, C. H. Entrapment of Chromosomes by Condensin Rings Prevents Their Breakage during Cytokinesis. *Dev. Cell* **27**, 469–478 (2013).
69. Nasmyth, K. Disseminating the Genome: Joining, Resolving, and Separating

Sister Chromatids During Mitosis and Meiosis. *Http://Dx.Doi.Org/10.1146/Annurev.Genet.35.102401.091334* 673–745 (2003).
70. Riggs, A. D. DNA Methylation and Late Replication Probably Aid Cell Memory, and Type 1 DNA Reeling Could Aid Chromosome Folding and Enhancer Function. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **326**, 285–297 (1990).
71. Dekker, J. Two ways to fold the genome during the cell cycle : insights obtained with chromosome conformation capture. *Epigenetics Chromatin* **7**, 1–12 (2014).
72. Zhang, Y., Isbaner, S. & Heermann, D. Mechanics of Sister Chromatids studied with a Polymer Model. *Front. Phys.* **1**, 1–11 (2013).
73. Alipour, E. & Marko, J. F. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res.* **40**, 11202–11212 (2012).
74. Rattner, J. B. & Lin, C. C. Radial loops and helical coils coexist in metaphase chromosomes. *Cell* **42**, 291–296 (1985).
75. Ohta, S., Wood, L., Bukowski-Wills, J. C., Rappsilber, J. & Earnshaw, W. C. Building mitotic chromosomes. *Curr. Opin. Cell Biol.* **23**, 114–121 (2011).
76. Cuylen, S. *et al.* Ki-67 acts as a biological surfactant to disperse mitotic chromosomes. *Nature* **44**, 1–19 (2016).
77. Scholzen, T. *et al.* The Ki-67 protein interacts with members of the heterochromatin protein 1 (HP1) family: a potential role in the regulation of higher-order chromatin structure. *J. Pathol.* **196**, 135–144 (2002).
78. Foti, R. *et al.* Nuclear architecture organized by Rif1 underpins the replication-timing program. *Mol. Cell* 1–14 (2015). doi:10.1016/j.molcel.2015.12.001
79. Haering, C. H., Löwe, J., Hochwagen, A. & Nasmyth, K. Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol. Cell* **9**, 773–788 (2002).
80. Ivanov, D. & Nasmyth, K. A topological interaction between cohesin rings and a circular minichromosome. *Cell* **122**, 849–860 (2005).
81. Tanaka, T., Cosma, M. P., Wirth, K. & Nasmyth, K. Identification of Cohesin Association Sites at Centromeres and along Chromosome Arms. *Cell* **98**, 847–858 (1999).
82. Waizenegger, I. C., Hauf, S., Meinke, a & Peters, J. M. Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase. *Cell* **103**, 399–410 (2000).
83. Liang, Z. *et al.* Chromosomes Progress to Metaphase in Multiple Discrete Steps via Global Compaction/Expansion Cycles. *Cell* **161**, 1124–1137 (2015).
84. Toselli-Mollereau, E. *et al.* Nucleosome eviction in mitosis assists condensin loading and chromosome condensation. *EMBO J.* **35**, 1–17 (2016).
85. Grigoryev, S. A., Bascom, G., Buckwalter, J. M., Schubert, M. B. & Woodcock, C. L. Hierarchical looping of zigzag nucleosome chains in metaphase chromosomes. *PNAS* (2015). doi:10.1073/pnas.1518280113
86. Grigoryev, S. A. Nucleosome spacing and chromatin higher-order folding. *Nucleus* **3**, 493–9 (2012).
87. Wang, F. & Higgins, J. M. G. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol.* **23**, 175–84 (2013).
88. Wilkins, B. J. *et al.* A cascade of histone modifications induces chromatin condensation in mitosis. *Science* **343**, 77–80 (2014).
89. Devaiah, B. N. *et al.* BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin. *Nat. Struct. Mol. Biol.* **23**, 1–12 (2016).
90. Jeong, Y. S., Cho, S., Park, J. S., Ko, Y. & Kang, Y.-K. Phosphorylation of serine-10 of histone H3 shields modified lysine-9 selectively during mitosis. *Genes Cells* **15**, (2010).
91. Alabert, C. *et al.* Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes Dev.* **29**, 585–590 (2015).
92. Lin, S., Yuan, Z.-F., Han, Y., Marchione, D. M. & Garcia, B. A. Preferential phosphorylation on old histones during early mitosis in human cells. *J. Biol. Chem.* **291**, jbc.M116.726067 (2016).

93. Shema, E. *et al.* Single-molecule decoding of combinatorially modified nucleosomes. *Science (80-.).* **352**, 717–721 (2016).
94. Bradbury, E. M., Inglis, R. J., Matthews, H. R. & Sarner, N. Phosphorylation of very-lysine-rich histone in Physarum polycephalum. Correlation with chromosome condensation. *Eur. J. Biochem.* **33**, 131–9 (1973).
95. Antonin, W. & Neumann, H. Chromosome condensation and decondensation during mitosis. *Curr. Opin. Cell Biol.* **40**, 15–22 (2016).
96. Johansen, K. M. & Johansen, J. Regulation of chromatin structure by histone H3S10 phosphorylation. *Chromosome Res.* **14**, 393–404 (2006).
97. Wang, F. *et al.* Histone H3 Thr-3 Phosphorylation at Centromeres in Mitosis. *Science* **330**, 231–236 (2010).
98. Wang, F. *et al.* Haspin inhibitors reveal centromeric functions of Aurora B in chromosome segregation. *J. Cell Biol.* **199**, 251–68 (2012).
99. Wang, F. *et al.* A positive feedback loop involving Haspin and Aurora B promotes CPC accumulation at centromeres in mitosis. *Curr. Biol.* **21**, 1061–9 (2011).
100. Yamagishi, Y., Honda, T., Tanno, Y. & Watanabe, Y. Two histone marks establish the inner centromere and chromosome bi-orientation. *Science* **330**, 239–43 (2010).
101. Gurden, M. D., Anderhub, S. J., Faisal, A. & Linardopoulos, S. Aurora B prevents premature removal of spindle assembly checkpoint proteins from the kinetochore: A key role for Aurora B in mitosis. *Oncotarget* (2016). doi:10.18632/oncotarget.10657
102. Martin, C. & Zhang, Y. The diverse functions of histone lysine methylation. *Nat. Rev.* **6**, 838–49 (2005).
103. Fischle, W., Wang, Y. & Allis, C. D. Binary switches and modification cassettes in histone biology and beyond. *Nature* **425**, 475–9 (2003).
104. Hirota, T., Lipp, J. J., Toh, B.-H. & Peters, J.-M. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438**, 1176–80 (2005).
105. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–20 (2001).
106. Qian, J., Beullens, M., Lesage, B. & Bollen, M. Aurora B defines its own chromosomal targeting by opposing the recruitment of the phosphatase scaffold Repo-Man. *Curr. Biol.* **23**, 1136–43 (2013).
107. Varier, R. a *et al.* A phospho/methyl switch at histone H3 regulates TFIID association with mitotic chromosomes. *EMBO J.* **29**, 3967–78 (2010).
108. Hon, G. C., Hawkins, R. D. & Ren, B. Predictive chromatin signatures in the mammalian genome. *Hum. Mol. Genet.* **18**, 12–14 (2009).
109. Orphanides, G., Lagrange, T. & Reinberg, D. The general transcription factors of RNA polymerase II. *Genes Dev.* **10**, 2657–83 (1996).
110. Vermeulen, M. *et al.* Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4. *Cell* **131**, 58–69 (2007).
111. Eswaran, J. *et al.* Structure and functional characterization of the atypical human kinase haspin 2VUW 3IQ7 3DLZ inhibitor iodotubercidin & AMP. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 20198–20203 (2009).
112. Prévost, M. *et al.* Quantitative fragmentome mapping reveals novel, domain-specific partners for the modular protein RepoMan (recruits PP1 onto mitotic chromatin at anaphase). *Mol. Cell. Proteomics* **12**, 1468–86 (2013).
113. Qian, J., Lesage, B., Beullens, M., Van Eynde, A. & Bollen, M. PP1/Repo-man dephosphorylates mitotic histone H3 at T3 and regulates chromosomal aurora B targeting. *Curr. Biol.* **21**, 766–73 (2011).
114. Lin, Z., Luo, X. & Yu, H. Structural basis of cohesin cleavage by separase. *Nature* **532**, 131–134 (2016).
115. Kruhlak, M. J. *et al.* Regulation of Global Acetylation in Mitosis through Loss

of Histone Acetyltransferases and Deacetylases from Chromatin. *J. Biol. Chem.* **276**, 38307–38319 (2001).
116. Chiang, C.-M. Brd4 engagement from chromatin targeting to transcriptional regulation: selective contact with acetylated histone H3 and H4. *F1000 Biol. Rep.* **1**, 98 (2009).
117. Dey, A., Nishiyama, A., Karpova, T., McNally, J. & Ozato, K. Brd4 Marks Select Genes on Mitotic Chromatin and Directs Postmitotic Transcription. *Mol. Biol. Cell* **21**, 4042–4056 (2009).
118. Tropberger, P. *et al.* Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell* **152**, 859–872 (2013).
119. Dey, A., Chitsaz, F., Abbasi, A., Misteli, T. & Ozato, K. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 8758–8763 (2003).
120. Zhao, R., Nakamura, T., Fu, Y., Lazar, Z. & Spector, D. L. Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation. *Nat. Cell Biol.* **13**, 1295–304 (2011).
121. Campos, E. I. *et al.* The program for processing newly synthesized histones H3.1 and H4. *Nat. Struct. Mol. Biol.* **17**, 1343–1351 (2010).
122. Ahmad, K. & Henikoff, S. Centromeres are specialized replication domains in heterochromatin. *J. Cell Biol.* **153**, 101–109 (2001).
123. Sánchez, P. & Losada, A. New clues to understand how CENP-A maintains centromere identity. *Cell Div.* **6**, 11 (2011).
124. Maheshwari, S. *et al.* Naturally Occurring Differences in CENH3 Affect Chromosome Segregation in Zygotic Mitosis of Hybrids. *PLoS Genet.* **11**, 1–20 (2015).
125. Roulland, Y. *et al.* The Flexible Ends of CENP-A Nucleosome Are Required for Mitotic Fidelity. *Mol. Cell* **63**, 674–685 (2016).
126. Cheeseman, I. M. *et al.* A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes Dev.* **18**, 2255–2268 (2004).
127. Bassett, E. A. *et al.* HJURP Uses Distinct CENP-A Surfaces to Recognize and to Stabilize CENP-A/Histone H4 for Centromere Assembly. *Dev. Cell* **22**, 749–762 (2012).
128. Dunleavy, E. M. *et al.* HJURP Is a Cell-Cycle-Dependent Maintenance and Deposition Factor of CENP-A at Centromeres. *Cell* **137**, 485–497 (2009).
129. Foltz, D. R. *et al.* Centromere-Specific Assembly of CENP-A Nucleosomes Is Mediated by HJURP. *Cell* **137**, 472–484 (2009).
130. McKinley, K. L. & Cheeseman, I. M. The molecular basis for centromere identity and function. *Nat. Rev. Mol. Cell Biol.* (2015). doi:10.1038/nrm.2015.5
131. Chow, C.-M. *et al.* Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. *EMBO Rep.* **6**, 354–60 (2005).
132. Ahmad, K. & Henikoff, S. The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly ment are not clear. A study in Tetrahymena concluded that no protein difference between histone H3 variants was required for replacement histone deposition and. *Mol. Cell* **9**, 1191–1200 (2002).
133. Ray-Gallet, D. *et al.* HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol. Cell* **9**, 1091–1100 (2002).
134. Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis. *Cell* **116**, 51–61 (2004).
135. Boggs, B. A., Allis, C. D. & Chinault, A. C. Immunofluorescent studies of human chromosomes with antibodies against phosphorylated H1 histone. *Chromosoma* **108**, 485–90 (2000).
136. Bradbury, E. M., Inglis, R. J. & Matthews, H. R. Control of cell division by very lysine rich histone (F1) phosphorylation. *Nature* **247**, 257–261 (1974).

137. Bleher, R. & Martin, R. Nucleo-cytoplasmic translocation of histone H1 during the HeLa cell cycle. *Chromosoma* **108**, 308–316 (1999).
138. Halmer, L. & Gruss, C. Effects of cell cycle dependent histone H1 phosphorylation on chromatin structure and chromatin replication. *Nucleic Acids Res.* **24**, 1420–1427 (1996).
139. Clausell, J., Happel, N., Hale, T. K., Doenecke, D. & Beato, M. Histone H1 subtypes differentially modulate chromatin condensation without preventing ATP-dependent remodeling by SWI/SNF or NURF. *PLoS One* **4**, e0007243 (2009).
140. Gréen, A., Lönn, A., Peterson, K. H., Ollinger, K. & Rundquist, I. Translocation of histone H1 subtypes between chromatin and cytoplasm during mitosis in normal human fibroblasts. *Cytometry. A* **77**, 478–84 (2010).
141. Ohsumi, K., Katagiri, C. & Kishimoto, T. Chromosome condensation in Xenopus mitotic extracts without histone H1. *Science (80-.).* **262**, 2033–2035 (1993).
142. Thoma, F., Koller, T. H. & Klug, A. Involvement of Histone H1 in the Organization of the Nucleosome and of the Salt-Dependent Superstructures of Chromatin. **83**, (1979).
143. Thoma, F. & Koller, T. H. Influence of histone H1 on chromatin structure. *Cell* **12**, 101–107 (1977).
144. Maresca, T. J., Freedman, B. S. & Heald, R. Histone H1 is essential for mitotic chromosome architecture and segregation in Xenopus laevis egg extracts. *J. Cell Biol.* **169**, 859–69 (2005).
145. Holliday, R. & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226–232 (1975).
146. Holliday, R. Epigenetic Defects. *Science (80-.).* **238**, 163–170 (1987).
147. Curradi, M., Izzo, A., Badaracco, G. & Landsberger, N. Molecular mechanisms of gene silencing mediated by DNA methylation. *Mol. Cell. Biol.* **22**, 3157–73 (2002).
148. Wigler, M., Levy, D. & Perucho, M. The somatic replication of DNA methylation. *Cell* **24**, 33–40 (1981).
149. Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–92 (2012).
150. Maurano, M. T. *et al.* Role of DNA Methylation in Modulating Transcription Factor Occupancy. *Cell Rep.* **12**, 1184–1195 (2015).
151. Schübeler, D. Function and information content of DNA methylation. *Nature* **517**, 321–326 (2015).
152. Meng, Y. *et al.* The non-coding RNA composition of the mitotic chromosome by 5′-tag sequencing. *Nucleic Acids Res.* gkw195 (2016). doi:10.1093/nar/gkw195
153. Chen, D. *et al.* Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins. *J. Cell Biol.* **168**, 41–54 (2005).
154. Caravaca, J. M. *et al.* Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes Dev.* **27**, 251–60 (2013).
155. Topham, C. *et al.* MYC Is a Major Determinant of Mitotic Cell Fate. *Cancer Cell* **28**, 129–140 (2015).
156. Young, D. W. *et al.* Mitotic retention of gene expression patterns by the cell fate-determining transcription factor Runx2. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3189–3194 (2007).
157. Zaret, K. S. Genome reactivation after the silence in mitosis: Recapitulating mechanisms of development? *Dev. Cell* **29**, 132–134 (2014).
158. Teves, S. S. *et al.* A Dynamic Mode of Mitotic Bookmarking by Transcription Factors. *Elife* 066464 (2016). doi:10.1101/066464
159. Chubb, J. R., Boyle, S., Perry, P. & Bickmore, W. A. Chromatin motion is constrained by association with nuclear compartments in human cells. *Curr.*

Biol. **12**, 439–445 (2002).
160. Steensel, B. van & Dekker, J. Genomics tools for unraveling chromosome architecture. *Nat. Biotechnol.* **28**, 1089–1095 (2010).
161. Kind, J. *et al.* Single-cell dynamics of genome-nuclear lamina interactions. *Cell* **153**, 178–192 (2013).
162. Öst, A. *et al.* Paternal diet defines offspring chromatin state and intergenerational obesity. *Cell* **159**, 1352–1364 (2014).
163. Banerjee, S., Smallwood, a & Hultén, M. ATP-dependent reorganization of human sperm nuclear chromatin. *J. Cell Sci.* **108 (Pt 2**, 755–765 (1995).
164. Brunner, A. M., Nanni, P. & Mansuy, I. M. Epigenetic marking of sperm by post-translational modification of histones and protamines. *Epigenetics Chromatin* **7**, 2 (2014).
165. Gill, M. E., Erkek, S. & Peters, A. H. F. M. Parental epigenetic control of embryogenesis: A balance between inheritance and reprogramming? *Curr. Opin. Cell Biol.* **24**, 387–396 (2012).
166. Thurman, R. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75–82 (2012).
167. Beagan, J. A. *et al.* Local Genome Topology Can Exhibit an Incompletely Rewired 3D-Folding State during Somatic Cell Reprogramming. *Cell Stem Cell* **18**, 611–24 (2016).
168. Wijchers, P. J. *et al.* Cause and Consequence of Tethering a SubTAD to Different Nuclear Compartments. *Mol. Cell* 1–13 (2016). doi:10.1016/j.molcel.2016.01.001
169. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–8 (2012).
170. Rosa, A. & Everaers, R. Structure and dynamics of interphase chromosomes. *PLoS Comput. Biol.* **4**, (2008).
171. Falconer, E. & Lansdorp, P. M. Strand-seq: A unifying tool for studies of chromosome segregation. *Semin. Cell Dev. Biol.* **24**, 643–652 (2013).
172. Grandy, R. A. *et al.* Genome-wide studies reveal that H3K4me3 modification in bivalent genes is dynamically regulated during the pluripotent cell cycle and stabilized upon differentiation. *Mol. Cell. Biol.* **36**, 615–627 (2016).
173. Zhang, Y. *et al.* H3K36 histone methyltransferase Setd2 is required for murine embryonic stem cell differentiation toward endoderm. *Cell Rep.* **8**, 1989–2002 (2014).
174. Falconer, E. *et al.* DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nat. Methods* **9**, (2012).
175. Falconer, E. *et al.* Identification of sister chromatids by DNA template strand sequences. *Nature* **463**, 93–97 (2010).
176. Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255**, cp1- (1975).
177. Conboy, M. J., Karasov, A. O. & Rando, T. A. High incidence of non-random template strand segregation and asymmetric fate determination in dividing stem cells and their progeny. *PLoS Biol.* **5**, 1120–1126 (2007).
178. Karpowicz, P. *et al.* Support for the immortal strand hypothesis: Neural stem cells partition DNA asymmetrically in vitro. *J. Cell Biol.* **170**, 721–732 (2005).
179. Potten, C. S., Owen, G. & Booth, D. Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J. Cell Sci.* **115**, 2381–2388 (2002).
180. Shinin, V., Gayraud-Morel, B., Gomès, D. & Tajbakhsh, S. Asymmetric division and cosegregation of template DNA strands in adult muscle satellite cells. *Nat. Cell Biol.* **8**, 677–687 (2006).
181. Smith, G. H. Label-retaining epithelial cells in mouse mammary gland divide asymmetrically and retain their template DNA strands. *Development* **132**, 681–7 (2005).
182. Xu, M. *et al.* Partitioning of Histone H3-H4 tetramers during DNA replication

dependent chromatin assembly. *Science (80-.).* **328**, 1–5 (2010).
183. Tran, V., Lim, C., Xie, J. & Chen, X. Asymmetric distribution of histones during Drosophila male germline stem cell asymmetric divisions. *Chromosom. Res.* **338**, 679–82 (2013).
184. Tran, V., Lim, C. & Xie, J. Asymmetric Division of Drosophila. *Science (80-.).* **1309**, 2010–2013 (2012).
185. Xie, J. *et al.* Histone H3 Threonine Phosphorylation Regulates Asymmetric Histone Inheritance in the Drosophila Male Germline. *Cell* **163**, 1–14 (2015).

Chapter 3

CTCF sites display cell cycle dependent dynamics in factor binding and nucleosome positioning

Marlies E. Oomen[1], Anders S. Hansen[2], Yu Liu[1], Xavier Darzacq[2], Job Dekker[1,3,]

Published in *Genome research.* (2019) 29: 236-249.

1. Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA
2. Department of Molecular and Cell Biology, Li Ka Shing Center for Biomedical and Health Sciences, CIRM Center of Excellence, University of California, Berkeley, CA 94720, USA
3. Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, MD 208156789, USA

Abstract

CCCTC-binding factor (CTCF) plays a key role in the formation of topologically associating domains (TADs) and loops in interphase. During mitosis TADs are absent, but how TAD formation is dynamically controlled during the cell cycle is not known. Several contradicting observations have been made regarding CTCF binding to mitotic chromatin using both genomics and microscopy-based techniques. Here we have used 4 different assays to address this debate. First, using 5C we confirmed that TADs and CTCF loops are readily detected in interphase, but absent during prometaphase. Second, ATAC-seq analysis showed that CTCF sites display greatly reduced accessibility and lose the CTCF footprint in prometaphase, suggesting loss of CTCF binding and rearrangement of the nucleosomal array around the binding motif. In contrast, transcription start sites remain accessible in prometaphase, although adjacent nucleosomes can also become repositioned and occupy at least a subset of start sites during mitosis. Third, loss of site-specific CTCF binding was directly demonstrated using CUT&RUN. Histone modifications and histone variants are maintained in mitosis, suggesting a role in bookmarking of active CTCF sites. Finally, live-cell imaging, fluorescence recovery after photobleaching and single molecule tracking showed that almost all CTCF chromatin binding is lost in prometaphase. Combined, our results demonstrate loss of CTCF binding to CTCF sites during prometaphase and rearrangement of the chromatin landscape around CTCF motifs. This, combined with loss of cohesin, would contribute to the observed loss of TADs and CTCF loops during mitosis and reveals that CTCF sites, key architectural cis-elements, display cell cycle stage-dependent dynamics in factor binding and nucleosome positioning.

Introduction

Several studies have observed a key role of CCCTC-binding factor (CTCF) in organizing the linear genome in topologically associating domains (TADs) and loops in interphase vertebrate cells[1–3]. CTCF is an 11 zinc finger protein that binds a well-defined motif to which it can bind only in one direction[4]. Nucleosomes flanking CTCF bound sites are strongly positioned [5]. In addition, flanking nucleosomes contain histone modifications such as H3K4 methylation and histone variants such as H2A.Z[6,7]. Although CTCF has about 42,000 predicted binding sites in the human genome, only a subset of CTCF sites are bound in a given cell type.

It has been proposed that topologically associating domains (TADs) and CTCF-loops are formed as a result of cohesin-dependent loop extrusion[8–10]. According to this model, when cohesin is loaded on the chromatin, it will be able to form a loop between two loci and will keep extruding until it is blocked by CTCF, which will function as a boundary element. Whether cohesin is blocked by bound CTCF and whether two CTCF-occupied sites can form a loop depends on the orientation of the CTCF motif: looping is mostly observed between CTCF sites in a convergent orientation[11–14]. CTCF loops often define TADs that are implicated in gene regulation.

Chromosome organization changes dramatically during mitosis. The structural features of interphase chromosomes described by 5C and Hi-C, such as TADs and A- and B-compartments are lost in prometaphase[15]. Current models, based on modeling Hi-C data for prometaphase cells combined with extensive earlier imaging data[16–18], propose that mitotic chromosomes are organized as arrays of nested loops that are helically arranged around a spiraling central axis[19]. These loops can be generated by a process of loop extrusion mediated by condensin complexes [8,19,20].

Although it is clear that loss of CTCF causes genome-wide loss of TADs in interphase[2], whether the loss of TADs during mitosis is due to regulation of CTCF is currently unclear. First, it is possible that condensin-mediated loop extrusion, unlike cohesin-mediated extrusion, is not blocked by CTCF. It has been shown that condensin II in interphase cells does not accumulate at CTCF sites unlike cohesion[21]. This suggests that CTCF cannot block condensin II mediated loop-extrusion. Alternatively, TAD boundaries could be absent because CTCF, and most of cohesin, dissociates from chromatin during mitosis. Along these lines, CTCF becomes highly phosphorylated in mitosis[22–24] and in vitro assays show that DNA binding capability of phosphorylated CTCF is dramatically reduced[25,26].

There have been several studies to examine chromatin/protein factor binding in mitotic cells using both microscopy and genomics techniques such as ChIP-seq and DNase I sensitivity assays. Several studies suggest that most factors

lose site-specific binding to the chromatin during mitosis[27,28]. However, other studies, mainly using imaging or western blot analysis of chromosome associated proteins, report maintenance of factor binding in mitosis[29–32]. There are several reasons that could explain these conflicting results. First, formaldehyde fixation can affect protein association with mitotic chromosomes, and therefore prevent observation of factor binding by both microscopy and ChIP-seq[29,33,34]. Additionally, when performing population wide genomic studies in mitosis, cells need to be synchronized using drugs, cell sorting or modified cell lines[19,35–37]. It is important to obtain pure synchronized populations, as contamination of interphase cells, especially in studies using immunoprecipitation, can lead to an overestimation of signal in mitosis. Furthermore, although microscopy has advantages over population wide studies, microscopy does not capture information on site specific binding of factors. Moreover, it is important to distinguish co-localization of factors with the mitotic chromatin from site specific binding, which could function as a mitotic bookmark[34,38,39]. In this study, we used a combination of several genomics techniques and live cell imaging to study cell cycle dynamics of CTCF mediated looping interactions, CTCF binding and local chromatin state flanking CTCF binding sites.

Results

5C shows loss of TADs and CTCF loops in prometaphase

It has been shown that interphase structures such as TADs and compartments, are lost in mitosis[15,19,40,41]. However, the resolution of these previous studies was not sufficient to investigate specific looping interactions, e.g. between CTCF sites. We therefore applied a targeted 5C approach[42] that allows high resolution analysis (10 to 15 kb) for domains up to several megabases in interphase and mitosis. To obtain mitotic cells, we synchronized HeLa S3 cells by first arresting cells in early S-phase using a thymidine block, followed by an arrest in prometaphase using nocodazole[15]. We confirmed cell cycle state of non-synchronous and mitotic (prometaphase) cell populations using flow cytometry (Supplemental Fig. S1) and quantified the mitotic index (percentage of cells with condensed chromosomes) using fluorescence microscopy of DAPI stained cells. HeLa S3 asynchronous populations have a mitotic index count of ~5%, whereas a nocodazole arrested culture contained 95-98% mitotic cells. We also biochemically purified mitotic chromatin from prometaphase cells[43]. When purified mitotic chromosomes were examined using fluorescence microscopy with DAPI staining, no contaminating interphase cells were detected (representative examples in Supplemental Fig. S1).

We performed 5C with a pool of primers targeting each end of each restriction fragment (a “double” alternating design[44]), which produces a complete interaction

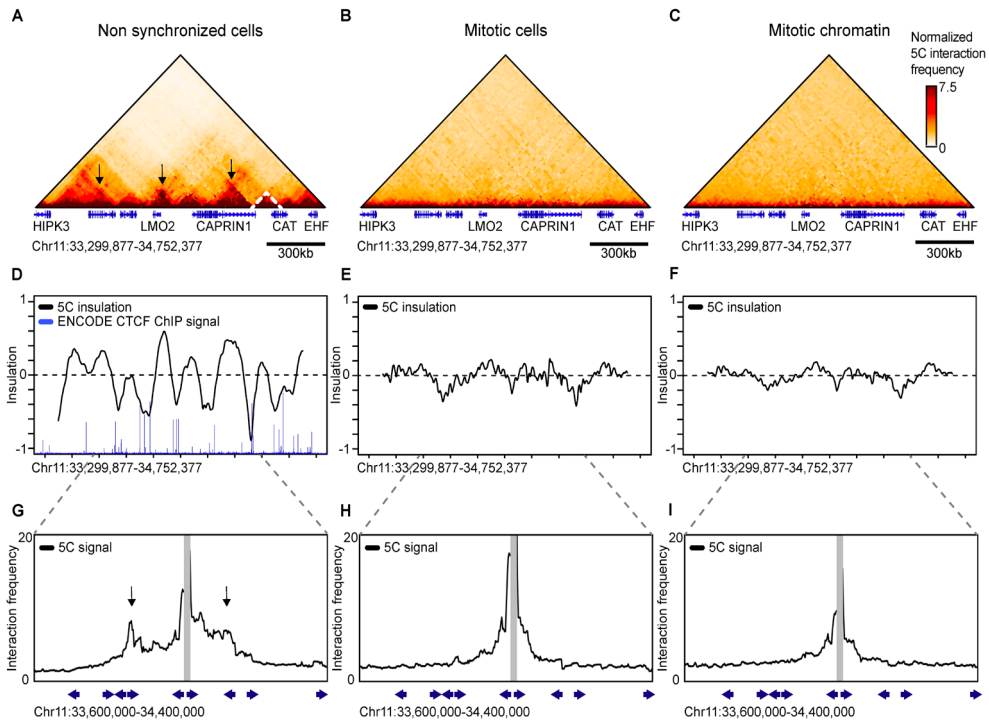


Figure 1. Topologically associating domains (TADs) and CTCF-loops are lost in prometaphase. 5C data of Chromosome 11 33,299,877-34,752,377 shows TADs (dashed lines) and CTCF loops (arrows) in interphase (**A**), however these structures are lost in nocodazole arrested mitotic cells (**B**) and purified mitotic chromatin (**C**). (**D-F**) Insulation profiles. The insulation profile for non-synchronized cells shows a strong pattern alternating peaks centered within TADs and valleys at TAD boundaries (**D**). TAD boundaries are co-localized with bound CTCF as is shown by ENCODE CTCF ChIP-seq. Insulation profiles for mitotic cells and mitotic chromatin do not show deep minima indicating TAD boundaries are absent (**E-F**). 5C interaction profiles anchored on one CTCF-bound site (15 kb bin spanning Chr11: 34,012,377-34,027,377). Peaks along these profiles (arrows) indicate CTCF-loops observed in interphase (**G**). CTCF loops are not detected in mitotic cells and mitotic chromatin (**H-I**). Blue arrows represent the position and orientation of CTCF motifs.

map for all restriction fragments throughout the two 2MB regions (Supplemental Fig. 2). In interphase cells, TADs are readily detected as domains of increased interaction frequencies between loci flanked by CTCF-bound sites (Fig. 1A, representative TAD marked with dashed line). By evaluating the insulation profile along the locus, we can identify TAD boundaries and quantify the strength of TADs[19,45]. Insulation score is low at TAD boundaries and high at loci inside TADs (Fig. 1D, an example of a TAD is indicated with white dotted lines in Fig. 1A). As has been shown before, TAD boundaries are enriched in CTCF binding[1,11,13]. CTCF-looping interactions are

detected by their appearance as “dots” of elevated interaction frequency ((Fig. 1A), several loops marked with arrows). To illustrate CTCF-loops we plotted the interaction frequencies of one CTCF site with its flanking loci (Fig. 1G). We find that interaction frequencies generally decay with genomic distance, but that peaks appear at other CTCF sites, consistent with loop formation. It has been shown that loops between CTCF sites typically occur between motifs that are in convergent orientation[12–14,46]. Looping interactions we observe in our 5C data are consistent with this (Fig. 1G).
5C analysis reveals that TADs are no longer observed in mitotic cells and mitotic chromatin (Fig. 1B-C), in agreement with previous observations[15,19]. The loss of TADs can be seen both visually in the interaction heatmap, as well as by calculation of the insulation profile (Fig. 1E-F). Importantly, looping interactions between CTCF sites are also no longer detected in mitosis (Fig. 1H-I).

Analysis of chromatin accessibility in interphase and prometaphase
To investigate chromatin characteristics at CTCF binding sites, we determined chromatin accessibility using ATAC-seq[47]. ATAC-seq fragments capture information in several different ways. First, ATAC-seq data provides information about genome-wide nucleosome positioning and spacing. This information is captured in the length distribution of all fragments (Fig. 2A). Very short fragments ranging from 24-80 bp are frequently observed. These fragments represent accessible regions in between nucleosomes or in between a nucleosome and another chromatin bound protein. Larger fragments typically form an enrichment of sizes that are multiples of approximately 195 bp. This reflects the nucleosomal array, which has been seen before in previous studies using ATAC-seq[48]. In mitosis we observed a similar nucleosomal array. However, interestingly, the nucleosomal array is more pronounced, as the enrichment for fragments that are multiples of 195 bp, represented by the peaks in Fig. 2A, is stronger. This suggests that the spacing of nucleosomes genome wide is more regular in mitosis compared to interphase.

A second type of information captured in ATAC-seq data is on accessibility of specific classes of sites. We called peaks on ATAC-seq data using HOMER[49]. ATAC-seq data have been shown to display peaks in regions of open chromatin similar to peaks generated by DNase I sensitivity assays[48]. We can separate regions of the genome based on their function assigned by ChromHMM[50,51]. ATAC-seq peaks are found in different types of functional elements, such as active promoters, enhancers or CTCF insulator regions (Fig. 2B). We then compared peaks found in interphase to peaks found in mitotic cells and mitotic chromatin. In general, there is large loss of peaks in prometaphase compared to interphase. However, when we compare which types of functional elements lose peaks, we see that active promoters maintain

significant accessibility, whereas the other functional elements examined lose most accessibility in mitosis (Fig. 2B). The maintenance of significant accessibility of promoter regions and concomitant loss of accessibility at enhancers during mitosis has been previously observed using DNase I sensitivity assays[28,52]. In addition to this, we find an even more substantial loss of accessibility at CTCF binding sites. This was observed in both mitotic cells and purified mitotic chromatin. Fig. 2C shows an example of ATAC-seq signal at individual TSSs and CTCF motif (Fig. 2C).

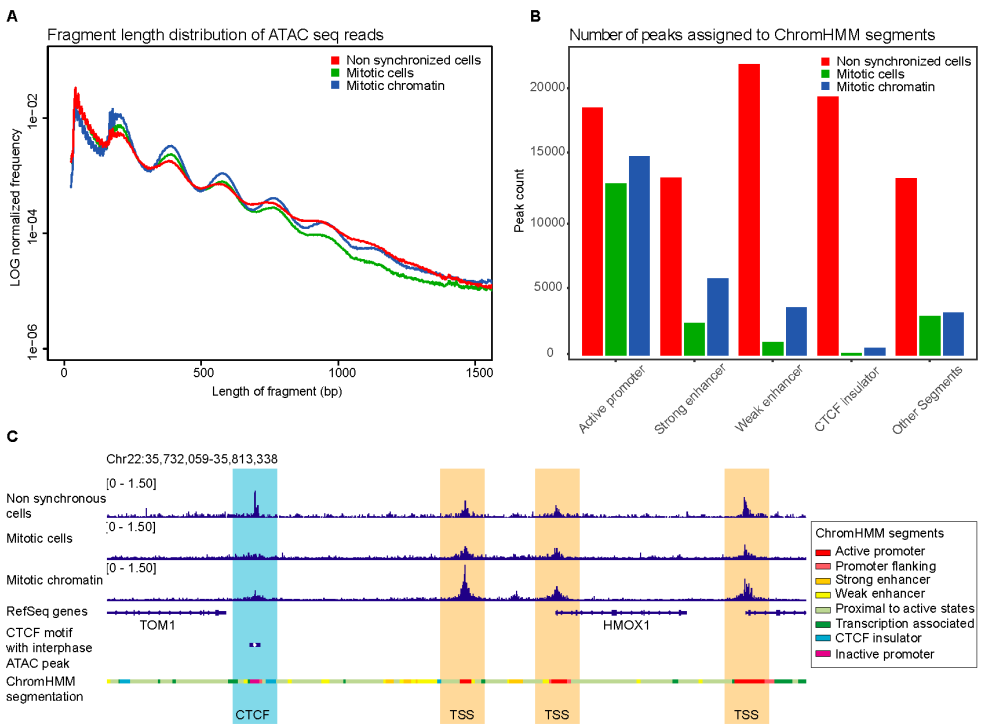


Figure 2. ATAC-seq data show accessibility at CTCF sites and enhancers is reduced in mitosis, though maintained at TSSs. (**A**) Fragment length distribution of ATAC-seq reads genome-wide in non-synchronized cells, mitotic cells and purified mitotic chromatin. (**B**) Distribution of number of peaks called in non-synchronized cells, mitotic cells and mitotic chromatin and their position on ChromHMM segments. (**C**) Example of a representative region illustrating maintenance of accessible chromatin at TSSs in mitotic conditions while ATAC-seq signal is lost at CTCF motifs.

V-plots reveal protein occupancy at CTCF motifs and nucleosome positioning in interphase

ATAC-seq also captures information on protein binding footprints and nucleosome positioning flanking hypersensitive sites. This information is captured by the length of fragments at a site of interest. We used V-plots to represent this data. V-plots have been used to plot MNase data as a way to display chromatin binding by site specific factors and positioning of flanking nucleosomes on different length scales [53]. We have plotted our ATAC-seq data with the fragment length on the y-axis and positioning of the midpoint of the fragment on the x-axis representing the distance to the binding site of interest. In order to investigate local chromatin state at and around sites of CTCF binding, we made V-plots of our data on CTCF motifs[4] that are accessible, i.e. have an ATAC-seq peak, in interphase (Fig. 3A). In non-synchronized cells, we observe an enrichment of 80-100 bp fragment at the CTCF binding sites (asterisk), which represents the footprint of bound CTCF and possibly associated proteins like cohesin. Similar footprints have been found using MNase digestion[5]. The footprint can also be observed when the fragment length distribution is plotted for all fragments with their midpoint on bound CTCF motifs (Fig. 3D, asterisk). When we plot the lengths of reads with one end on a CTCF motif, we observe that many fragments have a short length (Fig. 3E). These represent fragments generated by pairs of ATAC cleavages in between bound CTCF and the flanking nucleosomes.

The second type of information that can be derived from V-plots is regarding the positioning of flanking nucleosomes. For bound CTCF motifs in interphase, we observe enriched dots on the arms of the V in the V-plot (Fig. 3A, arrows). These dots indicate strong positioning of several nucleosomes flanking the bound CTCF motif, consistent with previous MNase results[5]. The series of enriched signals represent ATAC-seq fragments covering one, two, three and four flanking nucleosomes, but are longer than expected for a typical nucleosomal array. We attribute this size discrepancy to mean that some of these fragments can cover not only one or more nucleosomes, but also the flanking bound CTCF site. This becomes even clearer when fragment lengths are plotted of reads which have one of their read ends near a bound CTCF motif (Fig. 3E, red arrow). We observe an enrichment of fragments that are around 220 bp, instead of the expected 195 bp for a canonical mononucleosome. Similar results were found when V-plots were made for all CTCF motifs or for CTCF motifs with peaks from available CTCF ChIP-seq ENCODE data[51] (Supplemental Fig. S3A). Taken together, these analyses show that during interphase CTCF is bound to its motif and nucleosomes flanking CTCF-occupied sites form a regularly spaced array.

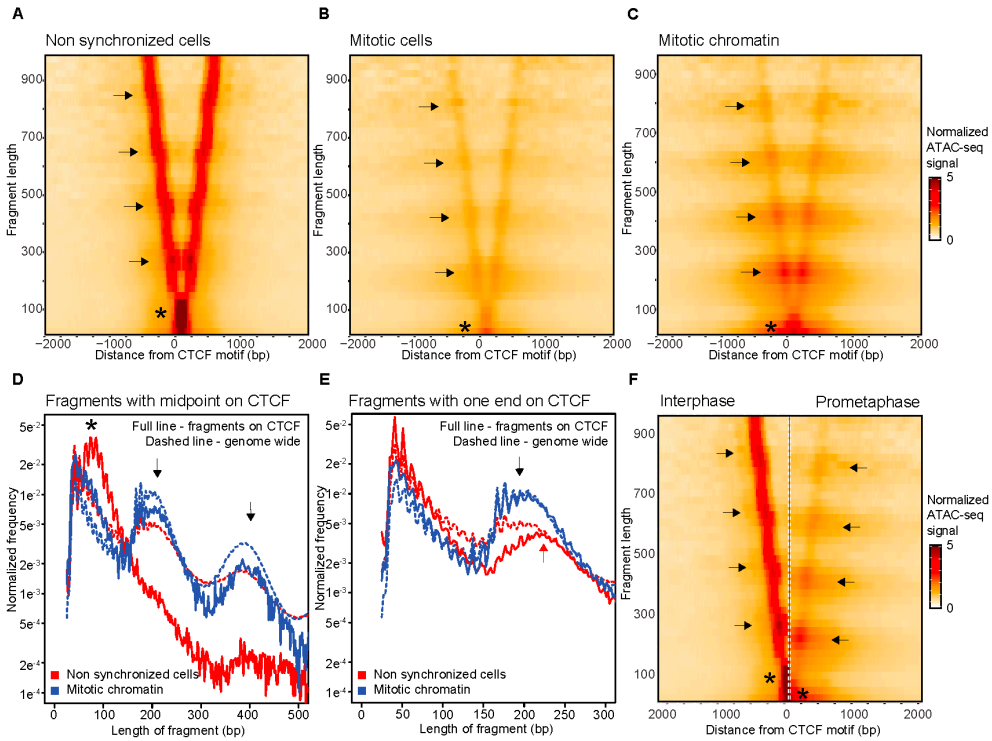


Figure 3. ATAC-seq data represented in V-plots show loss of CTCF binding in mitosis and rearrangement of nucleosomes flanking CTCF motifs. (**A-C**) ATAC-seq data represented in V-plots aggregated at CTCF sites. The lengths of ATAC-seg reads are plotted on the y-axis, and the distance between their midpoints and the CTCF motif is shown on the x-axis. (**A**) V-plot for interphase ATAC-seq data. Asterisk indicates the CTCF footprint. Enriched dots along the arms of the V (arrows) represent strongly positioned nucleosomes. (**B-C**): V-plots for ATAC-seq data from mitotic cells and purified mitotic chromatin aggregated at CTCF motifs with interphase ATAC-seq peaks. Asterisks mark the loss of the CTCF footprint. Arrows indicate positioning of flanking nucleosomes. (**D**) Distribution of fragment lengths of reads that have their midpoint on a CTCF motif. Dashed line: genome-wide average read length distribution. Arrows: read lengths representing 1 and 2 nucleosomes. (**E**) Distribution of fragment length of reads with either read end near a CTCF motif with interphase ATAC-seq peak compared to the genome-wide average (dashed line). In interphase reads representing one flanking nucleosome are longer (red arrow) as compared to reads representing one flanking nucleosome in mitosis (black arrow). (**F**) Side-by-side comparison of V-plots for non-synchronized and mitotic chromatin. The shift in nucleosome positioning is highlighted using arrows. Asterisks mark loss of CTCF footprint in mitosis.

V-plots show loss of binding at CTCF motifs and rearrangement of nucleosomes in prometaphase

We then created V-plots for ATAC-seq data generated from mitotic cells and purified mitotic chromatin at CTCF motifs that have an interphase ATAC-seq peak (Fig. 3B-C).

As expected from the peak calling assessment, overall accessibility is reduced (and no longer significant) in mitotic conditions. However, some accessibility remains. We do note that ATAC-seq signal at CTCF sites for purified mitotic chromatin is stronger compared to ATAC-seq signal in mitotic cells. This higher signal to noise ratio is possibly caused by lower background signals overall due to stronger nucleosome interactions as a result of detergents in the chromosome purification buffer[43]. Several additional features are observed. First, there is a change in the CTCF footprint. The enrichment of 80-100 bp fragment lengths observed in interphase is no longer present. Instead there is an enrichment of small fragments around 25-75 bp (asterisks). This suggests that CTCF is no longer bound to the motif, where it protected the site from ATAC-seq cleavage in interphase. We observed this in both mitotic cells and purified mitotic chromatin. The loss of 80-100 bp fragments is also observed when the length distribution is plotted for fragments with their midpoint on interphase bound CTCF motifs (Fig. 3D, compare to interphase length distribution marked by asterisk).

Second, the positioning of nucleosomes around interphase bound CTCF motifs also changes during mitosis (Fig. 3B-C, arrows). Loss of CTCF binding would create a large accessible region between the flanking nucleosomes, which can in turn cause nucleosomes to move inwards. This movement of nucleosomes would cause the linkers in the nucleosomal array to become larger than average. We observe this phenomenon in several ways. First, nucleosomes are able to occupy CTCF sites in mitosis. This can be seen by a gain of mono- and dinucleosome sized fragments in the length distribution plot in Fig. 3D (arrows). Nucleosome occupancy at the CTCF motif is not observed in interphase when CTCF typically occupies the motif. Second, as discussed above, in interphase we observe a peak at 220 bp in the fragment length distribution for fragments with one end in a CTCF motif that is larger than typical for a mononucleosome (Fig. 3E, red arrow). In mitosis, this mononucleosome peak becomes more similar to the genome-wide average (195 bp); again suggesting that CTCF is no longer bound (Fig. 3E, black arrow). This size discrepancy becomes even more obvious when V-plots for non-synchronized cells and purified mitotic chromatin are plotted side by side (Fig. 3F, compare arrows). Lastly, there is a change in nucleosome spacing and positioning during mitosis (Fig. 3B-C). In interphase, the locations of nucleosomes are observed as enriched dots along the arms of the V with a strong positioning relative to the CTCF motif. Conversely, in mitosis, the enriched dots become less pronounced and we observe horizontal bands of elevated fragment frequency in the heatmap running several kb up and down stream of the CTCF motif. Given that these V-plots are normalized for genome-wide average fragment length frequency, a banding pattern only emerges when the spacing between several of the flanking nucleosomes differs

and/or is more variable from the genome-wide average. Specifically, the banding pattern indicates that there is an increase in nucleosomes spaced by larger linkers than the genome-wide average (Fig. 3E, dashed line). These observations confirm that several flanking nucleosomes are able to move inwards, creating longer linkers between them. This creates a local nucleosomal array around CTCF sites with larger than average linkers.

Taken together, our data suggest that CTCF is no longer bound in prometaphase and nucleosomes rearrange. To exclude that there are certain subgroups of CTCF motifs that maintain binding in mitosis, we plotted V-plots for several obvious classes of CTCF motifs, e.g. motifs proximal or distal to TSSs or CTCF motifs that maintain an ATAC-seq peak in mitosis. In addition, we use *k*-means clustering for ATAC fragments of 75-150 bp at CTCF sites to determine whether there are subgroups of motifs that maintain a CTCF footprint in mitosis (Supplemental Fig. S4). None of these methods found a specific group of CTCF motifs that maintain CTCF binding. Finally, we confirmed the loss of CTCF binding and rearrangement of nucleosomes in multiple differentiated cell lines (Supplemental Fig. S5-6).

Transcription start sites maintain significant accessibility, but show variable loss of factor binding and nucleosome repositioning in prometaphase

We also made several observations about the chromatin landscape at accessible TSSs in interphase and mitosis (Supplemental Fig. S7). First, we see an enrichment of fragments with a length below 150 bp compared to the genome-wide average in interphase (Supplemental Fig. S7A, asterisk). We do not observe an enrichment of fragments of a defined size as we do for CTCF (compare supplemental Fig. S7D to Fig. 3D). This can be explained by the fact that contrary to CTCF sites, TSSs can be bound by a more diverse set of proteins. Second, we do not observe enrichments of nucleosome sized fragments centered on the TSS (Supplemental Fig. S7D), confirming that in interphase typically no nucleosomes are bound at TSSs. Furthermore, the length of fragments covering the two nucleosomes directly flanking the TSS appears less defined compared to those flanking CTCF sites (compare Fig. 3E to Supplemental Fig. S7E, red arrow). There is a local enrichment of fragments a little over 200 bp in length. These fragments represent ATAC-seq products covering factors bound to the TSS plus the neighboring nucleosome, resulting in a length that is over the genome-wide expected 195 bp covering a single nucleosome. In addition, the enrichment of fragments over 200 bp is less pronounced compared to genome-wide average for single nucleosome fragments, suggesting that heterogeneity in factors binding to the TSS can result in a wider distribution of fragment lengths covering the TSS and the flanking nucleosome.

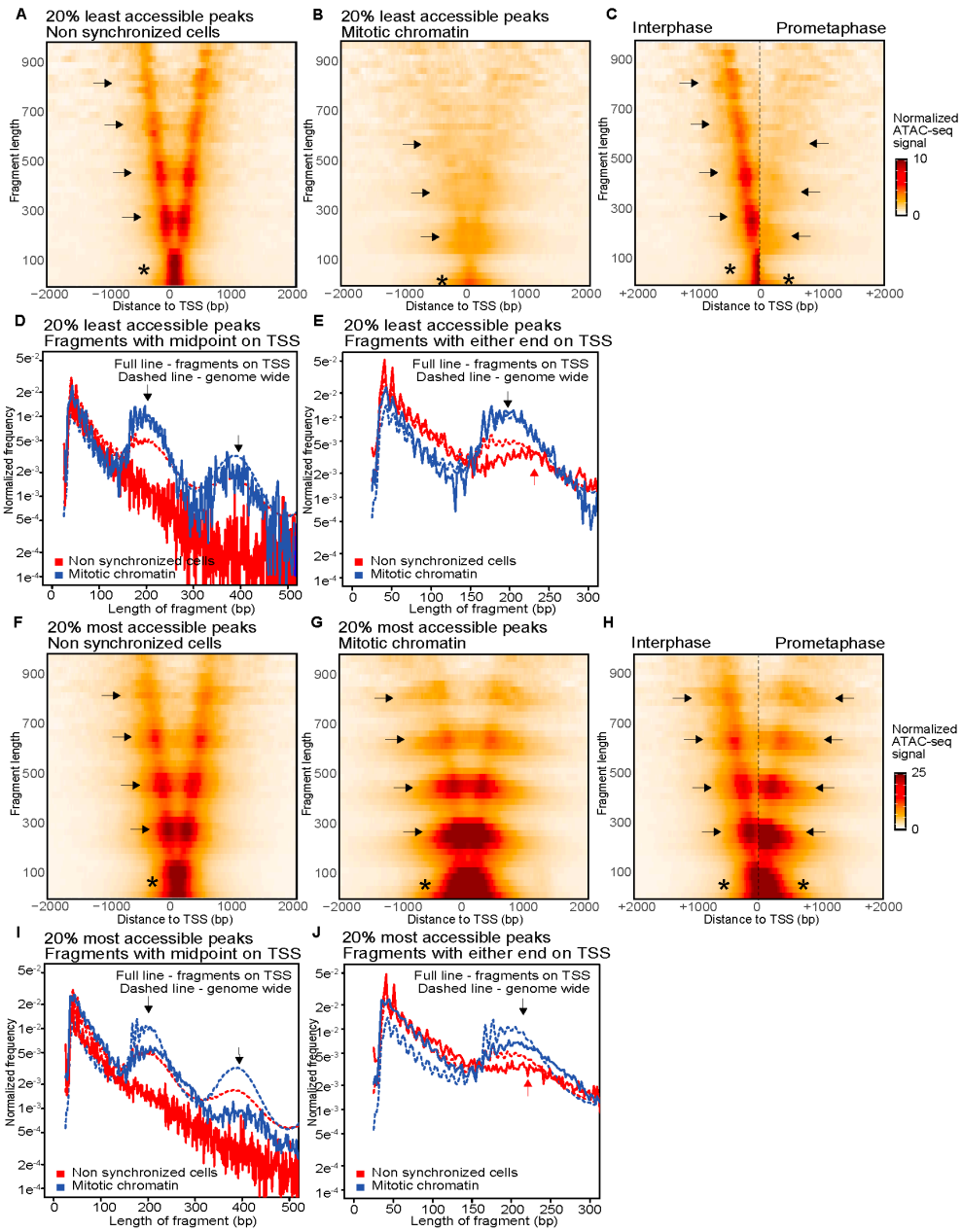


Figure 4. ATAC-seq V-plots show nucleosome repositioning, while accessibility is maintained at transcription start sites. Panels **A-E** show ATAC-seq data for the set of TSSs that display the lowest 20 percent accessible peaks in mitotic chromatin, panels **F-J** show ATAC-seq data for the set of TSSs that display the highest 20 percent accessible peaks in mitotic chromatin (**A-B**) V-plots of ATAC-seq signal aggregated at TSSs in non-synchronized cells (**A**) and mitotic chromatin (**B**). (**C**) Side-by-side comparison of the V-plots of low accessible TSSs for non-synchronized cells and mitotic chromatin. (**D**) Distribution of fragment length of reads

with their midpoint on a TSS compared to the genome-wide average (dashed line). Arrows: read lengths representing 1 and 2 nucleosomes. (**E**) Distribution of fragment length of reads with either end in a TSS in mitotic chromatin compared to the genome-wide average (dashed line). In interphase reads representing one flanking nucleosome are longer (red arrow) as compared to reads representing one flanking nucleosome in mitosis (black arrow). (**F**-**G**) V-plots of ATAC-seq signal aggregated at TSSs that display the highest 20% accessibility in mitotic chromatin for non-synchronized cells (**F**) and mitotic chromatin (**G**). (**H**) Side-by-side comparison of the V-plots shown in (**F**) and (**G**). (**I**) Distribution of fragment length of reads with their midpoint on a TSS with highest 20% highest accessible peaks in mitotic chromatin compared to genome-wide average (dashed line). Arrows: read lengths representing 1 and 2 nucleosomes. (**E**) Distribution of fragment length of reads with either end in a TSS with highest 20% highest accessible peaks in mitotic chromatin compared to the genome-wide average (dashed line). For mitotic chromatin reads representing one flanking nucleosome are of similar length (black arrow) as compared to reads representing one flanking nucleosome in interphase (red arrow).

In mitosis, complementary to our peak calling analysis (Fig. 2B-C), V-plots show that overall accessibility at TSSs is reduced, though maintained at higher levels than at CTCF sites (compare Fig. 3B-C to Supplemental Fig. S7B-C). Interestingly, we observe that in mitotic chromatin, nucleosome sized fragments are observed at the TSS (Supplemental Fig. S7D, arrows), suggesting that nucleosomes start occupying TSSs. This led us to explore further how mitotic accessibility at TSSs is related to nucleosome positioning. We compared the accessibility and nucleosome positioning in the top 20 percent most accessible and 20 percent least accessible TSSs in mitosis. In interphase, these two groups behave highly similarly, except for signal intensity as expected (compare Fig. 4A and D to Fig. 4F and I). In mitosis, both groups of TSSs can be occupied by nucleosomes, however the level of nucleosome occupancy is inversely related to the level of remaining accessibility in mitosis (Fig. 4D and I, black arrows). For the group of TSSs with low levels of remaining accessibility, the nucleosome occupancy in mitosis is comparable to the genome-wide average (Fig. 4D, dotted line), while the group of TSSs with the highest remaining accessibility in mitosis, the nucleosome occupancy is lower than the genome-wide average (Fig. 4I, dotted line). During mitosis, the set of TSSs with the lowest accessibility resembles CTCF sites in several ways. First, while in interphase the fragments representing mononucleosomes flanking the TSS are longer than genome-wide average, in mitosis this length shortens to the canonical 195 bp (Fig. 4E, compare black and red arrows). This shortening of fragments covering flanking nucleosomes can also be observed in V-plots (Fig. 4C, compare arrows). Lastly, as for CTCF sites, we observe a gain of horizontal banding pattern in the V-plot, again indicating that there is an increase in nucleosomes spaced by larger linkers than the genome-wide average. Importantly and in contrast to CTCF sites, these TSSs are significantly more accessible than the genome-wide average, despite the fact that

this set of TSSs is now occupied by nucleosomes.

The set of TSSs that maintain high levels of accessibility in mitosis however behaves differently. For this set of TSSs we continue to observe an enrichment of short fragments (below 150 bp) compared to the genome-wide average and larger fragments covering the flanking nucleosomes (Fig. 4J, arrows), suggesting that some factors remain bound to the TSSs. This is also observed by a lack of downwards shift of nucleosome sized fragments along the arms in the V (Fig. 4H, compare arrows). However, as indicated above, some increase in nucleosome occupancy at the TSS is observed in mitosis, suggesting that this set of TSSs is heterogeneous with some sites gaining nucleosomes, while other sites maintain interphase nucleosome positioning. The presence of a subset of TSSs with repositioned nucleosomes in mitosis is also reflected in the gain of horizontal bands in the V-plot (Fig. 4H). As discussed above, the horizontal banding pattern is a result of repositioned nucleosomes that have altered linker lengths than the genome-wide average.

CUT&RUN directly observes loss of CTCF binding in prometaphase

Next, we set out to detect CTCF binding directly. Given concerns about formaldehyde induced artifacts that affect protein binding to mitotic chromosomes (discussed above, and[29,33,34]), we decided not to use ChIP-seq. Instead, we use CUT&RUN[54]. CUT&RUN uses a protein A-MNase fusion to target proteins of interest labelled by primary antibodies. Upon addition of calcium ions, protein A-MNase will digest DNA bound by the protein of interest, after which short digested DNA fragments can be isolated, amplified and sequenced. In contrast to ChIP-seq, CUT&RUN can be done on unfixed cells and it does not require any pull down of molecules. CUT&RUN signal for CTCF shows a very clear peak at accessible CTCF motifs in interphase. However, in mitosis, the signal diminishes to almost background levels (Fig. 5A). Even prolonged digestion did not reveal CTCF binding in mitotic (prometaphase) cells (Supplemental Fig. S8). We analyzed CTCF binding in different obvious classes of CTCF motifs, e.g. sites with an ENCODE CTCF ChIP-seq peak or sites proximal to TSSs (Supplemental Fig. S9). None of these sites showed CTCF CUT&RUN signal in prometaphase cells. Furthermore, when we represent CTCF CUT&RUN signal of each motif sorted on signal strength, there are no CTCF sites found that maintain CTCF binding (Supplemental Fig. S10). This suggests that CTCF binding does not just become weaker in prometaphase, but that essentially all motif-specific binding is lost.

Next, we wanted to determine whether CTCF gains binding at other sites specifically in prometaphase. Whereas de novo peak calling of CTCF CUT&RUN data in interphase identified 7824 peaks, in mitotic cells only 107 peaks were

identified (Supplemental Fig. S11). This suggests there is very little site specific binding in mitosis. It has been described that CTCF becomes highly phosphorylated in mitosis[22]. To ensure the loss of CUT&RUN signal is not caused by the inability of the CTCF antibody to recognize phosphorylated CTCF, we performed western blot with lysates of non-synchronized cells, mitotic cells and mitotic chromatin (Supplemental Fig. S12). The antibody detected CTCF in both non-synchronized and mitotic cell extracts. Interestingly, we did not detect CTCF in extracts from mitotic chromatin (Supplemental Fig. S12). This indicates again that CTCF binding in mitosis is lost and when isolating chromatin from prometaphase arrested cells, CTCF is not co-purified with the chromatin.

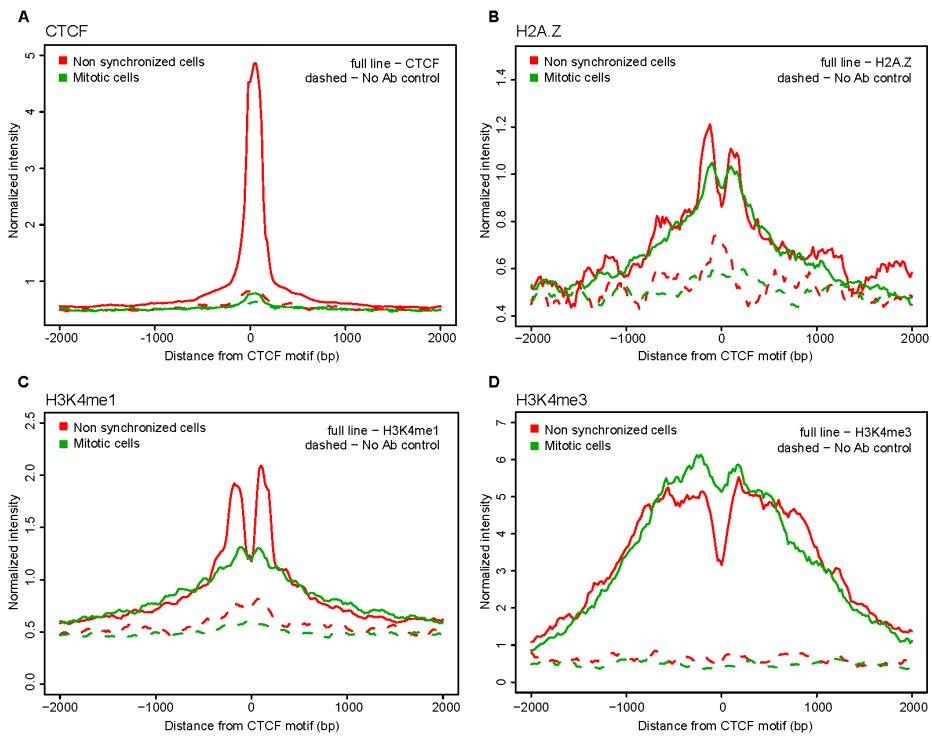


Figure 5. CUT&RUN data shows loss of CTCF binding at CTCF motifs in mitosis, while H2A.Z and H3K4 methylation marks are maintained. (**A**) Aggregation plot of CTCF CUT&RUN signal of reads shorter than 120 bp after 10 min digestion. Reads are aggregated on CTCF motifs that display an interphase ATAC-seq peak. (**B-D**) CUT&RUN signal for H2A.Z (**B**), H3K4Me1 (**C**), H3K4Me3 (**D**) of reads longer than 120bp after 30 min digestion. Reads are aggregated at CTCF motifs with interphase ATAC-seq peak.

Histone marks and variants at active CTCF sites are maintained in mitosis

There are several histone characteristics described that are associated with CTCF bound motifs in interphase. The histone variant H2A.Z is often observed in nucleosomes flanking CTCF sites and TSSs[5,7,55]. Additionally, several histone modifications have been found around bound CTCF motifs; for example H3K4me1 for motifs distal to TSSs and H3K4me3 for motifs proximal to TSSs[5]. In order to assay the presence of these histone marks and variants, we performed CUT&RUN. For H2A.Z and H3K4 mono- and trimethylation we observe two distinct peaks around the CTCF motif in non-synchronized cells (Fig. 5B-D). This confirms our observations in V-plots of ATAC-seq data, where the CTCF flanking nucleosomes are strongly positioned relative to distance from the CTCF motif in interphase due to CTCF binding.

Many histone marks and variants have been suggested to serve as mitotic bookmarks for factor binding (reviewed in[56]). Both H2A.Z and H3K4 methylation states have been observed throughout the cell cycle[7,31,55,57]. Using CUT&RUN we observed that H2A.Z binding and H3K4 methylation states are maintained at CTCF sites during mitosis, even though CTCF binding is temporarily lost (Fig. 5B-D). We note that the level of H3K4me1 is reduced during mitosis, in contrast to H3K4me3 for which signal strength is similar in interphase and mitosis. It has been previously shown that H3K4me3 typically gets rapidly reestablished on histones on both sister chromatids in S-phase, while H3K4me1 levels do not get restored until after cell division[58]. A 50% reduction in H3K4me1 levels as a result of dilution over the two sister chromatids is consistent with the reduced levels observed by CUT&RUN. As described, we observe two strong peaks in H2A.Z and H3K4 methylation signal flanking the CTCF motif in interphase, however in mitosis these peaks become weaker and form two broader peaks with a less defined valley in between. This is in concordance with ATAC-seq data, where we observed rearrangement of the nucleosomes and filling in of the nucleosome depleted regions (NDRs) at CTCF in mitosis.

We then analyzed H2A.Z and H3K4me3 levels at TSSs detected by CUT&RUN in interphase and mitosis. In line with previous studies and our CUT&RUN data at and around CTCF sites, we found that H2A.Z and H3K4me3 levels are maintained at TSSs in prometaphase (Supplemental Fig. S13;[7,55–59]). Interestingly, but not surprisingly, we again observe loss of depletion of histone signal at TSSs in prometaphase, similar to CTCF sites. This is an indication that nucleosomes reposition and occupy at least a subset of TSSs, consistent with our observations in ATAC-seq V-plots (Fig. 4).

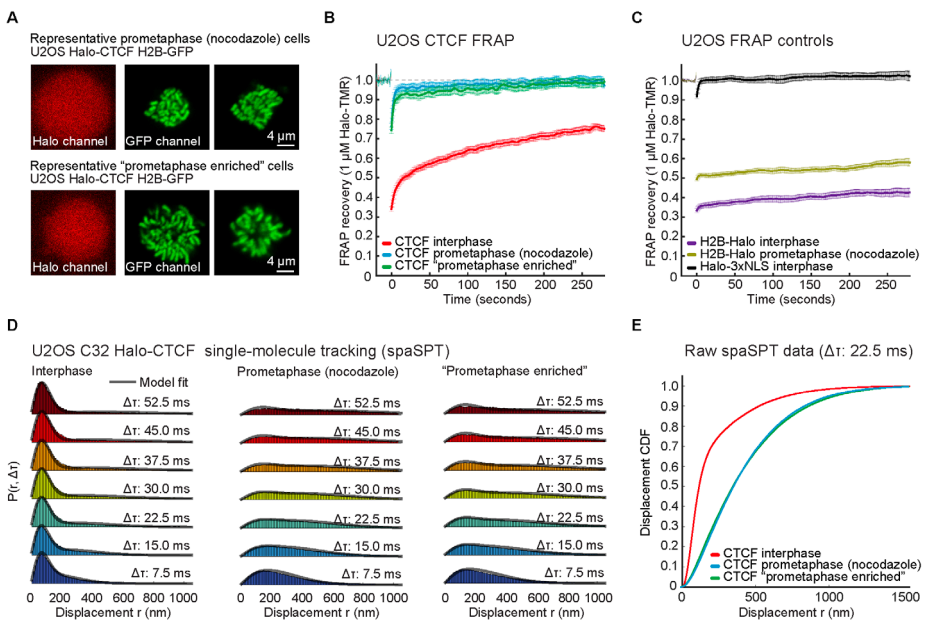
Live cell imaging shows that CTCF is not enriched on prometaphase chromosomes and that essentially all specific binding is lost

Genomic and imaging studies have occasionally reported contradicting findings when it comes to factor binding to prometaphase chromatin. To independently validate our observation using genomics techniques that essentially all site specific CTCF binding is lost during prometaphase, we turned to live-cell imaging and applied a three-pronged approach based on long-term time-lapse imaging, Fluorescence Recovery After Photobleaching (FRAP) and Single-Particle Tracking (SPT). We previously generated and validated a U2OS cell line where all CTCF alleles have been Halo-tagged (C32 Halo-CTCF[60]). To visualize mitotic cells, we additionally stably integrated a histone H2B-GFP transgene. We performed ATAC-seq on this cell line and observed the expected loss of accessibility at CTCF sites in mitosis and loss of the CTCF footprint represented in V-plots (Supplemental Fig. S5E-F and Supplemental Fig. S6G-H). First, we used multi-hour time-lapse fluorescence microscopy to observe Halo-CTCF (Supplemental Movie S1-2) and H2B-GFP (Supplemental Movie S2) in actively dividing cells. Although CTCF was clearly enriched on mitotic chromosomes during most phases of mitosis (e.g. telophase), CTCF localization appeared to be diffuse during prometaphase. Second, to quantify CTCF binding dynamics we used FRAP. As for the genomics experiments, we used nocodazole to arrest cells in prometaphase. As we observed with time-lapse microscopy, CTCF showed a diffuse localization without clear enrichment on mitotic chromosomes during prometaphase (Fig. 6A; upper panel). To rule out any artifacts due to nocodazole drug treatment, we also identified cells in prometaphase without drug-treatment based on their H2B-GFP localization ("prometaphase enriched") and similarly observed diffuse CTCF localization without enrichment on chromatin.

We then performed FRAP, bleaching a ~1 μm circle and quantifying the recovery[61]. In interphase, Halo-CTCF showed slow recovery consistent with a high fraction of chromatin binding to specific CTCF sites with an apparent residence time of a few minutes (Fig. 6B[60]). However, in prometaphase, Halo-CTCF showed ~90-95% recovery within seconds as well as a fraction of the population (~5-10%) which showed slower recovery suggesting stable binding by only a small subpopulation (Fig. 6B). We validated our FRAP approach and show that the difference in recovery was not due to improper drift-correction using histone H2B controls (Fig. 6C; note that the H2B bleach depth is slightly lower in prometaphase due to "gaps" between chromosomes, but that the rate of recovery is unchanged). Thus, in prometaphase, although a small population approaching our detection limit (~5-10%) does appear to bind specific sites with a residence time in the minute range, the vast majority of specific CTCF binding is clearly lost, which is consistent with the genomics experiments. We

conclude that nearly all specific CTCF binding is lost in prometaphase.

Third and finally, we sought to verify this using an independent technique and used stroboscopic photo-activation single-particle tracking (spaSPT). spaSPT makes it possible to observe single CTCF molecules in live cells and to visualize both bound (specific and non-specific) and freely diffusing molecules without motion-blur bias[60,62]. We then tracked single CTCF molecules at 134 Hz in interphase (Supplemental Movie S3), in nocodazole-arrested prometaphase (Supplemental Movie S4) and in "prometaphase-enriched" cells (Supplemental Movie S5) and quantified the distribution of displacements between frames. Analysis of the



F

U2OS C32 Halo-CTCF Spot-On inferred results	Fraction bound (total)	D_{FREE} (µm²/s)
Interphase (non-synchronized)	54.5% ± 7.0%	2.55 ± 0.29
"Prometaphase enriched"	13.8% ± 1.2%	3.11 ± 0.43
Prometaphase (nocodazole)	10.5% ± 2.4%	3.31 ± 0.33

Figure 6. Live-cell imaging shows large loss of CTCF binding in mitosis. (**A**) Halo-CTCF and H2B-GFP localization of representative U2OS cells for prometaphase arrested cells and non-arrested cells selected as prometaphase ("prometaphase enriched"). (**B**) FRAP for Halo-CTCF in interphase, prometaphase arrested and prometaphase enriched cells. (**C**) Controls showing FRAP for H2B-Halo in interphase and prometaphase arrested cells and for Halo-3xNLS in interphase cells. (**D**) Single particle tracking displacement statistics for halo-CTCF in different timeframes in interphase, prometaphase arrested and prometaphase enriched cells. (**E**) Displacement cumulative distribution function (CDFs) derived from single particle tracking at Δτ 22.5ms for interphase, prometaphase arrested and prometaphase enriched cells. (**F**) Fraction bound of Halo-CTCF calculated using the Spot-On model in interphase, prometaphase arrested and prometaphase enriched cells[62].

cumulative distribution function (CDF) revealed that nearly all CTCF chromatin binding was lost in prometaphase (Fig. 6D-E; bound molecules typically appear below 150nm). To quantitatively analyze the spaSPT data we fit a 2-state kinetic model (Spot-On) wherein CTCF can exist in either a bound (both specific and transient non-specific binding) and free state (Fig. 6D[60,62]). This revealed that whereas ~55% of CTCF molecules are chromatin bound in interphase, only ~10-14% are bound in prometaphase (Fig. 6F). Since this bound population includes transient non-specific binding, these results independently confirm the FRAP results (Fig. 6B). In summary, our live-cell imaging results independently confirm that nearly all specific CTCF binding is lost during prometaphase and, additionally, that this observation is not an artifact of nocodazole treatment.

Discussion

Our study presents a comprehensive analysis of CTCF binding and CTCF motif-flanking nucleosomes in mitosis using several complimentary techniques and multiple differentiated cell lines. Genomics and live cell imaging techniques show that overall CTCF binding is lost in mitosis, as are TADs and loops between CTCF sites. We did not find any subgroup of CTCF sites that maintain binding in mitosis, nor did we find any new sites of mitotic CTCF binding. This, together with the known loss of cohesin during prophase, could explain why TADs and CTCF loops are not observed in mitosis[63–66]. In addition to this, we found that nucleosomes flanking interphase CTCF sites rearrange in prometaphase, resulting in nucleosomes occupying the CTCF motif and forming an array of nucleosomes with larger and more variable linkers (Fig. 7). Furthermore, similar phenomena can occur at TSSs, although these sites remain hyperaccessible as observed by ATAC-seq. Epigenetic marks, such as histone variant H2A.Z and H3K4 methylation marks, are maintained at both CTCF sites and TSSs in mitosis.

Previous studies found evidence for CTCF binding to mitotic chromosomes using imaging and chromatin fractionation approaches[30,31,67]. Additionally, proteomics studies of isolated mitotic chromatin detect CTCF, although at reduced levels compared to interphase chromatin[19,68]. Importantly however, all these approaches measure general mitotic chromatin association and do not capture information on site specific binding[34,38,39]. Our live cell imaging data also indicate that CTCF remains associated with chromatin during several stages of mitosis, however in prometaphase CTCF binding dynamics are changed and the vast majority of specific and stable binding is lost. This is complementary to our findings using genomics techniques, where we also observe loss of CTCF binding at interphase sites and we do not find any mitotic site specific binding. It is possible that CTCF remains associated

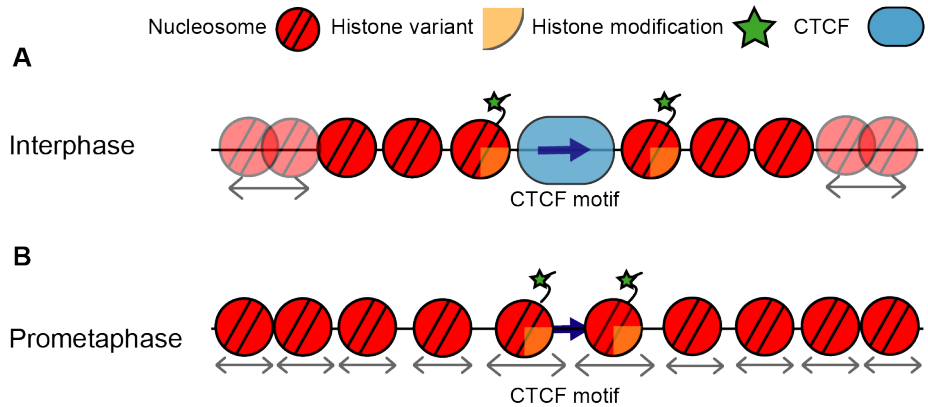


Figure 7. Dynamics of CTCF binding and chromatin organization around active CTCF motifs throughout the cell cycle. (**A**) In interphase CTCF is bound to its motif and flanking nucleosomes are strongly positioned relative to the motif, in contrast to nucleosomes further away from the motif, as indicated by arrows. Flanking nucleosomes are characterized by histone variants and modifications. (**B**) In prometaphase, CTCF binding is temporarily lost and nucleosomes rearrange to fill in the nucleosome depleted region at the CTCF motif. This increases the linker lengths between adjacent nucleosomes. Epigenetic marks however are maintained, possible functioning as bookmarks that enable inheritance of active CTCF motifs throughout the cell cycle. Arrows underneath nucleosomes indicate that the position of these nucleosomes can vary between cells.

with mitotic chromatin, although in a non-specific and highly dynamic manner. First, mitotic chromatin retention could enable proper segregation of CTCF levels over the daughter cells. Second, maintained chromatin association can enable efficient reestablishment of CTCF binding upon mitotic exit. A recent study observed a rapid raise of CTCF levels associated to the chromatin in late anaphase, as for many other chromatin binding factors[67]. The hypothesis that chromatin binding factors retaining chromatin association in mitosis, though losing motif specific binding, has been tested using imaging techniques in recent studies[34,39]. Additionally, we note that CTCF may show cell type specific dynamics in prometaphase. Our study observes CTCF cell cycle dynamics of differentiated cells, using both transformed and non-transformed cell lines.

Transcription start sites are highly accessible and free of nucleosomes in non-synchronized cells that are mostly in interphase. In contrast to CTCF sites, TSSs remain hyperaccessible during mitosis. This has also been observed by DNase I sensitivity assays[27,28]. Here we find that despite remaining highly accessible, nucleosome-sized ATAC-seq fragments are detected at TSSs in mitosis indicating that nucleosomes are able to occupy these sites, and that the spacing between

flanking nucleosomes becomes larger and more variable compared to genome-wide average, similar to what we observed at and around CTCF sites. These observations are consistent with a recent study that found that a large fraction of NDRs at TSSs become filled in by a nucleosome that is marked with H3K4 methylation[57]. The fact that TSSs remain hyperaccessible suggests either that TSSs become occupied by nucleosomes in only a subset of cells in the population, or that nucleosomes rearrange in all cells, creating linkers between adjacent nucleosomes at TSS that are relatively large and thus more accessible than the genome-wide average. Long linkers between the nucleosomes would be expected if flanking nucleosomes reposition across the relatively large interphase nucleosome-free regions around TSSs in the absence of de novo nucleosome assembly.

The level of nucleosome occupancy at the TSS is directly related to the remaining accessibility during mitosis. We find that the set of most accessible TSSs in mitosis has low levels of nucleosome occupancy, though this level is higher than in interphase. This set of most accessible TSSs also shows evidence that the TSS can frequently remain free of nucleosomes, with no changes in the positions of the directly flanking nucleosomes. One parsimonious interpretation of these data is that TSSs are variable in the extent to which nucleosomes reposition during mitosis. This variability can be at the level of single cells for individual TSSs, or at the level of subsets of TSSs where some sets remain bound by factors that maintain an open nucleosome-free site in most cells. TBP has been described as a factor that can maintain stable binding to at least a subset of promoters[69–71]. TBP, possibly together with histone modifications that remain stable in mitosis, may serve as bookmarks for re-activation of promoters in the subsequent cell cycle. Given that we do not find any evidence that CTCF or other factors remain associated with CTCF sites in mitosis, we propose that the continued presence of modified histones and histone variants such as H2A.Z around CTCF motifs, and the larger spacing between adjacent nucleosomes around the site are sufficient for marking these sites for re-binding of CTCF as cells exit mitosis.

Both H2A.Z and H3K4 methylation marks have been studied in regard to their role as mitotic bookmarks at promoters. H2A.Z has been described to form a heterodimer with H3.3 which can fill in the NDRs of promoter regions in mitosis[6,7,55]. The H3.3/H2A.Z heterodimer is found to be less stable than its canonical H3/H2A counterpart and often repositioned or removed by chromatin remodelers. It has been suggested that H3.3/H2A.Z could be a place holder for transcription factors in mitosis. Upon mitotic exit, chromatin remodelers can be recruited to sites of H3.3/H2A.Z and open up NDRs, which will enable transcription factors to bind at its interphase sites

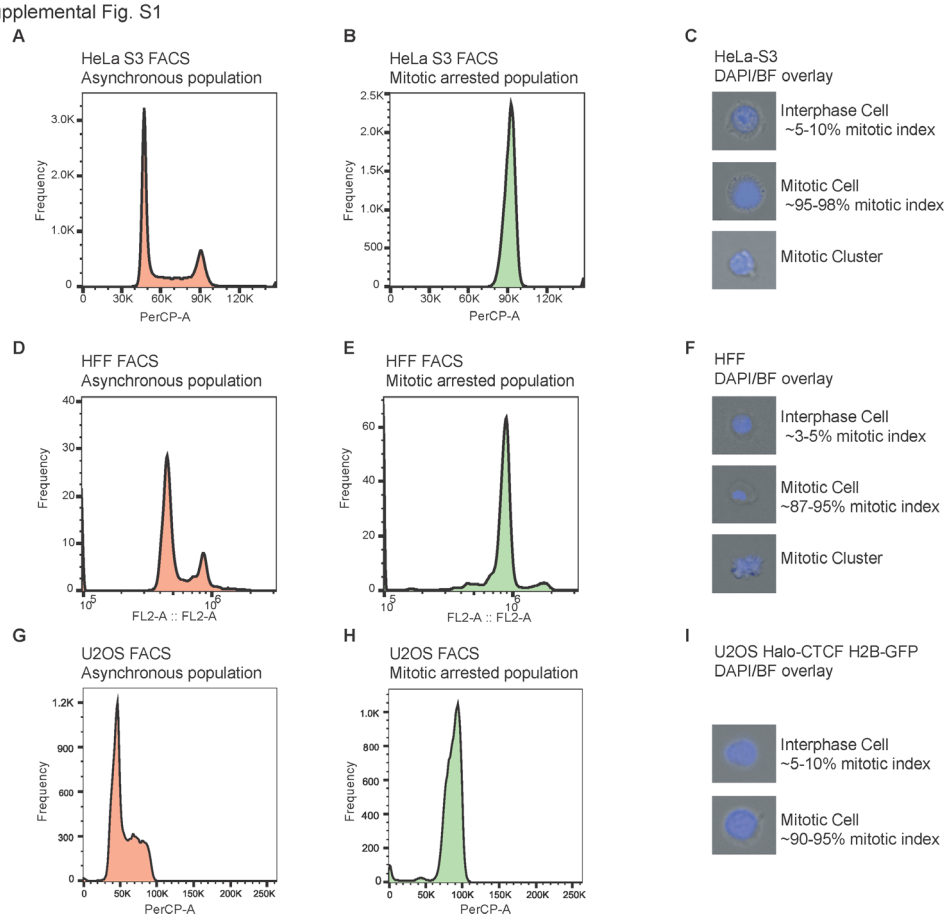
again[7]. Both H3K4 monomethylation and trimethylation are maintained in mitosis ((Lin et al.[58]; Varier et al.[59]) and Fig. 5 and Supplemental Fig. S13). Recently it was observed that a large fraction of NDRs at TSSs become filled in by a nucleosome that is marked with H3K4 methylation and loss of histone acetylation[57]. Therefore, the nucleosomes that fill in NDRs stand out compared to their flanking nucleosomes that will still have acetylation marks. This could be another mechanism for specific recruitment of chromatin remodelers upon mitotic exit to these bookmarked sites, which could enable reestablishment of interphase factor binding.

We find that chromatin organization around CTCF sites alternates between two distinct states during the cell cycle. To convert from one state to the other, a number of molecular events must likely take place. First, as cells enter mitosis, CTCF dissociates from the chromatin. This may involve phosphorylation of CTCF, which reduces its affinity for DNA in vitro[22,23,25,26]. The mechanism by which nucleosomes become repositioned is not known. It is possible that simply the removal of CTCF, and associated factors such as cohesin, is sufficient for nucleosomes to passively reposition and occupy the CTCF motif. Alternatively, and more interestingly, it is possible that specific remodeling enzymes act at CTCF sites during mitosis. Sometime during or after mitotic exit CTCF regains site specific binding. This likely involves dephosphorylation of CTCF. Rebinding at CTCF motifs could be facilitated by the relative large linkers between nucleosomes around previously bound CTCF sites and the presence of histone variants and histone modifications. CTCF rebinding correlates with repositioning of the flanking nucleosomes. How CTCF rebinding and nucleosome repositioning are mechanistically linked remains unknown. CTCF binding could passively lead to nucleosome repositioning or a process of active chromatin remodeling precedes CTCF binding.

The molecular events occurring at CTCF sites during the cell cycle coincide with large scale changes in higher order chromosomal folding, where TADs and loops are present in interphase and absent in prometaphase. The recently proposed model of loop extrusion explains how such localized events of CTCF binding can determine the formation of TADs and loops at the scale of hundreds of kilobases[8–10,60]. This model proposes that during interphase dynamic loop extrusion by cohesin is blocked by CTCF bound sites. This process leads to TAD formation and loops between convergent CTCF sites[11–14]. Taken together our data on CTCF and published data on cohesin[63–66], show that the absence of TADs and CTCF loops in mitosis can be explained by the dissociation of the entire interphase loop extrusion machinery. Upon mitotic exit condensins become inactivated and CTCF and cohesin re-associate rapidly[67], allowing the reestablishment of CTCF loops and TADs.

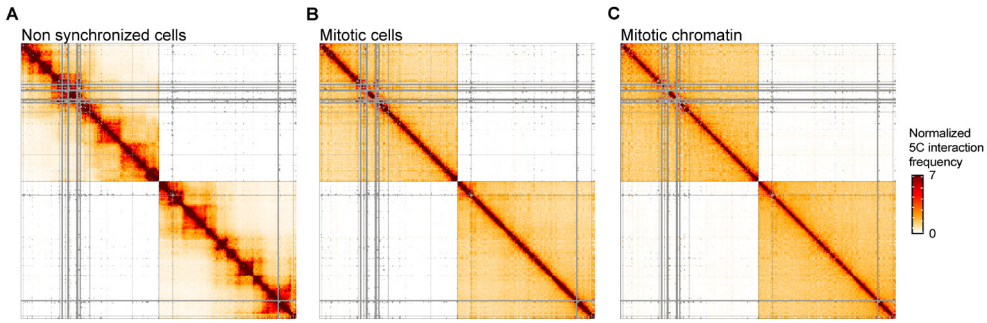
Supplemental figures

Supplemental Fig. S1



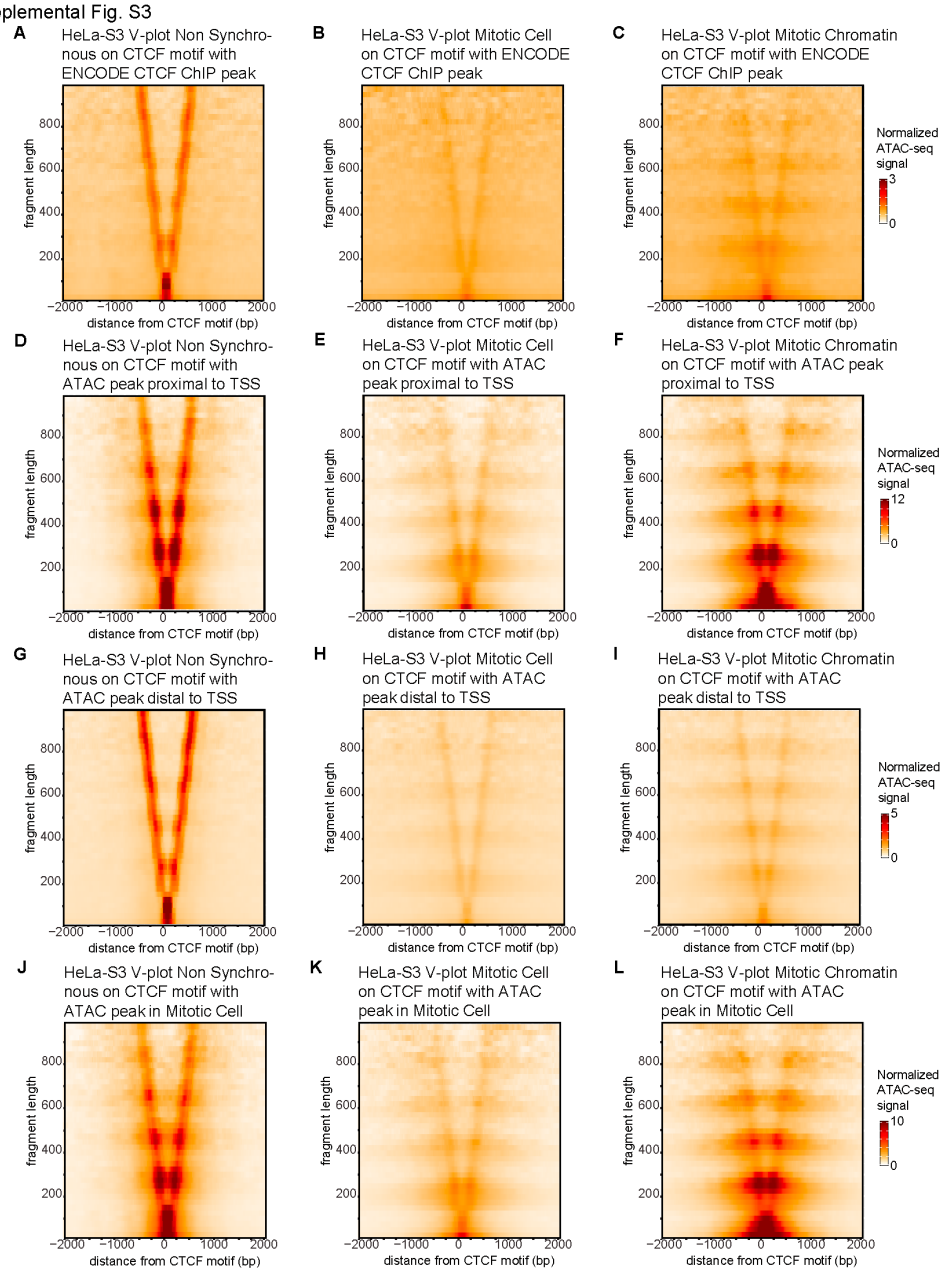
Supplemental Fig. S1 – Flow cytometry using propidium iodide staining and microscopy using DAPI staining show synchronization in M-phase by Nocodazole. (**A-B**) Cell cycle flow cytometry for non-synchronized (**A**) and nocodazole arrested (**B**) HeLa-S3 cell populations. (**C**) DAPI staining of an interphase cell, nocodazole arrested prometaphase cell and purified mitotic chromatin of HeLa-S3 cells and their average scored mitotic index in percentage of total population. (**D-E**) Cell cycle flow cytometry for non-synchronized (D) and nocodazole arrested (E) HFF cell populations. (**F**) DAPI staining of an interphase cell, nocodazole arrested prometaphase cell and purified mitotic chromatin of HFF and their average scored mitotic index in percentage of total population. (**G-H**) Cell cycle flow cytometry for non-synchronized (**G**) and nocodazole arrested (**H**) U2OS Halo-CTCF H2B –GFP cell populations. (**I**) DAPI staining of an interphase cell and nocodazole arrested prometaphase cell of U2OS Halo-CTCF H2B-GFP and their average scored mitotic index in percentage of total population. Note that the scales of x-axes of the different plots are different dependent on the setting of the FACS equipment.

Supplemental Fig. S2



Supplemental Fig. S2 – 5C interaction interaction maps obtained with non-synchronized HeLa-S3 cells (**A**), mitotic HeLa-S3 cells (**B**), and mitotic chromatin purified from mitotic HeLa-S3 cells (**C**). Heatmap shows two 2 MB regions located on chromosome 1 (hg19 chr1: 46740122-48740121 (upper left)) and chromosome 11 (hg19 chr11: 33003550-35003549 (lower right)). Data was binned in 20 kb bins.

Supplemental Fig. S3

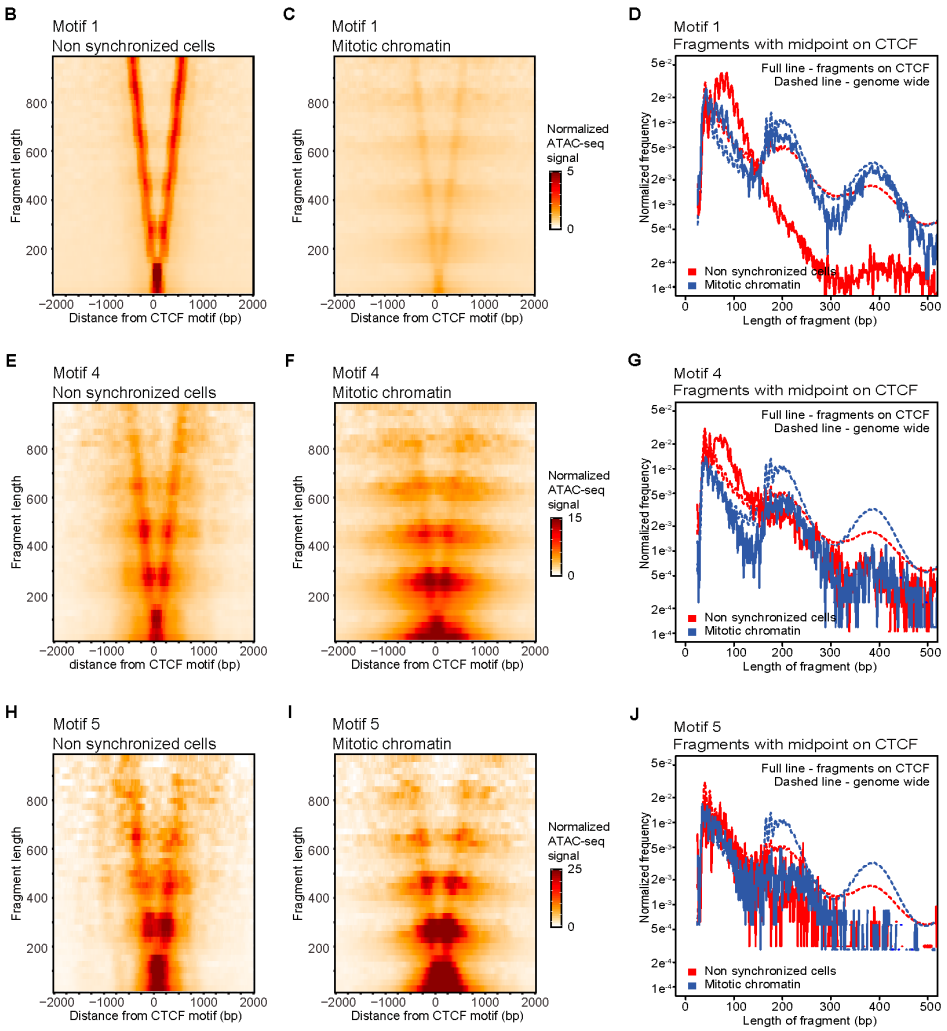


Supplemental Fig. S3 – (**A-C**) Normalized V-plots at CTCF motifs that display ENCODE CTCF ChIP peaks. V-plots are made based on ATAC-seq data from non-synchronized cells (**A**), mitotic cells (**B**) and mitotic chromatin (**C**). (**D-F**) Normalized V-plots at CTCF motifs within 2kb of TSS in non-synchronized cells (**D**), mitotic cells (**E**) and mitotic chromatin (**F**). (**G-I**) Normalized V-plots at CTCF motifs 2 kb or more away from any TSS in non-synchronized cells (**G**), mitotic cells (**H**) and mitotic chromatin (**I**). (**J-L**) Normalized V-plots at CTCF motifs with mitotic ATAC-seq peak in non-synchronized cells (**J**), mitotic cells (**K**) and mitotic chromatin (**L**).

Supplemental Fig. S4

A

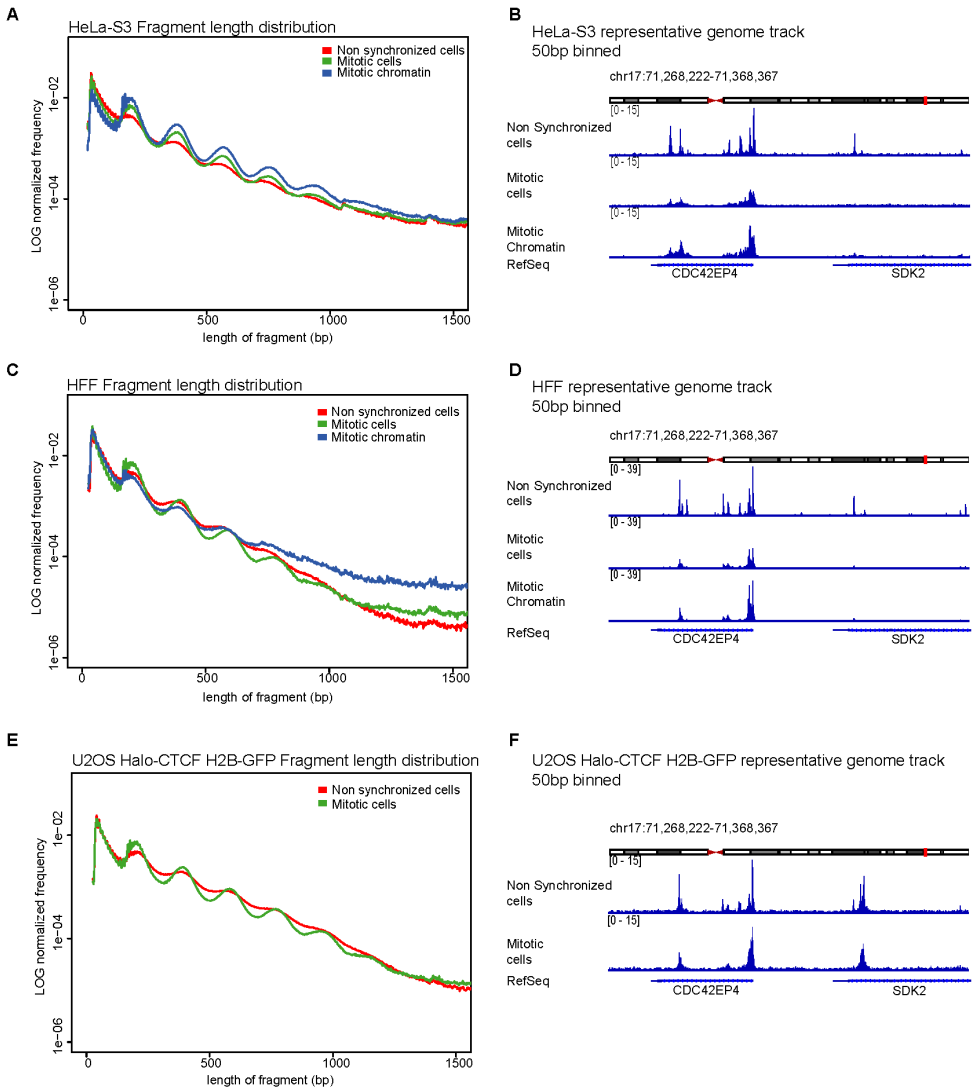
CTCF motifs clustered on signal 75-150bp in mitotic chromatin	Number of motifs
kmeans cluster 1	9775
kmeans cluster 2	59
kmeans cluster 3	21
kmeans cluster 4	612
kmeans cluster 5	104



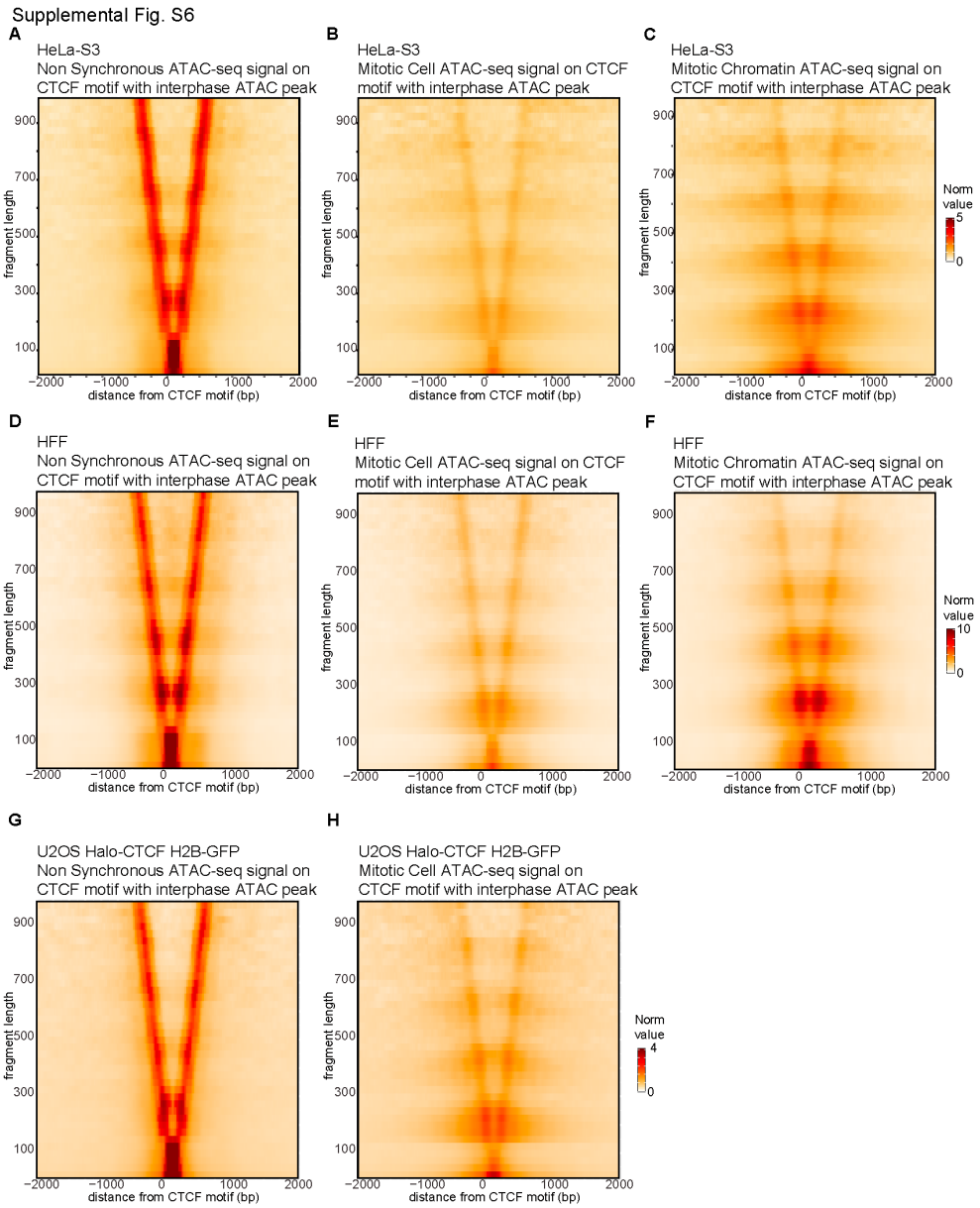
Supplemental Fig. S4 – (**A**) K-means clustering in 5 clusters of ATAC-seq signal of 75-150 bp length reads in mitotic chromatin at 2kb window from interphase accessible CTCF motifs. (**B-C**) V-plots of ATAC-seq data obtained with non-synchronized cells (**B**) and mitotic chromatin (**C**) at CTCF motifs in cluster 1 representing 9,975 CTCF motifs. (**D**) Fragment length distribution of reads with their midpoint on a CTCF motif from cluster 1. (**E-F**) V-plots of ATAC-seq data obtained with non-synchronized cells (**E**) and mitotic chromatin (**F**) on CTCF motifs in cluster 4 of kmeans clustering representing 612 CTCF motifs. (**G**) Fragment length

distribution of reads with their midpoint on a CTCF motif from cluster 4. (**H-I**) V-plots of ATAC-seq data obtained with non-synchronized cells (**H**) and mitotic chromatin (**I**) on CTCF motifs in cluster 5 of kmeans clustering representing 104 CTCF motifs. (**G**) Fragment length distribution of reads with their midpoint on a CTCF motif from cluster 5.

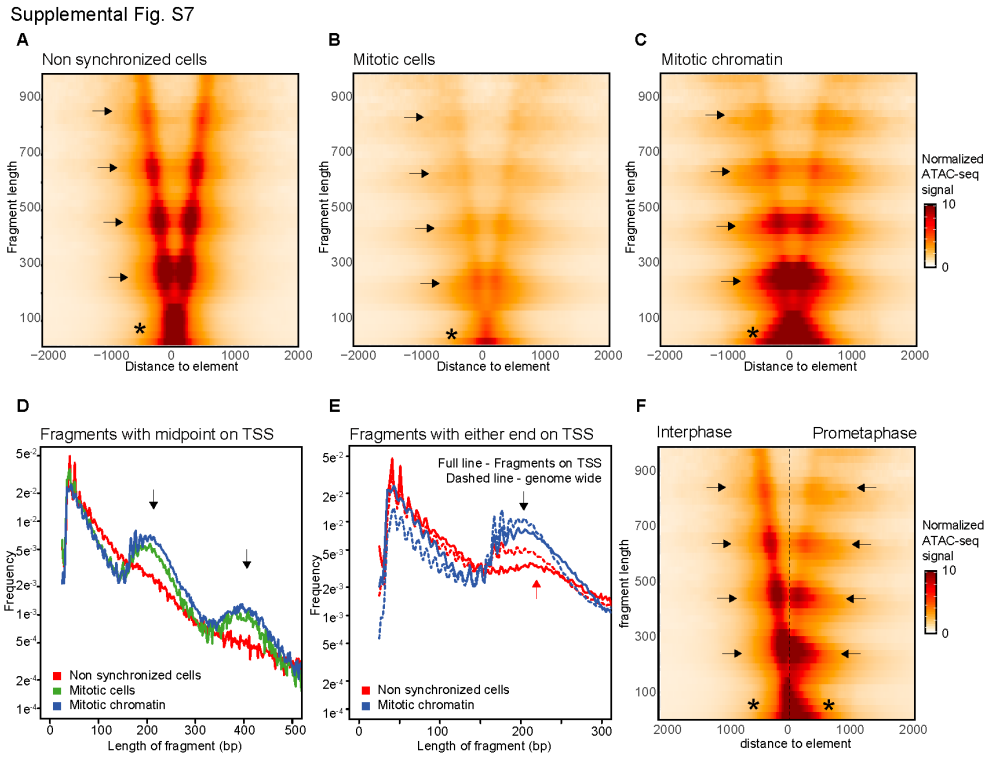
Supplemental Fig. S5



Supplemental Fig. S5 – ATAC-seq fragment length distributions and representative genome tracks for data obtained with HeLa-S3 (**A-B**), HFF (**C-D**) and U2OS Halo-CTCF H2B-GFP (**E-F**).

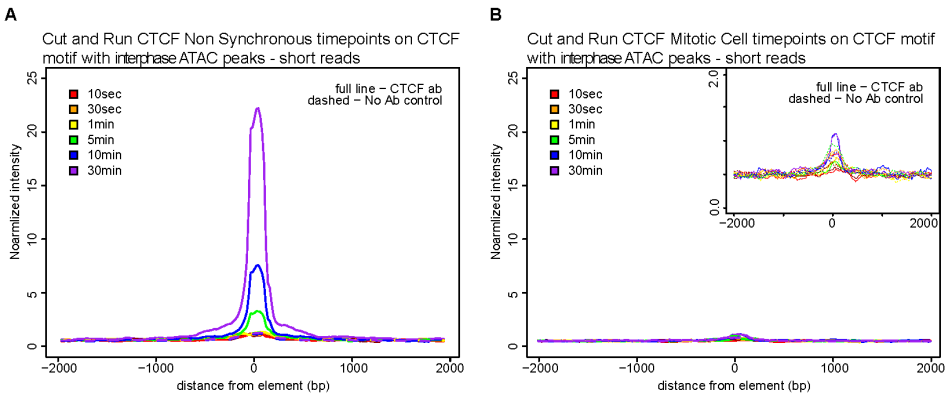


Supplemental Fig. S6 – Normalized V-plots of ATAC-seq data obtained with HeLa-S3 cells at CTCF motifs that display ATAC-seq peaks in interphase for non-synchronized cells (**A**), mitotic cells (**B**) and purified mitotic chromatin (**C**). Normalized V-plots of ATAC-seq data obtained with HFF cells at CTCF sites that display ATAC-seq peaks in interphase for non-synchronized cells (**D**), mitotic cells (**E**) and purified mitotic chromatin (**F**). Normalized V-plots of ATAC-seq data obtained with U2OS Halo-CTCF H2B-GFP cells at CTCF sites that display ATAC-seq peaks in interphase for non-synchronized cells (**A**), mitotic cells (**B**) and purified mitotic chromatin (**C**)



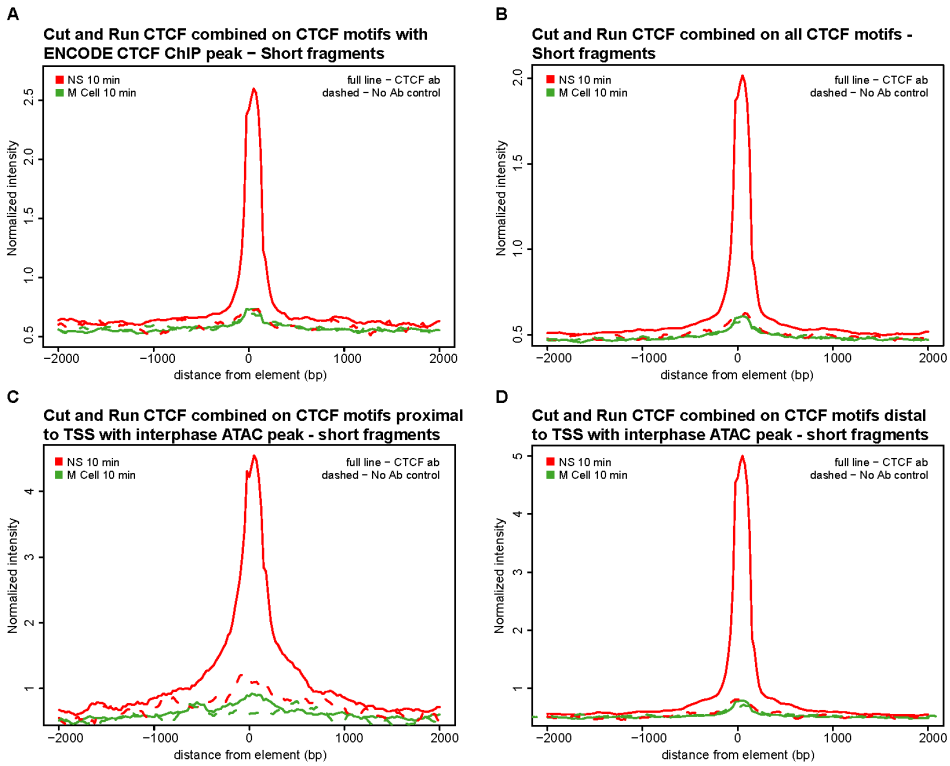
Supplemental Fig. S7 – ATAC-seq data obtained with HeLa-S3 in non-synchronozed cells (**A**), mitotic cells (**B**), and purified mitotic chromatin (**C**) represented as V-plots aggregated on all TSS that display an ATAC-seq peak in interphase. (**D**) Fragment length distribution of non-synchronized cells, mitotic cells and mitotic chromatin for all reads with their midpoint on any TSS. (**E**) Fragment length distribution of non-synchronized cells and mitotic chromatin for all reads with either end near any TSS compared to the genome-wide distribution (dashed line). (**F**) Side-to-side comparison of V-plots for non-synchronized cells and mitotic chromatin up to 2 kb upstream of TSS.

Supplemental Fig. S8



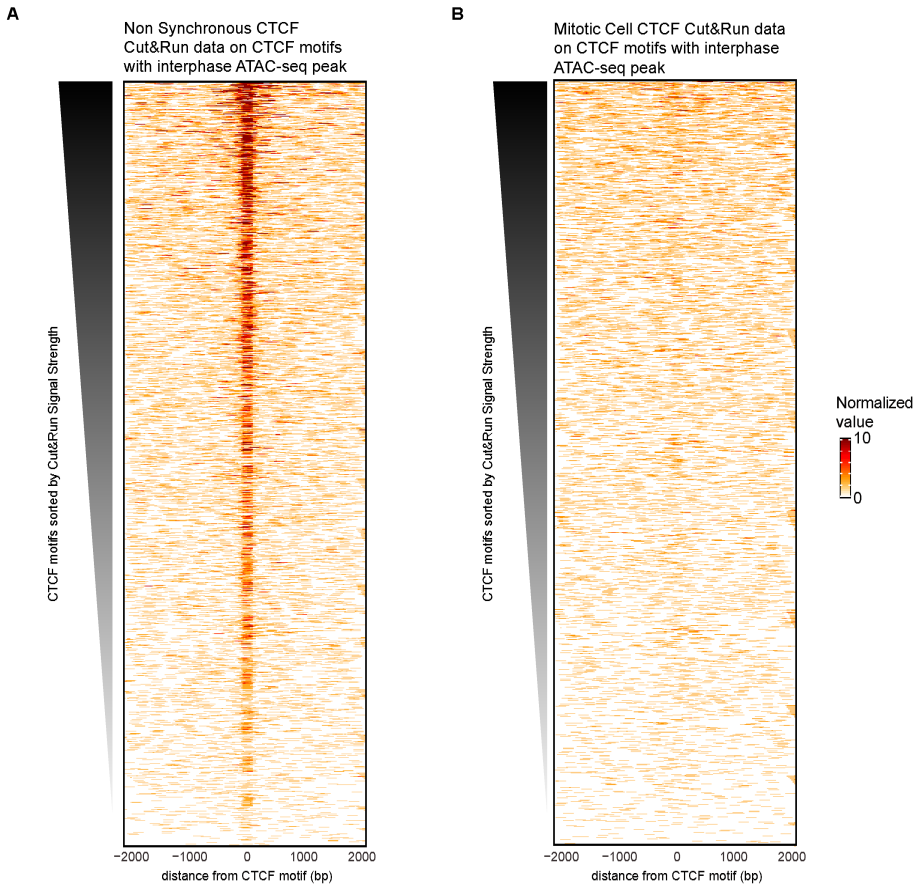
Supplemental Fig. S8 – Aggregation plots of CTCF CUT&RUN signal for reads shorter than 120bp obtained with different digestion times with pA-MNase (10 seconds to 30 min). Data are aggegated at CTCF motifs that display ATAC-seq peaks in interphase for HeLa-S3 non-synchronized cells (**A**) and mitotic cells (**B**). Insert in **B** shows CUT&RUN signal of mitotic cells at a y-axis from (0-2).

Supplemental Fig. S9



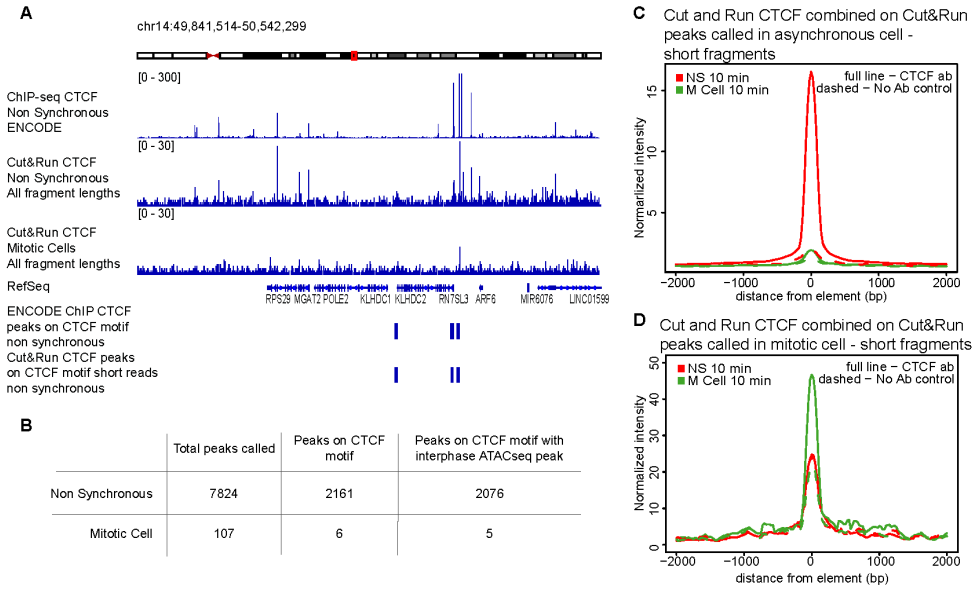
Supplemental Fig. S9 – (**A**) CTCF CUT&RUN signal obtained with non-synchronized and mitotic HeLa-S3 cells for reads shorter than 120bp aggregated at all CTCF motifs that display ENCODE CTCF ChIP peaks (6,893 sites). (**B**) CTCF CUT&RUN signal obtained with non-synchronized and mitotic HeLa-S3 cells for reads shorter than 120bp on all CTCF motifs (42,066 sites). (**C**) CTCF CUT&RUN signal obtained with non-synchronized and mitotic HeLa-S3 cells for reads shorter than 120bp on all CTCF motifs that display an ATAC-seq peak within 2 kb to a TSS in interphase (1,181 peaks). (**D**) CTCF CUT&RUN signal obtained non-synchronized and mitotic HeLa-S3 cells for reads shorter than 120bp on all CTCF motifs that display an interphase ATAC-seq peak further than 2 kb away from a TSS in interphase (9,459 peaks).

Supplemental Fig. S10



Supplemental Fig. S10 – Aggregation of CUT&RUN data heatmap obtained with HeLa-S3 cells for all reads shorter than 120bp on CTCF motifs that display an ATAC-seq peak in interphase sorted on CUT&RUN signal strength in interphase for non-synchronized cell (**A**) and mitotic cells (**B**).

Supplemental Fig. S11

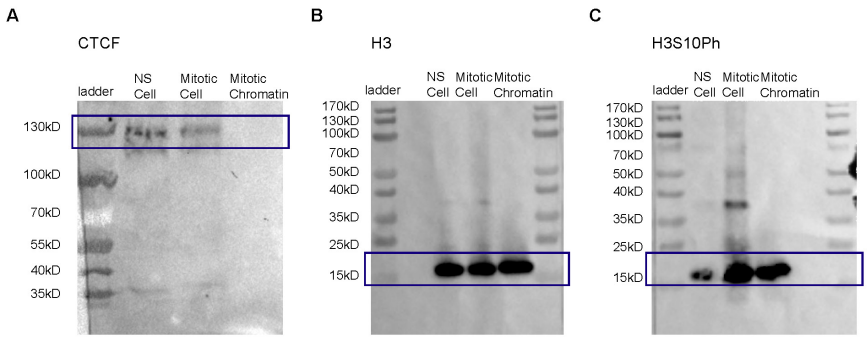


B

	Total peaks called	Peaks on CTCF motif	Peaks on CTCF motif with interphase ATACseq peak
Non Synchronous	7824	2161	2076
Mitotic Cell	107	6	5

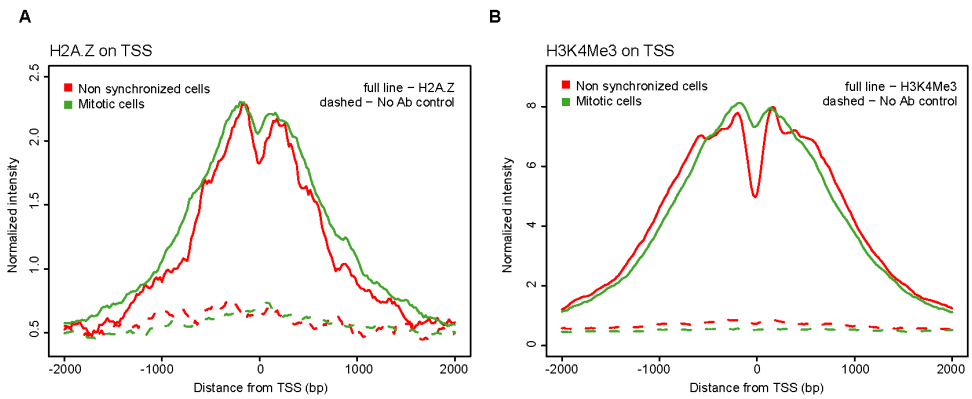
Supplemental Fig. S11 – De novo peak calling on CUT&RUN CTCF data obtained with non-synchronized and mitotic HeLa-S3 cells. (**A**) Visualization of ENCODE non synchronized CTCF ChIP-seq data and CUT&RUN data obtained with non-synchronized and mitotic HeLa-S3 cells for a representative region. The lower two tracks show the locations of CTCF motifs that display peaks called on ChIP-seq and CUT&RUN data respectively (CTCF motifs were stretched to 2kb to enhance visualization). (**B**) Number of peaks called on CTCF CUT&RUN data obtained with non-synchronized and mitotic HeLa-S3 cells and their overlap with CTCF motifs and ATAC-seq peaks. (**C**) CTCF CUT&RUN signal in non-synchronized and mitotic cells for reads shorter than 120bp on all peaks called in non-synchronized cells (7,824 peaks). (**D**) CTCF CUT&RUN signal in non-synchronized and mitotic cells for reads shorter than 120bp on all peaks called in mitotic cells (107 peaks).

Supplemental Fig. S12



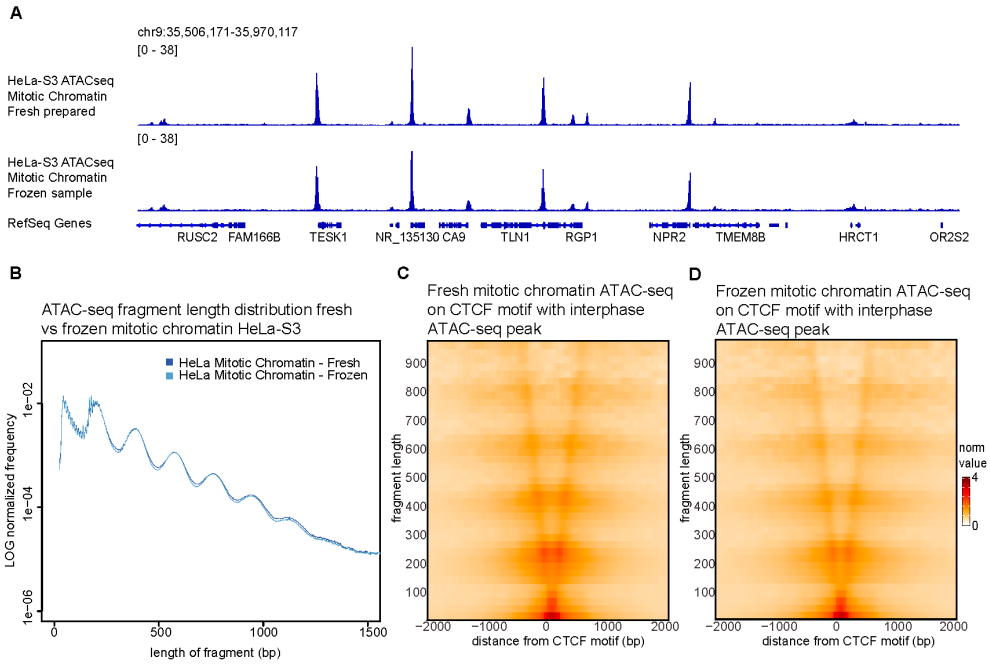
Supplemental Fig. S12 – Western blots of lysates from non-synchronized HeLa-S3 cells, mitotic HeLa-S3 cells and mitotic chromatin purified from mitotic HeLa-S3 cells probed with antibodies detecting CTCF (**A**), H3 (**B**) and H3S10ph (**C**).

Supplemental Fig. S13



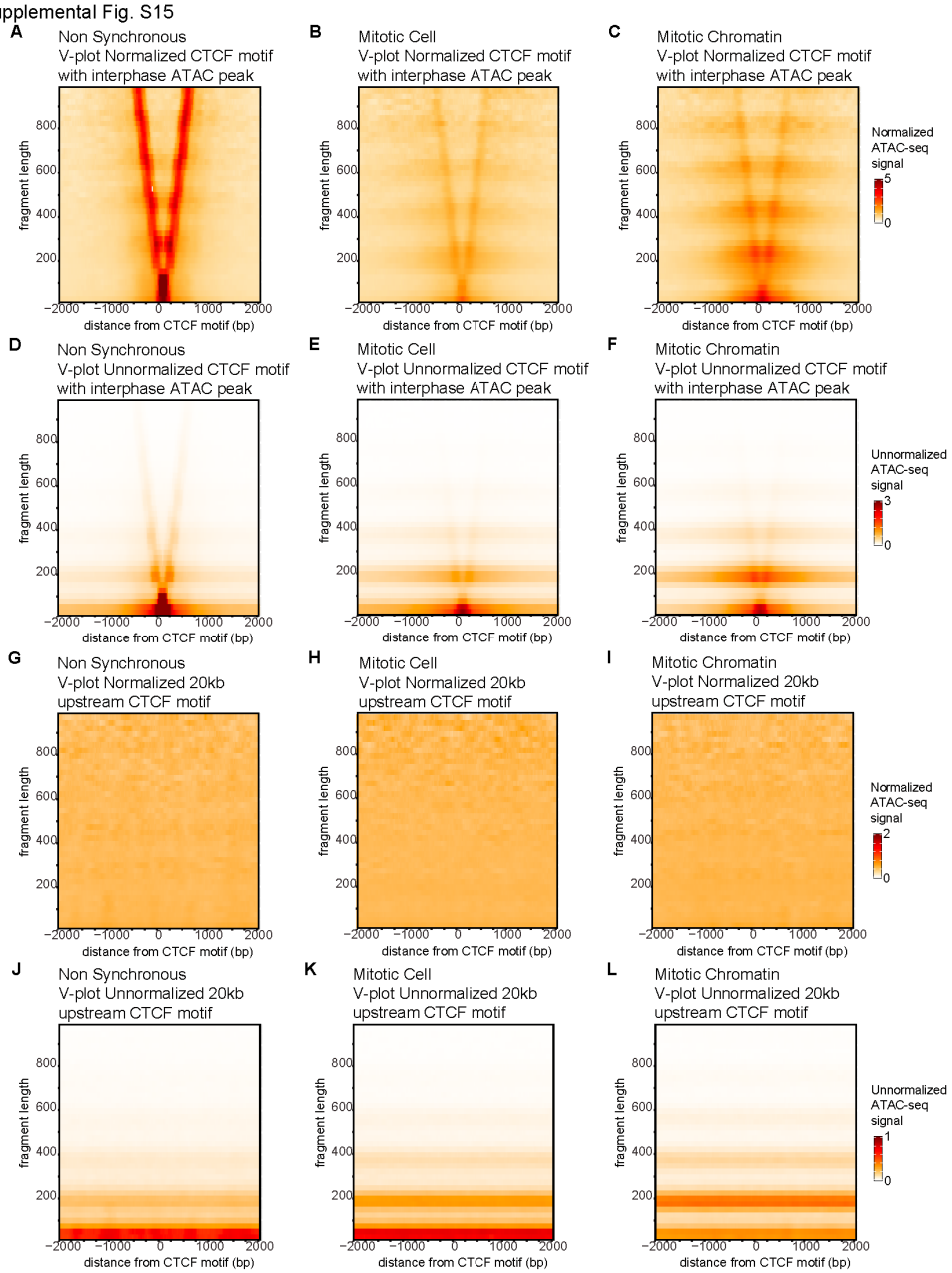
Supplemental Fig. S13 – (**A**) H2A.Z CUT&RUN signal of fragments larger than 120bp on TSSs that display ATAC-seq peaks in interphase for non-synchronized HeLa-S3 cells (red) and mitotic HeLa-S3 cells (green). (**B**) H3K4Me3 CUT&RUN signal of fragments larger than 120bp on TSSs that display ATAC-seq peaks in interphase for non-synchronized HeLa-S3 cells (red) and mitotic HeLa-S3 cells (green).

Supplemental Fig. S14



Supplemental Fig. S14 – Comparison of ATAC-seq data with freshly prepared and frozen stored mitotic chromatin from mitotically synchronized HeLa-S3. (**A**) representative genomic tracks, (**B**) fragment length distribution, (**C-D**) V-plots at CTCF motifs that display ATAC-seq peaks in interphase (**C-D**).

Supplemental Fig. S15



Supplemental Fig. S15 - V-plots normalized (**A-C**) and unnormalized (**D-F**) by the genome-wide fragment length distribution for HeLa-S3 ATAC-seq signal on all CTCF motifs that display an ATAC-seq peak in interphase for non-synchronized cells (**A/D**), mitotic cells (**B/E**) and purified mitotic chromatin (**C/F**). V-plots normalized (**G-I**) and unnormalized (**J-L**) by the genome-wide fragment length distribution for HeLa-S3 ATAC-seq signal on a region 20kb upstream of every CTCF motifs representing a random region in non-synchronized cells (**G/J**), mitotic cells (**H/K**) and purified mitotic chromatin (**I/L**).

Supplemental tables

Supplemental Table T1 – 5C mapping statistics

		Reads Sequenced	Valid reads mapped
R1	Asynchronous	100,986,155	99,609,624
	Mitotic Cell	67,042,262	65,620,645
	Mitotic Cluster	38,212,480	36,918,691
R2	Asynchronous	181,209,251	175,402,702
	Mitotic Cell	181,966,242	175,402,702
	Mitotic Cluster	173,399,970	166,197,088
Combined	Asynchronous	282,195,406	275,012,326
	Mitotic Cell	249,008,504	240,312,381
	Mitotic Cluster	211,612,450	203,115,779

Supplemental Table T2 – Antibodies used for western blot and Cut&Run

	Company + catalogue number	Western blot concentration	Cut&Run concentration
CTCF	CST #2899	1:1000	1:100
H2A.Z	Abcam #ab4174	x	1:100
H3K4Me1	Abcam #ab8895	x	1:50
H3K4Me3	Abcam #ab8580	x	1:100
H3	Abcam #ab1791	1:1000	x
H3S10Ph	Abcam #ab5176	1:1000	x

Methods

Cell culture and synchronization

HeLa S3 and HFF cells were grown at 37°C in Gibco glutamax DMEM supplemented with 10% heat inactivated fetal bovine serum (FBS) and penicillin-streptomycin. We previously generated a U2OS cell line where CTCF has been homozygously and endogenously with HaloTag (FLAG-Halo-hCTCF; C32) [61]. To additionally enable the visualization of mitotic cells, here we stably integrated a previously described transgene expressing histone H2B-GFP with Puro selection[29]. U2OS cells were cultured in low glucose DMEM with 10% FBS and penicillin-streptomycin. Cells were synchronized using thymidine and nocodazole treatments, which was tailored for each cell line. See supplemental methods for detailed description.

Flow cytometry and DAPI staining

Level of mitotic synchrony in all cultures grown for genomic studies was observed by performing flow cytometry for cell cycle analysis using propidium iodide staining in ethanol fixed cells. In addition to this, cells were also stained using DAPI to determine the fraction of prometaphase cells in total culture based on chromatin condensation.

Mitotic chromatin cluster purification

Mitotic chromatin clusters were purified according to previously a published protocol with minor changes[43]. A detailed description can be found in supplemental materials and methods. Mitotic chromatin clusters can be frozen at -80°C in 33.33% glycerol. Chromatin clusters were used for ATAC-seq both directly after chromatin cluster purification and after freezing. Supplemental Fig. S14 shows fresh or frozen clusters obtain highly similar results for ATAC-seq. For this reason, we continued using frozen chromatin clusters for all further ATAC-seq experiments. Mitotic chromatin clusters for 5C were immediately fixed in 1% formaldehyde before pelleting and storage at -80C.

Genomics studies

5C was performed according to published protocols[42]. We investigated the same regions as previously described and we used the same primer pool[44]. This 5C primer pool covers two 2 Mb regions located on Chromosome 1 (hg19 Chr1: 46740122-48740121) and Chromosome 11 (hg19 Chr11: 33003550-35003549). ATAC-seq was performed following a previously published protocol[47]. Additional information can be found in supplemental materials and methods. CUT&RUN was performed according to the published protocol[54]. As mitotic cells do not have a nuclear membrane, it was not possible to use concanavalin A beads. Instead, cells were spun at 600xg for 3

min at 4C for every wash or buffer exchange. Details of all genomics techniques can be found in the supplemental materials and methods.

Live cell imaging studies

FRAP and spaSPT experiments were performed and analyzed as previously described[60]. Time-lapse movies were recorded using phase, GFP (for H2B-GFP) and TMR (for Halo-CTCF) acquiring one frame every 2 or 5 minutes and lasted for a total time of at least 10 hours – enough that several cells in a field of view went through mitosis. CTCF was clearly enriched on mitotic chromosomes during most phases of mitosis. Raw data as well as details on all imaging techniques and analysis of spaSPT using Spot-On can be found in the supplemental materials and methods.

Bioinformatic analyses

All genomics data was mapped to hg19 to enable to usage of ENCODE datasets available. As we remove all blacklisted regions of hg19 as determined by the ENCODE project, we expect that our data mapped to hg38 would show highly similar results[51]. A detailed description of all bioinformatic analyses in this study can be found in the supplemental materials and methods. The Spot-On code is available at Gitlab https://gitlab.com/tjian-darzacq-lab/spot-on-matlab. Code used for analysis of ATAC-seq and CUT&RUN data can be found at Github: https://github.com/dekkerlab/CTCF_in_mitosis_GR_2018. Scripts used for 5C analysis can also be found at Github: https://github.com/dekkerlab/cworld-dekker.

Supplemental material

Extended methods and supplemental movies for this study are available at https://genome.cshlp.org/content/suppl/2019/01/17/gr.241547.118.DC1

Data access

All genomic data generated for this study has been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo) under accession number GSE121840.

All imaging data is publicly available at Zenodo (https://zenodo.org/record/1306976).

Acknowledgments

We especially thank Sarah Hainer and Tom Fazzio for advice on ATAC-seq data interpretation and providing reagents. We also thank Bas van Steensel for advice and suggestions. We thank Sheila Teves and all current and former Dekker lab members for helpful discussions; especially Betul Akgol Oksuz, Bryan Lajoie,

Ankita Nand and Hakan Ozadam for advice on data analysis and Allana Schooley for helpful comments regarding the manuscript. J.D. and XD. acknowledge support from the National Institutes of Health Common Fund 4D Nucleome Program (U54-DK107980 and UO1-EB021236). Additionally, this work was supported by grants from the National Human Genome Research Institute (HG003143) to J.D. and the California Institute of Regenerative Medicine (LA1-08013) to X.D. This work was performed in part at the UC Berkeley CRL Molecular Imaging Center, supported by the Gordon and Betty Moore Foundation. A.S.H. is a postdoctoral fellow of the Siebel Stem Cell Institute. J.D. is an investigator of the Howard Hughes Medical Institute.

Author contributions

M.E.O. and J.D. conceived and designed the project. Y.L and M.E.O. preformed 5C. M.E.O. generated ATAC-seq and CUT&RUN datasets. A.S.H. generated U2OS cell line and designed, performed and analyzed all imaging experiments. M.E.O. analyzed 5C, ATAC-seq and CUT&RUN datasets. M.E.O., A.S.H., X.D. and J.D. wrote the manuscript.

Competing interest statement

All authors declare no conflict of interest.

References

1. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–80 (2012).
2. Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-944.e22 (2017).
3. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–5 (2012).
4. Kim, T. H. *et al.* Analysis of the Vertebrate Insulator Protein CTCF-Binding Sites in the Human Genome. *Cell* **128**, 1231–1245 (2007).
5. Fu, Y., Sinha, M., Peterson, C. L. & Weng, Z. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet.* **4**, (2008).
6. Jin, C. *et al.* H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat. Genet.* **41**, 941–5 (2009).
7. Nekrasov, M. *et al.* Histone H2A.Z inheritance during the cell cycle and its impact on promoter organization and dynamics. *Nat Struct Mol Biol* **19**, 1076–1083 (2012).
8. Dekker, J. & Mirny, L. The 3D Genome as Moderator of Chromosomal Communication. *Cell* **164**, 1110–1121 (2016).
9. Sanborn, A. L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci.* (2015). doi:10.1073/pnas.1518552112
10. Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop

Extrusion. *Cell Rep.* 1–12 (2016). doi:10.1101/024620
11. Vietri Rudan, M. *et al.* Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. *Cell Rep.* **10**, 1297–1309 (2015).
12. Wit, E. De *et al.* CTCF Binding Polarity Determines Chromatin Looping. *Mol. Cell* **60**, 1–9 (2015).
13. Rao, S. S. P. *et al.* A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* **159**, 1665–1680 (2014).
14. Guo, Y. *et al.* CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* **162**, 900–910 (2015).
15. Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342**, 948–53 (2013).
16. Marsden, M. P. F. & Laemmli, U. K. Metaphase chromosome structure: Evidence for a radial loop model. *Cell* **17**, 849–858 (1979).
17. Earnshaw, W. C. & Laemmli, U. K. Architecture of metaphase chromosomes and chromosome scaffolds. *J. Cell Biol.* **96**, 84–93 (1983).
18. Adolph, K. W. Organization of chromosomes in mitotic HeLa cells. *Exp. Cell Res.* **125**, 95–103 (1980).
19. Gibcus, J. H. *et al.* A pathway for mitotic chromosome formation. *Science (80-.).* **359**, eaao6135 (2018).
20. Goloborodko, A., Imakaev, M. V., Marko, J. F. & Mirny, L. Compaction and segregation of sister chromatids via active loop extrusion. *Elife* **5**, 1–20 (2016).
21. Dowen, J. M. *et al.* Multiple structural maintenance of chromosome complexes at transcriptional regulatory elements. *Stem Cell Reports* **1**, 371–378 (2013).
22. Dephoure, N. *et al.* A quantitative atlas of mitotic phosphorylation. *Proc. Natl. Acad. Sci.* **105**, 10762–10767 (2008).
23. Dovat, S. *et al.* A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. *Genes Dev.* **16**, 2985–2990 (2002).
24. Rizkallah, R. & Hurt, M. M. Regulation of the Transcription Factor YY1 in Mitosis through Phosphorylation of Its DNA-binding Domain. *Mol. Biol. Cell* **20**, 4766–4776 (2009).
25. Sekiya, T., Murano, K., Kato, K., Kawaguchi, A. & Nagata, K. Mitotic phosphorylation of CCCTC-binding factor (CTCF) reduces its DNA binding activity. *FEBS Open Bio* **7**, 397–404 (2017).
26. Jantz, D. & Berg, J. M. Reduction in DNA-binding affinity of Cys2His2 zinc finger proteins by linker phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 7589–7593 (2004).
27. Hsiung, C. C. *et al.* Genome accessibility is widely preserved and locally modulated during mitosis. 1–29 (2015). doi:10.1101/gr.180646.114
28. Martinez-Balbas, M. A., Dey, A., Rabindran, S. K., Ozato, K. & Wu, C. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* **83**, 29–38 (1995).
29. Teves, S. S. *et al.* A Dynamic Mode of Mitotic Bookmarking by Transcription Factors. *Elife* 066464 (2016). doi:10.1101/066464
30. Burke, L. J. *et al.* CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin. *EMBO J.* **24**, 3291–300 (2005).
31. Liu, Y. *et al.* Widespread Mitotic Bookmarking by Histone Marks and Transcription Factors in Pluripotent Stem Cells. *Cell Rep.* **19**, 1283–1293 (2017).
32. Chen, D. *et al.* Condensed mitotic chromatin is accessible to transcription factors and chromatin structural protiens. *J. Cell Biol.* **168**, 41–54 (2005).
33. Pallier, C. *et al.* Association of chromatin proteins high mobility group box (HMGB) 1 and HMGB2 with mitotic chromosomes. *Mol. Biol. Cell* **14**, 3414–

26 (2003).
34. Festuccia, N. *et al.* Transcription factor activity and nucleosome organisation in mitosis. *BioRxiv* (2018). doi:10.1101/392241
35. Hochegger, H. *et al.* An essential role for Cdk1 in S phase control is revealed via chemical genetics in vertebrate cells. *J. Cell Biol.* **178**, 257–68 (2007).
36. Taylor, W. R. FACS-Based Detection of Phosphorylated Histone H3 for the Quantitation of Mitotic Cells. in *Checkpoint Controls and Cancer: Volume 2: Activation and Regulation Protocols* (ed. Schönthal, A. H.) 293–299 (Humana Press, 2004). doi:10.1385/1-59259-811-0:293
37. Jackman, J. & O'Connor, P. M. Methods for Synchronizing Cells at Specific Stages of the Cell Cycle. *Curr. Protoc. Cell Biol.* 1–20 (2001). doi:10.1002/0471143030.cb0803s00
38. Raccaud, M. & Suter, D. M. Transcription factor retention on mitotic chromosomes: regulatory mechanisms and impact on cell fate decisions. *FEBS Lett.* **592**, 878–887 (2018).
39. Raccaud, M. *et al.* Mitotic chromosome binding predicts transcription factor properties in interphase. *BioRxiv* (2018).
40. Nagano, T. *et al.* Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017).
41. Oomen, M. E. & Dekker, J. Epigenetic characteristics of the mitotic chromosome in 1D and 3D. *Crit. Rev. Biochem. Mol. Biol.* **0**, 1–20 (2017).
42. Dostie, J. *et al.* Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**, 1299–309 (2006).
43. Gasser, S. M. & Laemmli, U. K. Improved methods for the isolation of individual and clustered mitotic chromosomes. *Exp. Cell Res.* **173**, 85–98 (1987).
44. Hnisz, D. *et al.* Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
45. Crane, E. *et al.* Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* (2015). doi:10.1038/nature14450
46. Tang, Z. *et al.* CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. *Cell* **163**, 1–17 (2015).
47. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **2015**, 21.29.1-21.29.9 (2015).
48. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–8 (2013).
49. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).
50. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* **12**, 2478–2492 (2017).
51. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
52. Hsiung, C. C.-S. *et al.* Genome accessibility is widely preserved and locally modulated during mitosis - supplementary data. *Genome Res.* **25**, 213–25 (2015).
53. Zentner, G. E. & Henikoff, S. Surveying the epigenomic landscape, one base at a time. *Genome Biol.* **13**, 250 (2012).
54. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* 1–35 (2017). doi:10.1101/097188

55. Kelly, T. K. *et al.* H2A.Z maintenance during mitosis reveals nucleosome shifting on mitotically silenced genes. *Mol. Cell* **39**, 901–11 (2010).
56. Wang, F. & Higgins, J. M. G. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol.* **23**, 175–84 (2013).
57. Javasky, E. *et al.* Study of mitotic chromatin supports a model of bookmarking by histone modifications and reveals nucleosome deposition patterns. *Genome Res.* **28**, 1455–1466 (2018).
58. Lin, S., Yuan, Z.-F., Han, Y., Marchione, D. M. & Garcia, B. A. Preferential phosphorylation on old histones during early mitosis in human cells. *J. Biol. Chem.* **291**, jbc.M116.726067 (2016).
59. Varier, R. a *et al.* A phospho/methyl switch at histone H3 regulates TFIID association with mitotic chromosomes. *EMBO J.* **29**, 3967–78 (2010).
60. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife* **6**, (2017).
61. Hansen, A. S., Cattoglio, C., Darzacq, X. & Tjian, R. Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus* **1034**, 00–00 (2017).
62. Hansen, A. S. *et al.* Robust model-based analysis of single-particle tracking experiments with Spot-On. *Elife* **7**, 1–33 (2018).
63. Liang, Z. *et al.* Chromosomes Progress to Metaphase in Multiple Discrete Steps via Global Compaction/Expansion Cycles. *Cell* **161**, 1124–1137 (2015).
64. Waizenegger, I. C., Hauf, S., Meinke, a & Peters, J. M. Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase. *Cell* **103**, 399–410 (2000).
65. Losada, A., Hirano, M. & Hirano, T. Cohesin release is required for sister chromatid resolution, but not for condensin-mediated compaction, at the onset of mitosis. *Genes Dev.* **16**, 3004–3016 (2002).
66. Nagasaka, K., Hossain, M. J., Roberti, M. J., Ellenberg, J. & Hirota, T. Sister chromatid resolution is an intrinsic part of chromosome organization in prophase. *Nat. Cell Biol.* (2016). doi:10.1038/ncb3353
67. Cai, Y. *et al.* An experimental and computational framework to build a dynamic protein atlas of human cell division. *bioRxiv* (2018).
68. Ohta, S. *et al.* The Protein Composition of Mitotic Chromosomes Determined Using Multiclassifier Combinatorial Proteomics. *Cell* **142**, 810–821 (2010).
69. Chen, D., Hinkley, C. S., Henry, R. W. & Huang, S. TBP dynamics in living human cells: constitutive association of TBP with mitotic chromosomes. *Mol. Biol. Cell* **13**, 276–84 (2002).
70. Xing, H., Vanderford, N. L. & Sarge, K. D. The TBP-PP2A mitotic complex bookmarks genes by preventing condensin action. *Nat. Cell Biol.* **10**, 1318–1323 (2008).
71. Teves, S. S. *et al.* A stable mode of bookmarking by TBP recruits RNA Polymerase II to mitotic chromosomes. *Elife* **7**, (2018).

Chapter 4

Differences in mitotic chromosome organization between cell types and species

Marlies E. Oomen[1], Nick Owens[2], Thaleia Papadopoulou[2], Pablo Navarro[2], Job Dekker[1,3]

Unpublished results

1. Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA
2. Epigenetics of Stem Cells, Department of Developmental and Stem Cell Biology, Institut Pasteur, CNRS UMR3738, 75015 Paris, France
3. Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, MD 208156789, USA

Abstract

Mitotic chromosomes are often perceived as universal structures across species and cell types. However, contradicting recent genomics studies suggest that some features of mitotic chromosomes might be cell type or species specific. We have previously reported that CTCF binding in human differentiated cell lines is lost in mitosis. In contrast, studies have shown that in mouse stem cells and progenitor cells CTCF remain bound to a subset of sites during mitosis. We confirm these findings by performing CTCF footprint analysis of ATAC-seq data in mitotic mESCs. Additionally, we investigate the role of mitotically bound CTCF in chromosome organization by Hi-C. We do not find any remaining interphase structures such as TADs or CTCF loops at mitotic bookmarked CTCF sites in mESCs. This suggests that mitotic loop extruders condensin I and II are not blocked by bound CTCF, and thus that any remaining CTCF binding does not alter mitotic chromosome folding. Lastly, we compare mitotic Hi-C data generated in this study in mouse with publicly available data from human and chicken cell lines. We do not find any cell type specific differences; however, we do find a difference between species. The average size of the mitotic loops is much smaller in chicken (350 kb), compared to human (750 kb) and mouse (2 mb). Interestingly, we find that this difference in loop size is correlated with the average length of the q-arm in these species. This suggests that the dimensions of mitotic chromosomes can be modulated by loop extruding factors condensin I and II to facilitate proper condensation of the chromosome arms.

Introduction

The development of 3C-techniques[1–4] has brought a better understanding of the key features of chromosome organization in vertebrate cells. Interphase chromosomes are organized on megabase scale in A and B compartments and on a smaller scale of tens to hundreds of kilobase in topologically associated domains (TADs)[3,5–10]. TADs are proposed to be formed by loop extruding machines, such as cohesins, which can be blocked by the chromatin binding protein CCCTC-binding factor (CTCF) when bound to its motif[8,9,11–16]. Although the mechanisms that establish and maintain these structures are largely shared between different cell types and between different vertebrate species, the specific genomic regions that interact can differ strongly between species, cell types and even between sick and healthy cells[8,14,17–20].

In contrast to interphase chromatin, vertebrate mitotic chromosomes are often perceived as universal structures, independent of cell type or organism. Historically studied by microscopy[21–23] and in more recent years by genomics[24–26], we gained understanding on the fundamental principles of mitotic chromosome folding. In mitosis the interphase structures are completely dissolved, as both TADs and compartments can no longer be observed[24,25]. Instead, chromosomes are folded as helical loop arrays mediated by condensin I and II, which are not confined to any specific locations[24,27,28]. When studying mitotic chromosomes by Hi-C, a general distance decay without any site-specific features in heatmaps can be observed[24,25].

This might give the impression that all mitotic chromosomes are organized in a similar fashion. However, imaging and biochemical studies revealed that condensins play a more complex role during the rapid cell cycle of stem cells[29]. It has been shown in *Xenopus leavis* that mitotic chromosomes from sperm nuclei are folded as long and thin structures but become increasingly shorter and fatter throughout the early stages of development[30]. Additionally, depletion experiments in xenopus extracts show that the ratio of condensin I and II can affect the width-to-length ratio of chromosomes in mitosis[31]. Using genomics techniques, it was found that mitotic chromosomes can be cell type-specific on a more detailed scale as well, in chromatin accessibility, histone modifications and mitotically bound chromatin factors[32–35]. Along these lines, studies to determine whether architectural protein CTCF remains bound in mitosis show different result in different cell lines. In differentiated cell lines HeLa and HFF, we have previously reported complete loss of CTCF binding by both ATAC-seq and Cut&Run, as well as by imaging[32]. However, in mouse embryonic stem cells (mESCs), mouse blood progenitor cell line G1E-ER4 and mouse pluripotent cells (mPCs), a substantial fraction of CTCF sites remains bound in mitosis and this mitotic retention has been suggested to facilitate

more rapid transcription initiation of a specific set of genes as cells exit mitosis and reestablish interphase chromosome organization[36–39].

In this study, we explore differences in mitotic chromosome organization between cell types and species. As was reported previously, a subset of CTCF motifs remain occupied in mitosis in mESCs[39]. Using published CTCF ChIP-seq data[37], we first categorize CTCF binding sites in sets of sites that maintain CTCF binding in mitosis in mESC cells, sites that show reduced CTCF binding and sites that lose CTCF binding in mitosis. Using these categories, we confirm that CTCF can remain bound at certain sites in mitosis by reanalyzing previously published mESCs ATAC-seq data. Next, we set out to observe chromosome organization at mitotically bound CTCF sites and CTCF sites that lose CTCF binding in mitosis in both mESCs and differentiated mouse cell line C2C12 using Hi-C. Lastly, we explore species specific features of mitotic loop arrays of chromosomes in mouse, chicken and human Hi-C data.

Results

A fraction of CTCF sites remains bound to their cognate sites in mitotic stem cells

In the past few years several genomics studies have reported contradictory results on the cell cycle binding dynamics of CTCF[32,36–39]. These studies did not only differ in cell type, but also in which technique was used to study mitotic CTCF binding. Using ATAC-seq and Cut&Run, it was shown that in human differentiated cell lines all CTCF sites lose binding in mitosis[32]. In contrast, using ChIP-seq it was shown in mESC, mPCs and the mouse blood progenitor cell line G1E-ER4 that a substantial fraction of CTCF motifs remain occupied in mitosis[36–38]. It is possible that these differences are the result of the use of different methods. ChIP-seq requires pull down of the protein of interest, in this case CTCF, which can result in the overestimation of bound protein. ATAC-seq and Cut&Run are both performed on unfixed cells and do not rely on pull down. However, these studies do not only differ in genomics techniques and crosslinking conditions, but more notably, they differ in which cell line was used. Perhaps more interestingly, we hypothesized that this difference could result from a difference in cell types. This would suggest that pluripotent cells can maintain partial CTCF binding in mitosis, whereas somatic cell lines lose CTCF binding in mitosis entirely. In order to compare the different cell types, while using the same technique, we set out to compare and reanalyze ATACseq data of mitotic mouse stem cells generated in the study by Festuccia et al[37] using tools developed in our own study[32] (see also chapter 3 of this thesis).

In the study by Festuccia et al[37], ChIP-seq was performed to observe the binding of CTCF in non-synchronized and mitotically-synchronized mESC cells.

A non-synchronized population of cells predominantly contains cells in interphase and only a small fraction of mitotic cells. We will therefore interpret data from non-synchronized cells to represent features of interphase cells. Using ChIP-seq, a total of 51,805 CTCF motifs were found to be bound by CTCF in interphase. We further categorized these interphase-bound CTCF sites as bookmarked sites, where ChIP-seq signal is maintained in mitosis (10,799 sites), reduced sites, where mitotic ChIP-seq signal is lower yet still significant (18,704 sites) and lost CTCF sites, where ChIP-seq signal in mitosis is below the level required to call peaks (22,302 sites). We did not find CTCF sites that gained binding in mitosis but were not bound in interphase cells.

As described above, performing ChIP-seq in mitotic cells can result in over-estimation of signal. We therefore used the ATAC-seq data generated in the same study, to observe footprints at these CTCF sites categories and confirm the findings that were obtained with ChIP-seq. ATAC-seq signal can be represented as V-plots[32,40], which not only unveils chromatin accessibility, but also the footprint of any bound protein. When CTCF is bound to chromatin, it will occupy approximately 80 base pairs around its motif. Furthermore, it will push the neighboring nucleosomes away from the motif and into a well-positioned tight array on each side of the motif[32,41]. We can observe these phenomena when we represent ATAC-seq data of non-synchronized mESCs aggregated around CTCF sites that are known to be bound in interphase based on ChIP-seq data (figure 1a, see also chapter 3, figure 3a). First, the arms of the V cross at approximately 80bp fragment length, the known footprint size of CTCF[41]. Second, along the arms of the V, dots of enriched signal appear at regular interval (~280bp, ~460bp, ~640bp etc). This ATAC-seq signal indicates the array of well-positioned nucleosomes flanking the bound CTCF motif[41]. Previously, we found that in differentiated cell lines HeLa, U2OS and HFF, CTCF sites generally lost accessibility in mitosis[32]. When ATAC-seq signal was plotted as V-plots, we found that CTCF sites no longer showed enrichment at 80bp fragment length in mitosis (chapter 3, figure 3b). Instead, the fragment size dropped to much smaller fragment size. Furthermore, nucleosomes were no longer positioned in an array, but instead can rearrange their positioning and are no longer confined by distance from the motif.

When we create V-plots for all interphase-bound CTCF sites in both non-synchronized (figure 1a) and mitotic (figure 1e) mESCs, we see a less clear picture. First, more accessibility is maintained at CTCF sites in mitotic mESCs compared to differentiated cell lines (compare figure 1 to chapter 3, figure 3). When we look at a side-by-side comparison of V-plots of non-synchronized and mitotic cells at the CTCF motif (figure 1i), we do see that the size of the CTCF footprint and the positioning

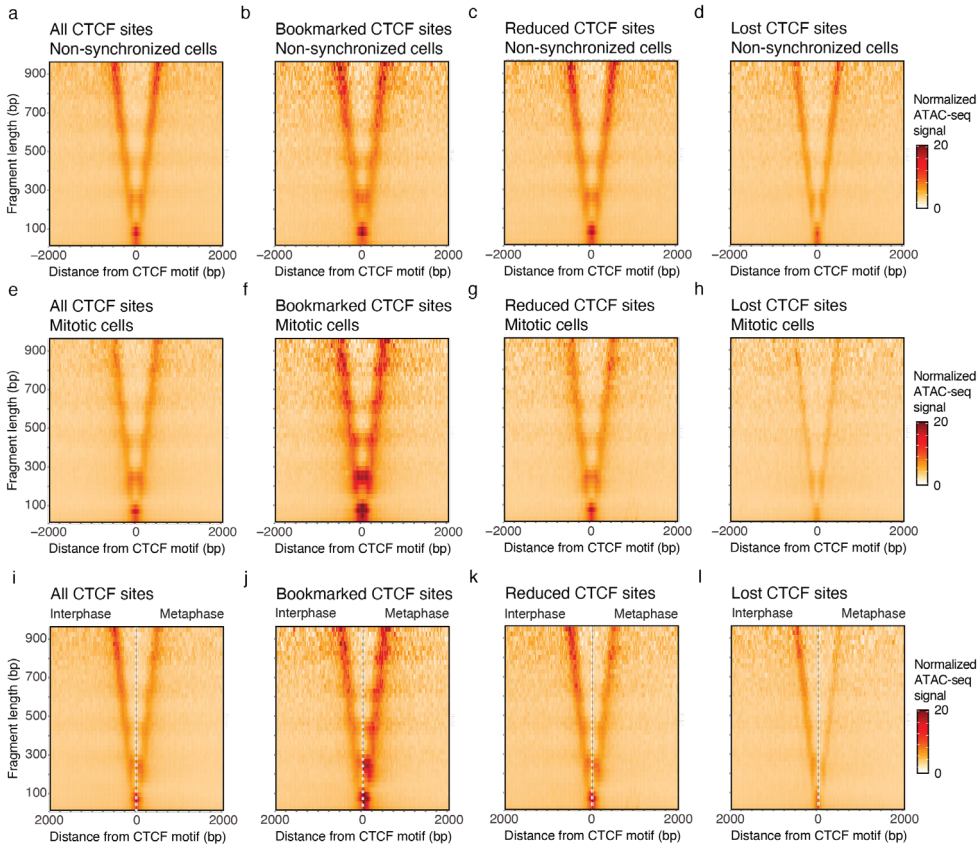


Figure 1 – ATAC-seq data in mESCs show that a group of CTCF motifs remain bound by CTCF in mitosis, whereas other CTCF motifs lose binding. (**a-d**) ATAC-seq data of non-synchronized mESCs represented in V-plots as a pile up on all interphase-bound CTCF sites (51,805 sites total) (a), bookmarked CTCF sites (10,799 sites) (b), CTCF sites with reduced CTCF binding (18,704 sites) (c) and CTCF sites that lose CTCF binding in mitosis (22,302 sites) (d). (**e-h**) ATAC-seq data of mESCs synchronized in mitosis represented in V-plots as a pile up on all interphase-bound CTCF sites (e), bookmarked CTCF sites (f), CTCF sites with reduced CTCF binding (g) and CTCF sites that lose binding in mitosis (h). (**i-l**) Side-by-side comparison of V-plots for non-synchronized and mitotically synchronized cells on all interphase-bound CTCF sites (i), bookmarked CTCF sites (j), reduced CTCF sites (k) and CTCF sites that lose binding in mitosis (l).

of the nucleosomes along the arms of the V drop down to shorter fragment sizes in mitosis. However, this change is less drastic than what we have observed before in differentiated cell lines. This suggests that there are CTCF sites that maintain mitotic binding as well as CTCF sites that lose binding during mitosis, as was observed by using ChIP-seq before[37,39]. Indeed, we find that mitotically bookmarked sites (figure 1b, f, j) maintain both ATAC-seq signal and strong CTCF footprint in mitosis. In

contrast, lost CTCF sites lose ATAC-seq signal and the fragment size of the CTCF footprint drops to shorter fragments, suggesting loss of CTCF binding (figure 1d, h, l). ATAC-seq signal at CTCF sites that showed reduced ChIP-seq signal in mitotic mESCs, show a more ambiguous footprint when plotted as V-plots (figure 1c, g, k). This suggest that this category contains both CTCF sites that lose binding in mitosis and other CTCF sites that maintain binding. Alternatively, this can also be explained by the fact that this ATAC-seq experiment is performed on a population of cells, and this can be the result of cell-to-cell variability in CTCF binding to this set of sites. For example, the same CTCF motif remains bound in mitosis in one cell, whereas it lost binding in mitosis in another cell. We also note that the ATAC-seq signal in non-synchronized cells is much less intense for lost CTCF sites compared to bookmarked CTCF sites. This could again be explained by cell-to-cell variability, where the CTCF sites in the lost CTCF site category is not bound in every cell in the population at the time of the ATAC-seq experiment.

Taken together this analysis of ATAC-seq data confirms that at 30-50% of all interphase-bound CTCF sites, CTCF maintains binding to the motif during mitosis in mouse stem cells. This is in contrast to previous observations by both genomics and imaging techniques in human differentiated cell lines, where CTCF binding to CTCF motifs is lost entirely[32].

Hi-C data shows loss of interphase organization in mitosis for both mouse stem cells and differentiated cells

The finding that a substantial fraction of CTCF sites maintains binding to mitotic chromosomes, raises the question whether CTCF can still function as an architectural protein in mitosis. In interphase cells, chromatin-bound CTCF can block loop extrusion mediated by cohesin[11]. This results in the formation of TADs and strong interactions between 2 CTCF sites (CTCF-CTCF loops), which can both be observed by Hi-C and other C-techniques[7,42,43]. In mitotic differentiated cell lines, where CTCF binding is lost, no TADs and CTCF-CTCF or any other site-specific loops are observed[24,25,32]. Maintained CTCF binding in mitotic stem cells creates the opportunity to study whether mitotic loop extruding machines condensin I and II, can also be blocked by CTCF, or whether they can shape the characteristic helical loop array uninterrupted by bound CTCF. We performed Hi-C on non-synchronized and mitotically synchronized mouse embryonic stem cells (figure 2a). In addition to this, we also performed Hi-C on the differentiated myoblast mouse cell line C2C12 (figure 2b) - a cell line derived from muscle tissue - following our hypothesis that differentiated cell lines lose binding of CTCF in mitosis entirely.

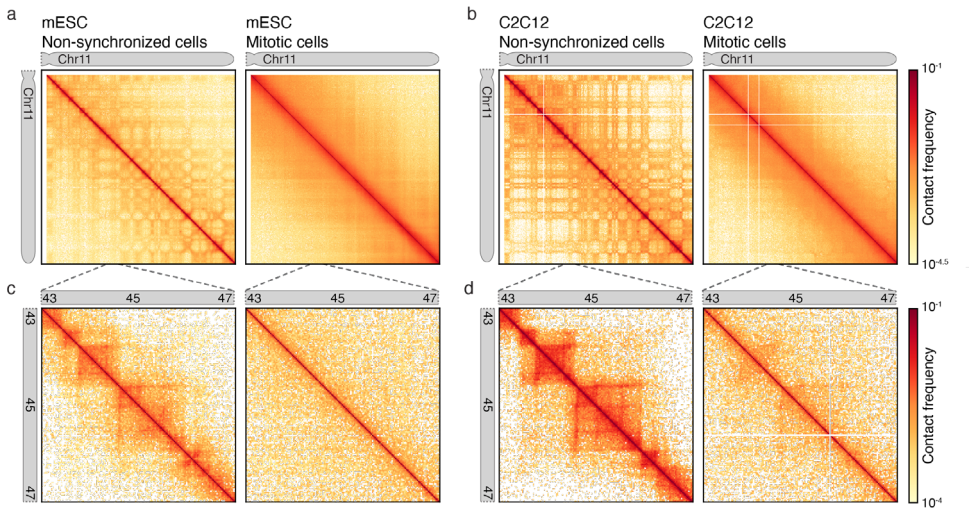


Figure 2 – Hi-C data shows compartments and TADs are lost in both mitotic mESCs and C2C12. (**a-b**) Hi-C heatmap of chr11 at 100kb bins for mESCs (a) and C2C12 (b) non synchronized cells (left panel) and mitotic arrested cells (right panel). (**c-d**) Zoom in Hi-C heatmap of chr11:43,000,000-47,000,000 at 25kb bins for mESC (c) and C2C12 (d) for non-synchronized cells (left panel) and mitotic arrested cells (right panel).

When we plot Hi-C data on a chromosome wide level (figure 2a-b), we see in interphase cells from both mESC and C2C12 clear compartment structures, represented as a checkerboard pattern in the heatmaps. Interestingly, the compartment signal in mESCs is much less pronounced compared to C2C12 cells. The strengthening of compartment signal during differentiation has recently been described in human cell lines[17]. When we next look at chromosome-wide heatmaps of mitotic cells, we find that compartments are lost in both C2C12 and mESCs. This is in line with the previous observations in differentiated human cell lines, where compartment signal is lost entirely in mitosis as well. We then continued to look at a 4Mb window within chr11 to observe TADs. Where in non-synchronized cells, TADs can be observed in both mESCs (figure 2c) and C2C12 cells (figure 2d), in mitosis these structures are lost. We note that there are faint structures visible in mitotic C2C12 cells. We do not expect that this is due to maintenance of TADs in mitosis, but it is more likely to be explained by the presence of a small number of interphase cells that contaminates the mitotically synchronized population.

Mitotic loop extrusion is not blocked by retained CTCF sites

Next, we set out to observe CTCF dependent loops in mitotic mESCs in order to investigate whether mitotic loop extruders condensin I and II are blocked

by bound CTCF. As described above no compartments and TADs are maintained in mitotic stem cells at individual genomic locations (figure 2). However, it requires high sequencing depth to observe individual CTCF-dependent loops in Hi-C data. To observe loop formation using less deeply sequenced Hi-C libraries, loops can be visualized by plotting the aggregate of Hi-C signal on either single CTCF sites (figure 3a-h) or on the pairwise interactions of CTCF sites (figure 3i-p). In line with the above described ATAC-seq analysis, we used CTCF sites that are categorized based on published ChIP-seq data[37,39] as mitotic bookmarked sites, mitotically reduced sites and sites that lose CTCF binding in mitosis.

When we aggregate Hi-C signal on all interphase-bound CTCF-sites on the diagonal a strong boundary can be observed at the center of the pile up plot in interphase cells (figure 3a). This represent the accumulation of insulating potential of CTCF at TAD boundaries, as it prevents the interaction of loci across the bound CTCF site[7,42]. Insulation can be the result of blocked loop extrusion at CTCF sites. This same phenomenon is observed when aggregating the signal in non-synchronized cells on mitotically bookmarked CTCF sites (figure 3b), reduced CTCF sites (figure 3c) and lost CTCF sites (figure 3d). We note that the insulation potential is strongest for bookmarked CTCF sites, compared to reduced and lost CTCF sites. Similar to the ATAC-seq experiments described above, Hi-C is performed on a population of cells. Therefore, a possible explanation could be that bookmarked CTCF sites are more likely to be bound by CTCF across the population, whereas lost CTCF sites are also captured in unbound state in the population. Contrary, when we plot these same pile-up plots for mitotic mouse stem cells, we see that all CTCF insulation is lost for each category of CTCF sites (figure 3e-h). This strongly implies that loop extrusion in mitosis is not blocked at sites where CTCF binding is maintained.

Likewise, we can plot the aggregation of Hi-C signal on pairwise CTCF interactions. We created a list of all possible pairwise interactions between two CTCF sites separated by up to 250kb. Typically, pairwise CTCF interactions are enriched in Hi-C interaction signal in interphase, as can be observed as a dot in the center of the pile-up plot representing loops between pairs of CTCF sites. Indeed, we see a clear enrichment at pairwise CTCF interactions in non-synchronized mESCs across all categories of CTCF sites (figure 3i-l). Similar to the observation found at an aggregation at single CTCF sites, this enrichment at pairwise CTCF sites is lost in mitosis (figure 3m-p). Combined these results suggest that although CTCF binding is maintained in mitosis at a substantial fraction of sites in mouse stem cells, CTCF does not have the ability to block mitotic loop extruders condensin I and II and therefore no CTCF-CTCF loops are formed.

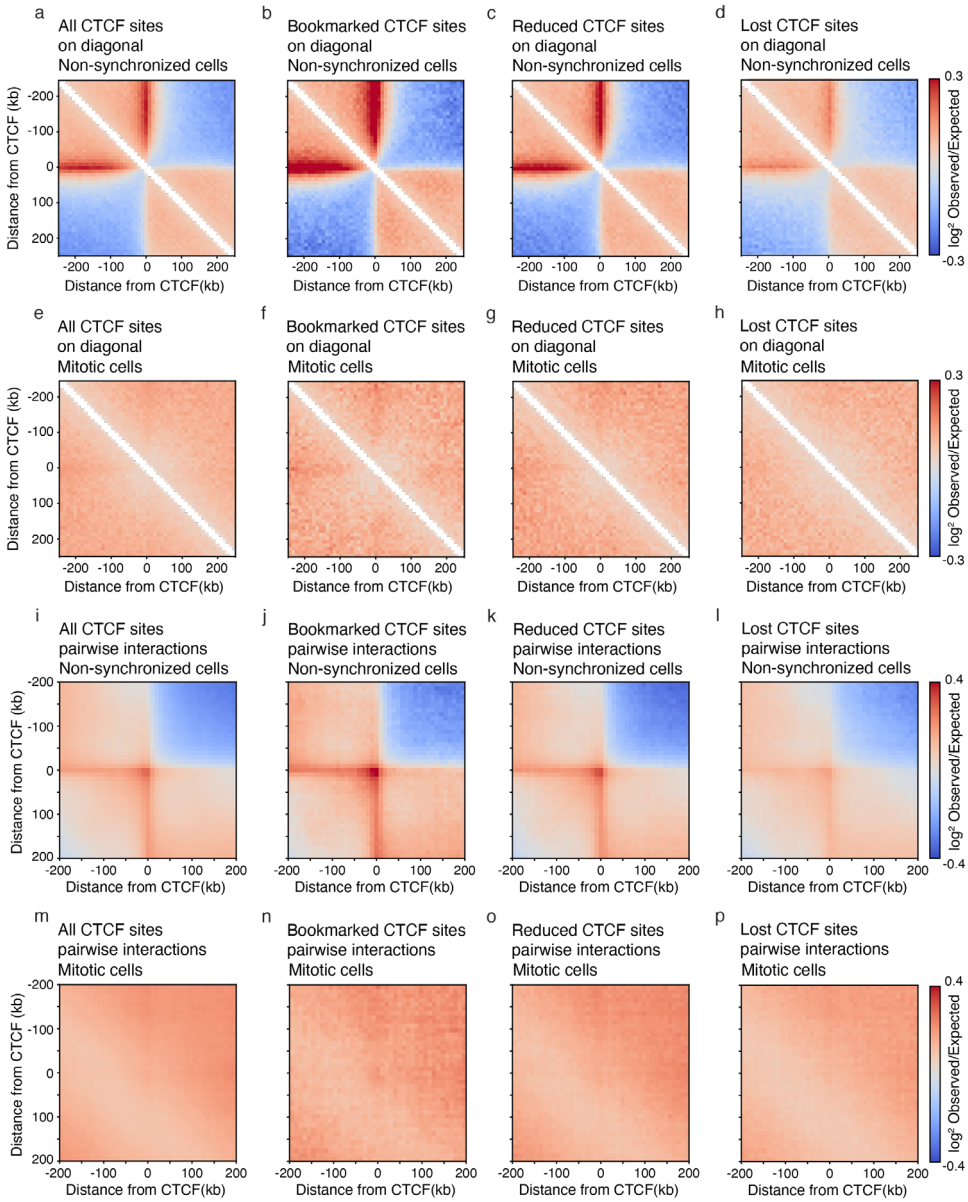


Figure 3 – Hi-C pile-up plots on single and pairwise CTCF sites show that loop extrusion by condensins in mitosis cannot be blocked by bound CTCF. (**a-d**) Aggregate of Hi-C signal binned at 10kb in non-synchronized mESCs on all interphase-bound CTCF sites (a), mitotic bookmarked sites (b), reduced CTCF sites (c), and CTCF sites that lose binding in mitosis (d). (**e-h**) Aggregate of Hi-C signal in mitotic mESCs on all interphase-bound CTCF sites (e), mitotic bookmarked sites (f), reduced CTCF sites (g), and CTCF sites that lose binding in mitosis (h). (**i-p**) Pile up of Hi-C signal in 10kb bins in non-synchronized (i-l) and mitotic (m-p) mESCs of pairwise interactions within 250kb at all interphase bound CTCF sites (i,m), bookmarked CTCF sites (j,n), reduced CTCF sites (k,o) and CTCF sites that lose binding in mitosis (l,p). All CTCF sites are plotted with respect to strand orientation of the motif.

Mitotic loop sizes differ between species

Hi-C data can be represented as a distance decay plot, where the interaction frequency *P* is plotted as a function of the genomic distance *s*. These *P*(*s*) plots have distinct shapes for both interphase and mitotic chromosomes[25]. By calculating the slope of *P*(*s*) and plotting the contact frequency derivative as a function of genomic distance, the average loop sizes present in interphase and mitosis can be revealed[24,26,44–46]. In addition to any differences between stem cells and differentiated cells, we were interested to study the loop characteristics of different species in interphase and mitosis. In addition to Hi-C data generated in this study, several studies have published Hi-C data on both non-synchronized and mitotic cells in different species[24,25,47], which enabled the comparison of chicken cells (cell line DT40), human cells (cell line HeLa) and mouse cells (cell lines mESCs, C2C12 and C127). We chose to plot the derivatives of P(*s*) for an acrocentric chromosome of similar length to allow for proper comparison between species (chr14 for both mouse and human and chr1 for chicken), although we did not find differences when calculating derivative plots for different chromosomes. In non-synchronized cell populations, all species behaved similarly (figure 4a). Loops in interphase have an average loop size of ~100kb across all species and cell lines (as highlighted with arrow in figure 4a). Interestingly, this is not the case for mitotic loops of these different species (figure 4b and zoom in figure 4c). Although there is no difference between the derivative plots of mouse stem cells mESCs and mouse differentiated cell lines C2C12 and DT40, a clear difference is observed between mitotic cells of human, mouse and chicken. A prominent feature of mitotic chromosome folding that can be observed in derivative plots, is the average size of the mitotic loops formed by condensins[24]. All mouse cell lines show an average mitotic loop size of 2 megabase (figure 4c, highlighted with circle), whereas human cell line HeLa shows a loop array size of 750kb in mitosis (figure 4c, highlighted with triangle) and chicken cell line DT40 has an average loop size of 350kb (figure 4c, highlighted with star). This suggests a different level of mitotic condensation between our three species.

We hypothesized that this difference in loop sizes could be related to chromosome length. Possibly, longer chromosomes require a higher level of condensation in order to ensure proper separation of sister chromatids in anaphase. When we plot all lengths of all chromosomes of the three species (figure 4d), it becomes clear that chicken chromosomes are on average much shorter than human and mouse chromosomes, however the mouse and human chromosomes have similar average chromosome length. The centromere is an important region of mitotic chromosomes where the mitotic spindle will attach, which will pull the sister chromatids apart during anaphase[48]. It is therefore more relevant to plot the

length of the longest arm of each chromosome, per definition the q-arm (figure 4e). Indeed, when we compare the q-arm length between our three species, we find that chicken has very short q-arms with an average length of 11Mb, followed by human chromosomes with an average q-arm length of 94Mb and an average q-arm length of 125Mb for mouse chromosomes. Combined, these results show that in the cell lines we investigated mitotic loop sizes are not related to cell type or differentiation state but instead differ between species. Moreover, our results suggest that there is a relationship between average q-arm length and the size of the mitotic loop array.

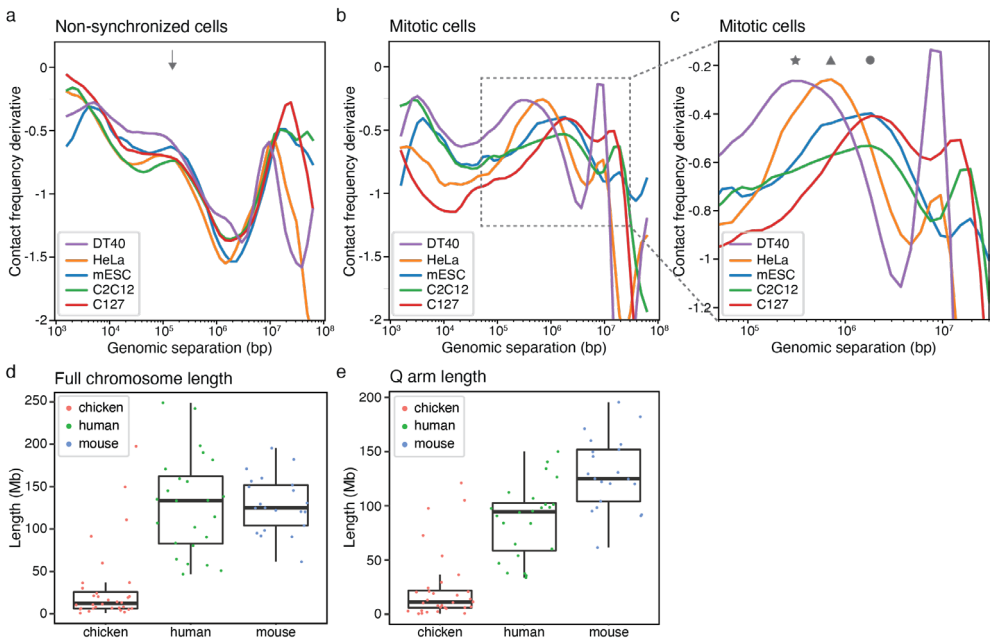


Figure 4 – Mitotic loop arrays species differ in average loop size between species. (**a**) Derivative of *P*(*s*) as a function of genomic separation in non-synchronized chicken cells (DT40), human cells (HeLa) and mouse cells (mESCs, C2C12 and C127). The arrow highlights the average loop size mediated by cohesin in interphase in all cell types and species (**b**) Derivative plots of Hi-C data from chicken cells (DT40), human cells (HeLa) and mouse cells (mESCs, C2C12 and C127) synchronized in mitosis. (**c**) A zoom-in of the derivative plot shown in figure 4b. The star highlights the average loop size observed in mitotic chicken cells, the triangle highlights the average loop size in mitotic human cells and the circle highlights the average loop size in mitotic mouse cells. (**d**) Boxplot of full chromosome lengths in chicken genome (galGal6), human genome (hg38) and mouse genome (mm10). Dots represent individual chromosomes. (**e**) Boxplot of all q-arm lengths in chicken genome (galGal6), human genome (hg38) and mouse genome (mm10). Dots represent individual chromosomes.

Discussion

In this study, we set out to explore mitotic chromosome organization in different cell types and vertebrate species. Although mitotic chromosomes are often perceived as universal structures, there are several characteristics that can differ between differentiation state and between species. First, we describe partial maintenance of CTCF binding in mitotic mouse stem cells. This is in contrast with our previous report on complete loss of mitotic CTCF binding in several differentiated cell lines (chapter 3 of this thesis)[32]. Interestingly, when mESCs are investigated by Hi-C, we observe that no interphase structures are maintained in mitosis despite maintained CTCF binding, suggesting that CTCF does not block mitotic loop extrusion by condensins. Lastly, we investigate whether mitotic chromosomes are differently organized between species. For this analysis, we used Hi-C data generated in mouse cell lines and publicly available data for mitotic human and chicken cell lines[24,47]. Although further experiments will be necessary, we find that the sizes of mitotic loops are different between species, but do not change between different cell lines of the same organism. Furthermore, our results suggest that mitotic loop size is correlated with the average length of the q-arm of chromosomes.

The result that stem cells maintain partial bookmarking of CTCF binding, raises the questions how CTCF binding is maintained in stem cells, if bookmarked CTCF has a function and why mitotic CTCF binding is not observed in differentiated cell lines. In addition to our study, several recent papers have reported similar retention of CTCF binding at a fraction of CTCF sites in mitotic mouse progenitor cells[36] and pluripotent cells[38] and even mouse sperm cells in meiosis II[49]. We hypothesized that CTCF loses binding to mitotic chromosomes in differentiated cell lines as a result of phosphorylation of zinc-fingers, which reduces the ability of CTCF to bind to its motif[32,50]. It is not known whether the CTCF protein in stem cells does not become phosphorylated or if CTCF retention is mediated in a different way. Additionally, we note that it has not yet been shown whether loss of CTCF binding in mitosis is an active process mediated by phosphorylation or chromatin remodelers or if this is a passive process as a result of general chromatin changes and condensation as cells progress through mitosis. The studies by the Blobel and Apostolou laboratories correlated the fraction of bookmarked CTCF sites with proximity to genes that are transcribed early in G1[36,38]. This could suggest that maintenance of CTCF binding in mitosis could promote fast transcription initiation upon mitotic exit. Additionally, mouse stem and progenitor cells have a much faster cell cycle compared to many differentiated cell lines (~12 hours in mESCs vs 24 hours in HeLa cells), which could necessitate fast transcription initiation upon mitotic exit.

Although we can only speculate about the potential function of maintained

CTCF binding upon G1 entry, we did not observe any function related to chromosome folding by bound CTCF in mitosis. When representing Hi-C data as individual loci or as pileups of Hi-C signal on CTCF sites, we did not find any evidence of TADs or CTCF loops as a result of maintained CTCF binding in mitosis. This is in line with modeling studies, which suggested that mitotic loop extruders condensin I and II are not blocked by bound CTCF and therefore explain the loss of TADs and loops in mitosis[11,51].

Lastly, we addressed the difference in average loop size in mitosis between different species. It has been proposed that mitotic loop arrays are formed by condensin I and II, where condensin II mediates loop formation in large loops with several smaller loops inside formed by condensin I[24]. Additionally, the ratio of condensin I and II modulate the level of condensation and the average loop sizes as has been observed as cell progress from prophase to mitosis[24], during development in mitotic *Xenopus* chromosomes[30] and when mitotic chromosomes are depleted from either condensin I or II[31]. Combined, this suggests that when chromosomes have longer chromosome arms, sister chromatids compact to a greater extent. This process can possibly be mediated by loading different ratios of condensin I and II on mitotic chromosomes. Although it has been described that vertebrate species appear to have different ratios of condensin I and II[52–56], to our knowledge this has not yet been systematically studied in relation to mitotic loop size and chromosome dimensions. This could be studied by proteomic or fluorescent imaging experiments to detect the amounts and ratios of condensin I and II.

Overall, we find that mitotic chromosomes indeed have characteristics that can differ between differentiation state and between species. On a detailed level we observe that CTCF binding is partially maintained at stem cells but is completely lost in differentiated cell lines. This confirms the hypothesis that condensins are not blocked by mitotically bound CTCF. On a larger chromosome-wide scale, we observe that the size of mitotic loops can differ between species, which could be correlated to the average length of the chromosome q-arm. Although all vertebrate mitotic chromosomes are folded as an array of loops mediated by condensin I and II, the ratio at which condensins are loaded onto chromosomes could modulate the dimensions of chromosomes to generate long and thin or short and fat chromosomes.

Methods

ATAC-seq analysis

ATAC-seq sequencing reads were trimmed to 24bp and aligned to reference genome mm10 using Bowtie2 with a maximum mapping length of 2000bp[57,58]. Paired-end reads were filtered for mapping quality, mitochondrial reads and PCR duplicates. ATAC-seq data was plotted as V-plots[40] on all interphase bound CTCF motifs (51805 sites) and on CTCF motifs categorized as bookmarked (10799 sites), reduced (18704 sites) or lost (22302 sites) in mitosis by N.O. and P.N. V-plots were produced as described in chapter 3 of this thesis[32]. To plot V-plots, CTCF motifs were oriented in the same direction.

Cell culture and synchronization conditions

Mouse embryonic stem cells (E14TG2a) were cultured and synchronized with a 6 hour nocodazole arrest following previous publications[33,37]. C2C12 cells were cultured in DMEM media supplemented with Glutamax-I, 10% heat-inactivated FBS and penicillin-streptomycin. C2C12 cells were synchronized with nocodazole arrest (50ng/mL) for 8 hours. Mitotic C2C12 and mES cells were harvested by mitotic shake off. Both mitotic and asynchronous cultures were fixed with 1% formaldehyde and stored at -80°C until processed for Hi-C.

Hi-C

Hi-C on mitotic and asynchronous cultures were performed according to previously published protocol[4]. Briefly cells were fixed and stored as described above. Crosslinked cells were thawed, lysed and digested with DpnII restriction enzyme overnight at 37°C. Restriction overhangs were filled with biotin-14-dATP supplemented with dTTP, dCTP and dGTP for 4 hours at 23°C, followed by ligation using T4 DNA ligase at 16°C for another 4 hours. Samples were then treated with proteinase K at 65°C overnight. DNA was cleaned up and purified using phenol:chloroform and ethanol precipitation. DNA was sonicated and size selection to average size of 100-350bp using AMpure XB beads, followed by end repair. Samples were enriched for biotin-tagged DNA fragments by pull down using streptavidin beads. After A-tailing, libraries were ligated with indexed Illumina TruSeq sequencing adapters, followed by pcr amplification. Finally, libraries were cleaned up from PCR primers using AMpure XP beads and sequenced using paired-end 50bp sequencing on an Illumina HiSeq 4000.

Hi-C mapping and downstream analysis

Hi-C sequencing files were mapped to reference genomes hg38 (HeLa data),

mm10 (C2C12, mESC and C127 data) and galGal6 (DT40 data) using publicly available distiller-nf mapping pipeline (https://github.com/mirnylab/distiller-nf) and downstream analysis tools pairtools (https://github.com/mirnylab/pairtools) and cooltools (https://github.com/mirnylab/cooltools). Briefly, reads were mapped using bwa-mem, pcr duplicates were removed and reads were filtered for mapping quality. Distance decay and derivative plots created using cooltools code by calculating contact frequency (P) as a function of genomic distance (*s*) using valid pairs. For further downstream analysis, interactions were binned in matrices at a range of different resolutions using cooler[59]. Iterative balancing was applied to all matrices, while ignoring the first two bins from the diagonal[60]. Pile up plots at single CTCF sites and pairwise CTCF interactions were produced using observed over expected signal binned at 10kb. Pairwise CTCF sites for pile up plots were predicted by pairing all CTCF sites within 250kb on the same chromosome within the CTCF category (CTCF sites bookmarked in mitosis, reduced in mitosis or lost in mitosis) following curation by N.O and P.N. Directionality of the CTCF motifs were taken into account and all motifs were orientated in the same direction.

Code availability

Hi-C mapping pipeline distiller-nf is available on Github: https://github.com/mirnylab/distiller-nf. Downstream analysis tools pairtools and cooltools are available through https://github.com/mirnylab/pairtools and https://github.com/mirnylab/cooltools. Code used for analysis of ATAC-seq data can be found at Github: https://github.com/dekkerlab/CTCF_in_mitosis_GR_2018.

Publicly available data used in this study

In addition to the Hi-C data that was generated for this study, we use several ATAC-seq and Hi-C datasets that are publicly available on the gene expression omnibus (GEO). ATAC-seq data in mESC[37] is available under accession number GSE122589. Hi-C data in asynchronous and mitotic DT40 cells and HeLa are available under GSE102740[24] and Hi-C data of mouse cell line C127 under GSE149677[47].

Author contributions

M.E.O and J.D conceived and designed the project. M.E.O performed all Hi-C experiments and analysis with input of J.D. and P.N. T.P. cultured and synchronized mESC cells and M.E.O cultured and synchronized C2C12 cells for Hi-C experiments. N.O. and P.N. curated list of CTCF classes based on published ChIP-seq data. M.E.O. reanalyzed published ATAC-seq datasets. M.E.O wrote the manuscript with input from J.D.

Acknowledgements

We thank all current and former members of the Dekker lab and the Navarro lab for helpful discussions and suggestions, especially Johan Gibcus and Bastiaan Dekker. We thank Eugenio Mattei for advice on computational analyses. This work was supported by a grant from the National Institutes of Health (HG003143 to J.D.). J.D. is an investigator of the Howard Hughes Medical Institute.

References

1. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–11 (2002).
2. Dostie, J. *et al.* Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**, 1299–309 (2006).
3. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–93 (2009).
4. Belaghzal, H., Dekker, J. & Gibcus, J. H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* **123**, 56–65 (2017).
5. Erdel, F. & Rippe, K. Formation of Chromatin Subcompartments by Phase Separation. *Biophys. J.* **114**, 2262–2270 (2018).
6. Michieletto, D., Orlandini, E. & Marenduzzo, D. Polymer model with epigenetic recoloring reveals a pathway for the de novo establishment and 3D organization of chromatin domains. *Phys. Rev. X* **6**, 1–15 (2016).
7. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–80 (2012).
8. Rao, S. S. P. *et al.* A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* **159**, 1665–1680 (2014).
9. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl. Acad. Sci. U. S. A.* 196261 (2018). doi:10.1073/pnas.1717730115
10. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–5 (2012).
11. Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 1–12 (2016). doi:10.1101/024620
12. Nora, E. P. *et al.* Targeted degradation of CTCF decouples local insulation of chromosome domains from higher-order genomic compartmentalization. *bioRxiv* (2016). doi:10.1111/j.1574-6941.2012.01443.x
13. Rao, S. S. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305-320.e24 (2017).
14. Dekker, J. & Mirny, L. The 3D Genome as Moderator of Chromosomal Communication. *Cell* **164**, 1110–1121 (2016).
15. Sanborn, A. L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci.* (2015). doi:10.1073/pnas.1518552112
16. Wit, E. De *et al.* CTCF Binding Polarity Determines Chromatin Looping. *Mol. Cell* **60**, 1–9 (2015).
17. Oksuz, B. A., Yang, L., Abraham, S., Venev, S. V & Krietenstein, N. Systematic evaluation of chromosome conformation capture assays. 0–42 (2020).
18. Smith, E. M., Lajoie, B. R., Jain, G. & Dekker, J. Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the CFTR Locus. *Am. J. Hum. Genet.* **98**, 185–201 (2016).
19. Lupianez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
20. Valton, A. L. & Dekker, J. TAD disruption as oncogenic driver. *Curr. Opin. Genet. Dev.* **36**, 34–40 (2016).
21. Earnshaw, W. C. & Laemmli, U. K. Architecture of metaphase chromosomes and chromosome scaffolds. *J. Cell Biol.* **96**, 84–93 (1983).
22. Marsden, M. P. F. & Laemmli, U. K. Metaphase chromosome structure: Evidence for a radial loop model. *Cell* **17**, 849–858 (1979).
23. Flemming, W. Zur Kenntnis der Zelle und ihrer Teilung-Erscheinungen. *Schr.*

Nat. Wiss. Ver. Schlesw.-Holst. **3**, 23–27 (1878).
24. Gibcus, J. H. *et al.* A pathway for mitotic chromosome formation. *Science (80-.).* **359**, eaao6135 (2018).
25. Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342**, 948–53 (2013).
26. Abramo, K. *et al.* A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. *Nat. Cell Biol.* **21**, 1393–1402 (2019).
27. Belmont, A. S. Mitotic chromosome structure and condensation. *Curr. Opin. Cell Biol.* **18**, 632–638 (2006).
28. Batty, P. & Gerlich, D. W. Mitotic Chromosome Mechanics: How Cells Segregate Their Genome. *Trends Cell Biol.* **xx**, 1–10 (2019).
29. Fazzio, T. G. & Panning, B. Condensin complexes regulate mitotic progression and interphase chromatin structure in embryonic stem cells. *J. Cell Biol.* **188**, 491–503 (2010).
30. Kieserman, E. K. & Heald, R. Mitotic chromosome size scaling in Xenopus. *Cell Cycle* **10**, 3863–3870 (2011).
31. Shintomi, K. & Hirano, T. The relative ratio of condensing I to II determines chromosome shapes. *Genes Dev.* **25**, 1464–1469 (2011).
32. Oomen, M. E., Hansen, A. S., Liu, Y., Darzacq, X. & Dekker, J. CTCF sites display cell cycle–dependent dynamics in factor binding and nucleosome positioning. *Genome Res.* 1–14 (2019). doi:10.1101/gr.241547.118.
33. Festuccia, N. *et al.* Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nat. Cell Biol.* **18**, (2016).
34. Wang, F. & Higgins, J. M. G. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol.* **23**, 175–84 (2013).
35. Hsiung, C. C. *et al.* Genome accessibility is widely preserved and locally modulated during mitosis. 1–29 (2015). doi:10.1101/gr.180646.114
36. Zhang, H. *et al.* Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* **576**, 158–162 (2019).
37. Festuccia, N. *et al.* Transcription factor activity and nucleosome organization in mitosis. *Genome Res.* **29**, 250–260 (2019).
38. Pelham-Webb, B. *et al.* Mitotic retention of H3K27 acetylation promotes rapid topological and transcriptional resetting of stem cell-related genes and enhancers upon G1 entry. *bioRxiv* (2020). doi:10.1101/2020.06.02.130104
39. Owens, N. *et al.* CTCF confers local nucleosome resiliency after dna replication and during mitosis. *Elife* **8**, 1–26 (2019).
40. Zentner, G. E. & Henikoff, S. Surveying the epigenomic landscape, one base at a time. *Genome Biol.* **13**, 250 (2012).
41. Fu, Y., Sinha, M., Peterson, C. L. & Weng, Z. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet.* **4**, (2008).
42. Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-944.e22 (2017).
43. Rao, S. S. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305-320.e24 (2017).
44. Haarhuis, J. H. I. *et al.* The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. 693–707 (2017). doi:10.1016/j.cell.2017.04.013
45. Gassler, J. *et al.* A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* **36**, 3600–3618 (2017).
46. Schwarzer, W. *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
47. Fitz-James, M. H. *et al.* Large domains of heterochromatin direct the formation of short mitotic chromosome loops. *Elife* **9**, 1–28 (2020).
48. McKinley, K. L. & Cheeseman, I. M. The molecular basis for centromere identity and function. *Nat. Rev. Mol. Cell Biol.* (2015). doi:10.1038/nrm.2015.5

49. Jung, Y. H. *et al.* Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. *Cell Rep.* **18**, 1366–1382 (2017).
50. Sekiya, T., Murano, K., Kato, K., Kawaguchi, A. & Nagata, K. Mitotic phosphorylation of CCCTC-binding factor (CTCF) reduces its DNA binding activity. *FEBS Open Bio* (2016). doi:10.1002/2211-5463.12189
51. Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A. & Mirny, L. A. Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harb. Symp. Quant. Biol.* **LXXXII**, 034710 (2018).
52. Vagnarelli, P. Mitotic chromosome condensation in vertebrates. *Exp. Cell Res.* **318**, 1435–41 (2012).
53. Ohta, S. *et al.* The Protein Composition of Mitotic Chromosomes Determined Using Multiclassifier Combinatorial Proteomics. *Cell* **142**, 810–821 (2010).
54. Ono, T. *et al.* Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. *Cell* **115**, 109–121 (2003).
55. Hirota, T., Gerlich, D., Koch, B., Ellenberg, J. & Peters, J. M. Distinct functions of condensin I and II in mitotic chromosome assembly. *J Cell Sci* **117**, 6435–6445 (2004).
56. Green, L. C. *et al.* Contrasting roles of condensin I and condensin II in mitotic chromosome formation. *J. Cell Sci.* **125**, 1591–1604 (2012).
57. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–9 (2012).
58. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–8 (2013).
59. Abdennur, N. & Mirny, L. A. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics* **36**, 311–316 (2019).
60. Imakaev, M. *et al.* Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* **9**, 999–1003 (2012).

Chapter 5

Detecting chromatin interactions along and between sister chromatids with SisterC

Marlies E. Oomen[1], Adam K. Hedger[2], Jonathan K. Watts[2], Job Dekker[1,3]

Published in *Nature Methods.* (2020) 17, 1002–1009

1 Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA;
2 RNA Therapeutics Institute and Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA.
3 Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA.

Abstract

Accurate chromosome segregation requires chromosome compaction with concordant disentanglement of the two sister chromatids. This process has been studied extensively by microscopy but has remained a challenge for genomic methods, such as Hi-C, because sister chromatids have identical DNA sequences. Here we describe SisterC, a chromosome conformation capture assay that can distinguish interactions between and within sister chromatids. The assay is based on BrdU incorporation during S-phase, which labels the newly replicated strands of the sister chromatids. This is followed by Hi-C, e.g. during different stages of mitosis, and then the selective destruction of BrdU containing strands by UV/Hoechst treatment. After PCR amplification and sequencing of the remaining intact strands, this allows for the assignment of Hi-C products as inter- and intra-sister interactions by read orientation. We performed SisterC on mitotically arrested *S. cerevisiae* cells. As expected, we find prominent interactions and alignment of sister chromatids at their centromeres. Along the arms, sister chromatids are less precisely aligned with inter-sister connections every ~35kb. In many instances, inter-sister interactions do not involve the interaction of two identical loci but occur between cohesin binding sites that can be offset by 5 to 25kb. Along sister chromatids, extruding cohesin forms loops up to 50kb. Combined, SisterC allows the observation of the complex interplay between sister chromatid compaction and sister chromatid segregation as the cell transitions from late S-phase to mitosis. SisterC should be applicable to study mitotic events in a wide range of organisms and cell types.

Introduction

During S-phase, when sister chromatids are formed, they are closely held together by the cohesin complex. Sister chromatids are initially also thought to be wrapped around each other and entangled. During the subsequent mitosis, sister chromatids become compacted and, in the process, become disentangled and segregated from each other, although they remain aligned side by side[1]. Classically, this process has been studied using microscopic methods by labeling sister chromatids differently using thymidine analogues[2,3]. It has been difficult to study this complex series of mitotic events using genomic techniques such as Hi-C, as sequencing-based methods cannot distinguish the identical sister chromatids, and therefore cannot differentiate interactions between and along sister chromatids. Recently an assay detecting sister chromatid exchange events allowed mapping of sister chromatid interactions genome wide in bacteria[4]. However, this approach requires extensive genome editing to introduce sister chromatid exchange markers throughout the genome.

Here we present a Hi-C-based assay, SisterC, that can detect and distinguish inter- and intra-sister interactions. We demonstrate the performance of the assay by studying mitotic *S. cerevisiae* cells. In *S. cerevisiae* the cohesin complex mediates inter-sister interactions at the centromere and along the chromosome arms ("cohesive cohesin")[5,6]. In addition, the cohesin complex compacts the sister chromatid arms by forming intra-sister loops by dynamic loop extrusion ("extruding cohesin")[7–11]. The latter is different from many other organisms, including mammals, where the condensin complex compacts chromosome arms. Condensin in yeast does act on the centromeres and rDNA loci[7]. SisterC reveals the alignment of sister chromatids, the positioning of inter-sister interactions and intra-sister loops and how the interplay between cohesive and extruding cohesin shapes the mitotic chromosome.

Results

SisterC library production after induction of single strand breaks at BrdU incorporation sites

Sister chromatids are identical in sequence but differ in which strand was newly replicated in S-phase. This difference can be leveraged to differentiate interactions between and within sister chromatids. BrdU containing DNA strands can be selectively degraded after Hoechst treatment and radiation with UV[3,12]. This property has been used before in Strand-seq[13], which allows the detection of sister chromatid exchange events[14,15], and the mapping of structural genomic variants such as polymorphic inversions[16] onto completely phased human genomes[17]. Here we describe SisterC, a method that combines Hi-C[18,19] on BrdU incorporated DNA and UV/Hoechst treatment

to distinguish interactions between sister chromatids (inter-sister interactions) and along sister chromatids (intra-sister interactions). Briefly, SisterC works as follows (figure 1a-c). When cells go through S-phase in the presence of BrdU, this results in cohesed pairs of sister DNA molecules where one molecule contains BrdU in the – strand (assigned as sister A), and the other molecule contains BrdU in the + strand (sister B) (figure 1a). Hi-C ligation of crosslinked and digested fragments of these DNA molecules results in a matrix of 16 possible ligation products between and within sister A and sister B that differ in the orientation of ligated fragments and the strand or strands that contain BrdU (figure 1b-c). Treatment of the ligation products with Hoechst and UV light creates single strand nicks in strands containing BrdU. Upon PCR amplification, this results in specific depletion of Hi-C products for which BrdU was present in both the + and – strands as for these products no intact template strand remains. Only 8 possible ligation products will still have one completely intact strand after UV irradiation, and these will therefore be PCR amplified (figure 1c). Four of these amplifiable ligation products will be interactions between sister chromatids and four products will be interactions along a sister chromatid, but they differ in fragment orientation. Strand orientations can be assigned after paired-end sequencing and mapping of genomic locations on both sides of the ligation junction. The orientation of SisterC fragments can then be used to differentiate interactions between and along sister chromatids: + + or – – read orientations represent inter-sister interactions, while + – or – + read orientations represent intra-sister interactions.

Here we choose to study mitotic sister chromatid interactions in *S. cerevisiae* using SisterC. Budding yeast has a relatively small and haploid genome. Furthermore, it can be synchronized in both late G1 and mitosis, which allow controlled incorporation of BrdU for exactly one S-phase. Lastly, its mappable centromeres and known cohesin binding sites along arms offer sites of particular interest, as these are sites at which sister chromatids are connected and intra-sister chromatid loops may form. As wild-type *S. cerevisiae* cells do not take up nucleosides from the environment, we use a strain that expresses human equilibrative nucleoside transporter (hENT) and *Drosophila* deoxyribonucleoside kinase DmdNK, which allows cells to take up BrdU from the environment and incorporate it into their DNA[14,20]. Cells were synchronized in late G1 using alpha factor, the culture was split and released in media containing either BrdU or thymidine. This was followed by nocodazole treatment to obtain cells arrested in mitosis or was followed by a second alpha factor incubation to obtain cells arrested in the subsequent G1 (supplemental figure 1).

To investigate the efficiency of SisterC, we performed several control experiments. First, to determine the amount of BrdU incorporation required to achieve efficient Hoechst/UV-induced strand destruction, we PCR amplified DNA

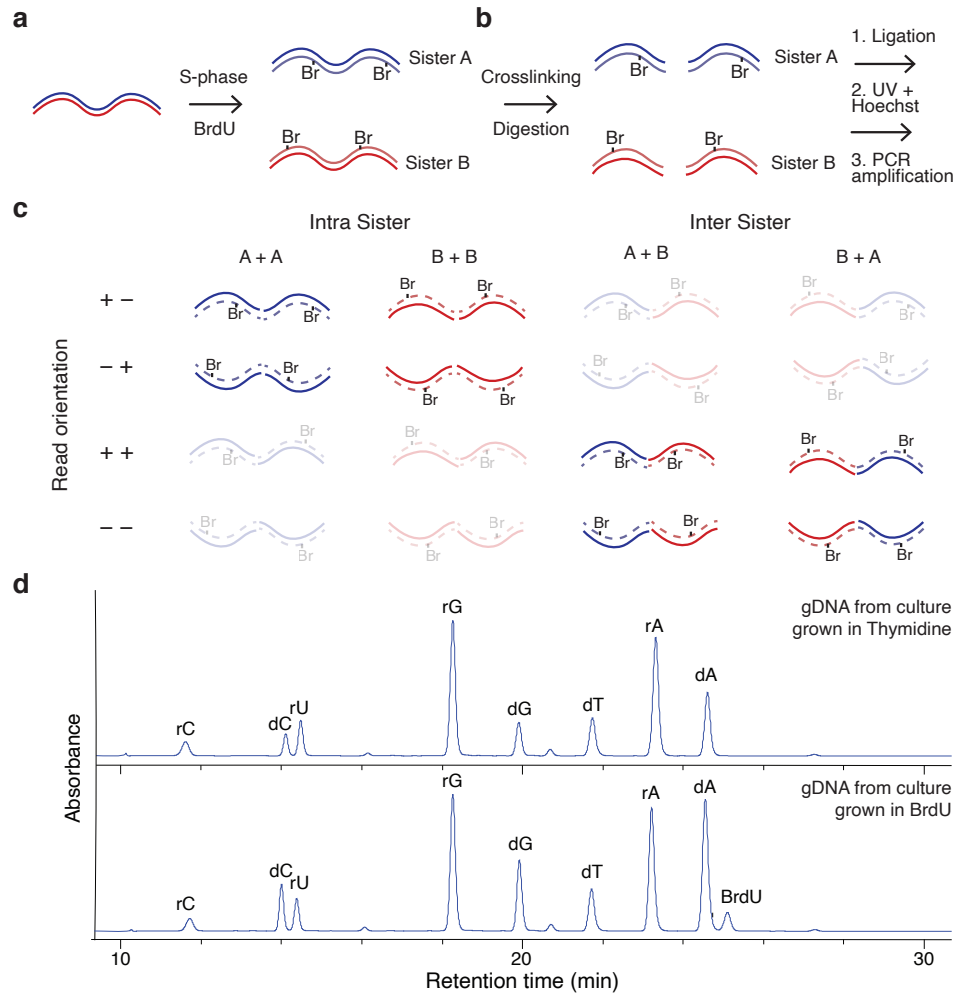


Figure 1 – Outline of SisterC chromosome conformation capture technique. (a) As cells go through replication BrdU is incorporated in the – strand for sister A and in the + strand for sister B. (**b**) During Hi-C/SisterC DNA is digested followed by proximity ligation, UV-Hoechst treatment and PCR-amplification. (**c**) Proximity ligation leads to 16 possible Hi-C products of inter- and intra-sister interactions. For SisterC libraries, DNA molecules were treated with Hoechst and radiated with UV, which introduces single strand nicks. This leads to destruction of DNA strands containing BrdU. PCR will lead to amplification only of these products with at least one full length template strand. 8 different ligation products are amplifiable. After paired-end sequencing, we can identify inter-sister (+ + and – –) and intra-sister interactions (– + and + –) by read orientation. (**d**) HPLC spectra of digested DNA isolated from yeast cultures that progressed through 1 round of DNA replication in BrdU- or thymidine-containing media; this enables quantification of BrdU incorporation. Percentages of each nucleoside are calculated using the extinction coefficient of each nucleoside. (**e**) Flow cytometry analysis of cell cycle profile and BrdU incorporation of harvested yeast cultures for preparation of SisterC libraries. (**f**) Table of BrdU content and percentage inter-sister interactions in each SisterC replicate preformed in this study.

fragments using a range of dTTP to BrdUTP ratios, resulting in products with BrdU incorporated in both strands. After treatment with Hoechst and UV radiation, we amplified these products again (supplemental figure 2). This showed that PCR amplification efficiency of DNA fragments containing more than 10 to 50% BrdUTP after treatment with both UV and Hoechst is greatly reduced, indicating that template strands were successfully broken. Second, we performed flow cytometry to detect BrdU incorporation and cell cycle profile of the synchronized yeast cultures (figure 1e and supplemental figure 3). We observe proper cell synchronization in mitosis and late G1 and uniform BrdU incorporation across the cell population. Third, we directly measured the BrdU incorporation efficiency in yeast cells by determining the base composition of genomic DNA using HPLC. HPLC allows to quantitatively measure each nucleotide present in genomic DNA extracted from cultures grown in BrdU or thymidine (figure 1d and supplemental figure 4). We identified all peaks in the HPLC spectra by mass spectrometry (supplemental figure 5) and detected all DNA and RNA nucleosides, including BrdU. Note that each nucleoside has a different extinction coefficient, which affects the area of the peak as measured by UV absorbance at 260nm in the HPLC spectra (this results in the relatively low BrdU peak area at 260nm; compare to HPLC spectrum measured at 279nm in supplemental figure 6). After adjustment of the peak area in the spectra by each extinction coefficient, we find that 82.7 up to 98.1% of all Ts in the newly replicated strand are replaced by BrdU (figure 1f and supplemental table 1). Lastly, we estimated the efficiency of selective depletion of Hi-C ligation products as result of UV/Hoechst treatment by producing SisterC libraries for cells synchronized in late G1 after BrdU incorporation during the previous S-phase. As each cell went through the cell cycle and divided into two cells, there are no longer sister chromatids in G1 cells. Using these cells to perform SisterC, we were able to estimate the percentage of false inter-sister interactions as these should no longer be present in a G1 SisterC library. Although we do not detect full depletion of all inter-sister interactions in G1 libraries, we do see a depletion down to approximately 20% (figure 1f and supplemental table 1). In contrast, in mitotic SisterC libraries we typically see around 35-40% captured inter-sister interactions (figure 1f and supplemental table 1), supporting that in SisterC inter-sister interactions are enriched in pairs with – – and + + read orientation and intra-sister interactions are enriched in pairs with – + and + – read orientation.

SisterC allows observation of interactions between and along sister chromatids

We carried out three biological replicates of SisterC experiments using DpnII with highly concordant results and one SisterC experiment using HindIII (supplemental figure 7 and supplemental table 1). After mapping and standard Hi-C data processing

(see methods), we examined the SisterC interaction frequency (*P*) as a function of genomic distance (*s*) separated by strand orientation (figure 2a-b and supplemental figure 8a-b). Here the definition of genomic distance for interactions between sister chromatids is the difference between their respective genomic coordinates, even though this involves two different DNA molecules. We compared the results to Hi-C controls (supplemental figure 8c-d) and SisterC negative controls of cells cultured in thymidine instead of BrdU (supplemental figure 8e-h). As expected from any Hi-C library, read orientation of interactions between loci separated by less than 1500bp are influenced by technical artifacts, such as unligated dangling ends (+ – orientation) and self-ligated fragments (– + orientation)[21]. Therefore, we exclude any interactions between loci separated by less than 1500bp in all our analyses (supplemental figure 8a-b). In the G1 SisterC libraries (figure 2a), we see a clear depletion in $P_{inter}(s)$ of inter-sister interactions as expected, as there are no longer sister chromatids in G1. Note that the $P_{inter}(s)$ of interactions assigned as inter-sister contacts runs parallel to interactions assigned as intra-sister. This suggests that the former interactions with + + and – – read orientations are in fact intra-sister interactions that were not successfully depleted by UV/Hoechst treatment.

Interestingly, we find a very different $P_{inter}(s)$ for inter-sister interactions in mitosis (figure 2b). Inter-sister interactions are less frequent than intra-sister interactions for loci separated by short genomic distances (below 30kb). $P(s)$ for inter-sister and intra-sister interactions converge for distances larger than 35kb. This describes two phenomena. First, sister chromatids in yeast are not perfectly aligned, as perfectly aligned sister chromatids would result in a minimal difference in $P(s)$ between inter-sister and intra-sister interactions (figure 2c, top model). Instead, interactions with genomic distance below 35kb occur more frequently within the same sister than interactions between sisters, suggesting sister chromatids are loosely aligned (figure 2c, bottom model, left panel). Second, interaction frequency of inter-sister and intra-sister interactions between loci separated by more than 35kb distance converge, which implies that the likelihood of an interaction being inter-sister or intra-sister has become identical (figure 2c, bottom model, right panel). Together, this suggests that although the sister chromatids are not perfectly aligned, sisters are loosely held together mediated by inter-sister interactions that are spaced every 35kb on average (figure 2c).

When mitotic SisterC data are visualized as an interaction heatmap, e.g. for chromosome XIII (figure 2d), we observe similar characteristics as seen in $P_{inter}(s)$. Inter-sister interactions show a weaker interaction signal around the diagonal, representing short-range interactions up to 30kb, compared to intra-sister interactions. At larger distances this difference in interaction frequency is no longer detectable.

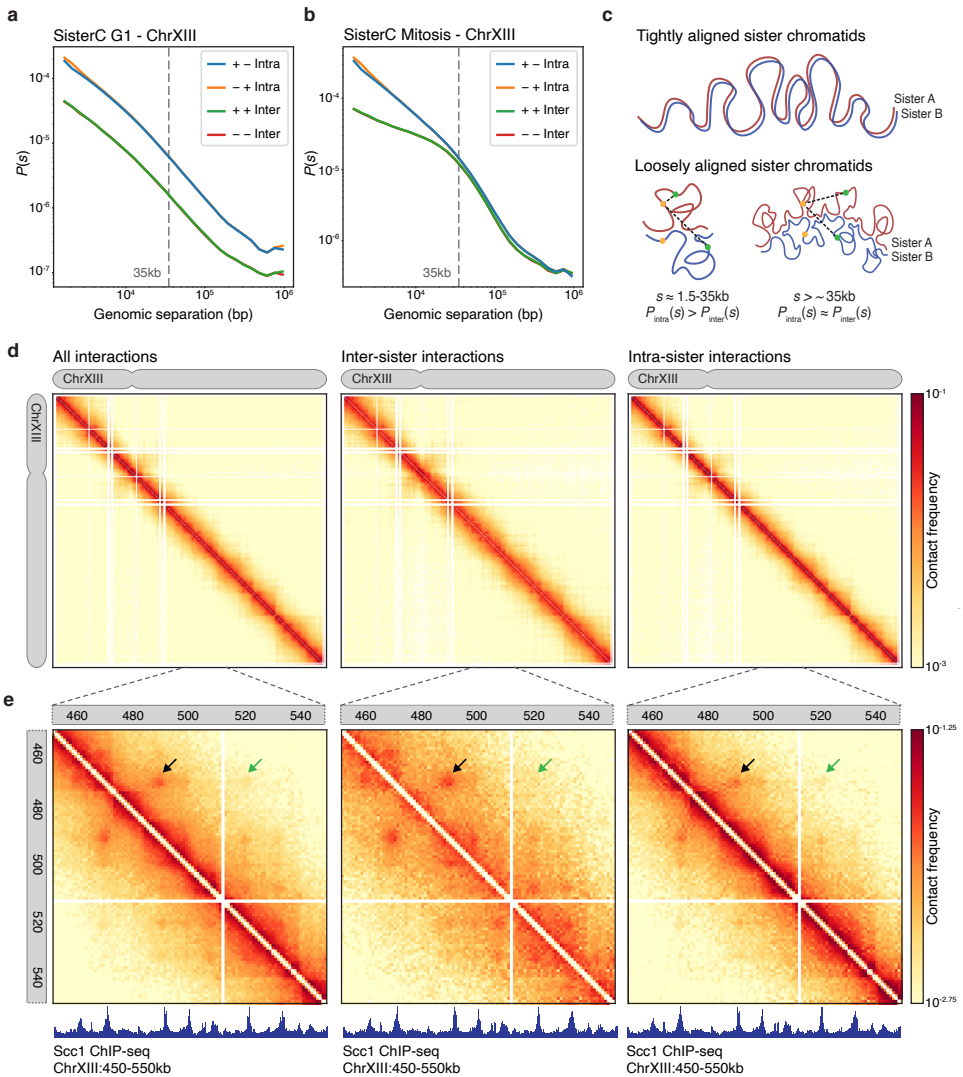


Figure 2 – SisterC allows differentiation of inter-sister and intra-sister interactions in mitotic chromosomes. (a) Chromatin interaction distance-dependent decay shows uniform depletion of read pairs with + + and – – orientation in G1 synchronized yeast (loci separated by more than1500bp). (**b**) Mitotic SisterC libraries show an offset of inter-sister interactions up to 35kb genomic separation, with interactions between loci separated by less than 35 kb depleted compared to intra-sister interactions. (**c**) Model of tightly aligned sister chromatids (top panel) and loosely aligned sister chromatids (bottom panel). (**d**) Chromosome-wide SisterC interactions on chrXIII of all read pairs (left panel), all read pairs assigned as inter-sister interactions (middle panel) and all read pairs assigned as intra-sister interactions (right panel), binned in 2kb bins. (**e**) SisterC interactions at 1kb resolution of all reads on zoomed in region of chrXIII:450,000-550,000 of all reads (left panel), inter-sister reads (middle panel) and intra-sister reads (right panel). Arrows highlight interaction of cohesin sites that are more prevalent as inter-sister interaction (black) or intra-sister interaction (green). Lower panels show Scc1 ChIP-seq track, a subunit of the cohesin complex.

However, zooming in to a smaller region on chromosome XIII, we observe features that appear stronger in either the inter-sister heatmap or the intra-sister heatmap (figure 2e). In the combined SisterC interaction map (the sum of inter- and intra-sister interactions, figure 2e left panel), we observe dots that represent interactions between cohesin binding sites as detected by ChIP-seq[22]. When inter-sister or intra-sister interactions are plotted separately, we observe that some of these interactions are more

prominent in the inter-sister dataset (black arrow, figure 2e), while others are more prominent in the intra-sister dataset (green arrow, figure 2e). This difference is further highlighted after correction for the distance dependent expected interaction frequency (supplemental figure 9). The finding that some of these interactions between cohesin binding sites are detected in both inter- and intra-sister datasets, albeit with different frequencies, could be related to the fact that Hoechst/UV depletion is not complete. Perhaps more interestingly, this could also be due to variability in the population in the positioning of inter- and intra-sister interactions. As any Hi-C library, SisterC detects interactions as a population average. The observation that cohesin binding sites pair in different frequencies as interactions between and along sister chromatids can be interpreted in this context. Where in one cell a given cohesin site interacts with a second cohesin site along the same chromatid to mediate an intra-sister interaction, in another cell this cohesin site interacts with the same second cohesin site but located on the other sister chromatid to form an inter-sister interaction. Figure 2e illustrates a second example of cell to cell variability: the inter-sister interactions highlighted by the black arrow and the intra-sister interaction highlighted with the green arrow, are mediated by the same cohesin binding site at position chrXIII:469,450. One interpretation is that in one cell this site is involved in an inter-sister interaction where in another cell this site is mediating an intra-sister interaction. Interestingly, inter-sister interactions seem to occur at a shorter genomic distance than intra-sister interactions, as will be described in more detail below.

Sister chromatid interactions at centromeres

We set out to explore SisterC data around centromeres as we expect enrichment of inter-sister interactions at these sites. Centromeres display very prominent binding of both condensin and cohesin, where they mediate inter-sister interactions and possibly intra-sister interactions[7]. When we aggregate all inter- and intra-sister interactions combined around all 16 centromeres (figure 3a), we observe a striking pattern. Regions directly adjacent on either side of the centromere interact frequently with sequences up to 10 to 15kb away from the centromere. When interactions between sisters and interactions along the sisters are plotted

separately (figure 3b-c), we find that inter-sister interactions contribute differently to the combined heatmap than intra-sister interactions do. First, intra-sister interactions are depleted in a 2kb window centered precisely at the centromere. Second, we observe that, compared to the genome wide average expected level, inter-sister interactions at centromeric regions are enriched up to 10kb away from the diagonal (figure 3b and 3d). Third, both intra- and inter-sister interactions contribute to the line-like features emanating from the centromere in the Hi-C maps. Line-like features in Hi-C maps have been interpreted as dynamic or variable loops with one fixed anchor and with the second anchor at various distances away[23]. In this instance the fixed inter-sister connection and intra-sister loop anchor is located directly adjacent to the centromere. This fixed site then is engaged in an inter-sister interaction with a site located on the other sister chromatid some variable distance away from the centromere but on the same chromosome arm. Similarly, within a sister chromatid the fixed loop anchor would engage with a second anchor at some variable distance. Averaged over 16 centromeres, lines appear, but at individual centromeres such intra-sister looping and inter-sister interactions at the same site can be observed as an enriched dot, an example is shown in supplemental figure 10.

As described above, we observe that on average $P_{inter}(s)$ is smaller than $P_{intra}(s)$ for loci separated for up to 35kb, while they converge for loci separated by larger distances. When we plot these two parameters for interactions anchored at all centromeres (figure 3d, solid lines) and compare it to interactions along chromosome arms (figure 3d, dashed lines), we observe near identical interaction frequencies of inter- and intra-sister interactions at all distances. This indicates that at centromeres sister chromatids are very precisely aligned so that the two sister centromeres act as a single entity, while loci along arms sister loci are much more loosely aligned. This can be explained at least in part because cohesin binding sites are spaced more closely at centromeres than along arms: spacing between ChIP-seq cohesin peaks is about 5kb around centromeres but along arms the spacing is ~ 15kb. Second, the cohesin ChIP peaks at centromeres are much higher than along arms, indicating that many or most cells in the population will have cohesin bound to those sites, while along arms there may be more cell-to-cell variation in cohesin site occupancy. In summary, a the precise and close alignment at centromeres may be the result of closely spaced and frequently bound cohesin sites at centromeric regions during mitosis[22,24].

Combined these observations lead us to propose that around centromeres interactions between and along sister chromatids are organized differently from those along chromosome arms. We note that we cannot differentiate between a situation where extruding loops are actively formed by loop extruding factors such

as extruding cohesin and condensin or if the observation of loops in peri-centromeric regions is a result of offset cohesive cohesin binding, which would passively expulse a loop on one sister chromatid (figure 3e, marked loops by asterisks). Both scenarios will be observed as line-like features in Hi-C heatmaps.

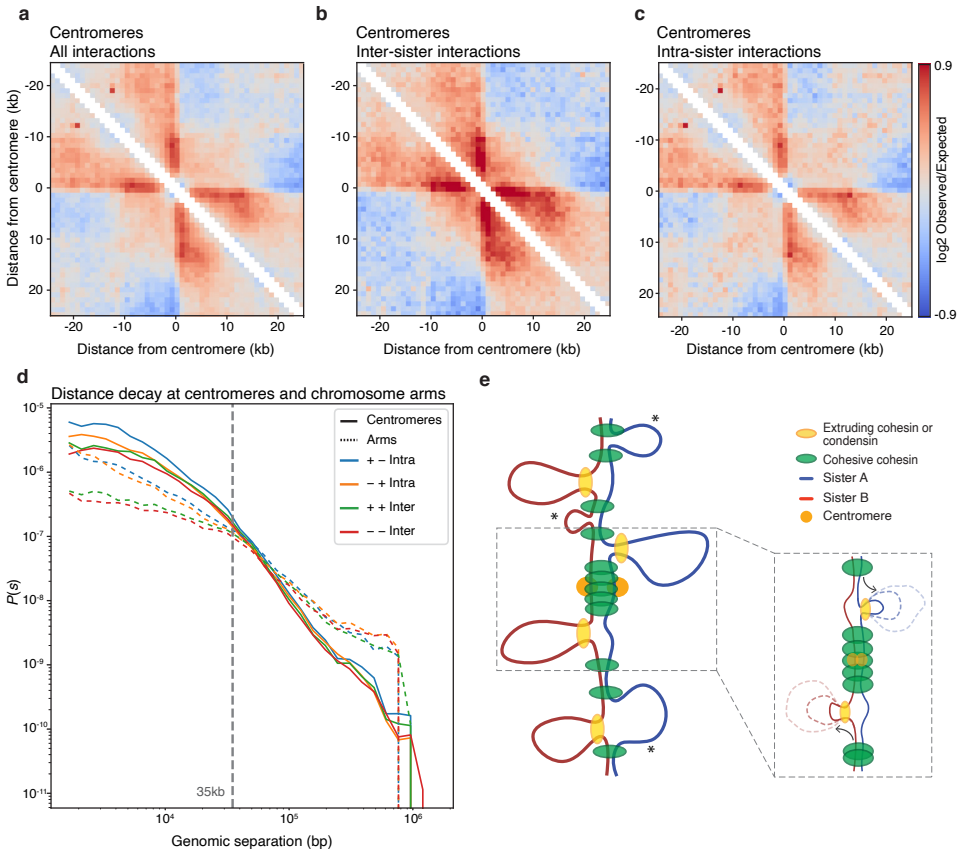


Figure 3 – SisterC data show that centromeres are more precisely and closely aligned than loci along chromosome arms. (**a-c**) Pile up plot on all yeast centromeres at 1kb resolution of all interactions (**a**), inter-sister interactions (**b**) and intra-sister interactions (**c**). (**d**) Distance decay of interactions anchored on centromeres shows inter- and intra-sister interactions are very similar for all genomic distance, while interactions along the chromosome arms show reduced inter-sister interaction for loci separated by up to 35kb (dashed lines; dashed vertical line indicates the 35 kb offset). (**e**) Proposed model of centromeric sister chromatid conformation where closely spaced binding by cohesin and condensin molecules mediate tight and aligned interactions of the sister chromatids at the centromere. Away from the centromere a less dense array of interactions between cohesin sites result in an offset of inter-sister interactions. Note that loops could be formed actively by extruding cohesin or passively by loop expulsion (loops marked by asterisks). As centromeres function as a fixed loop anchor, intra-sister loops can only be formed emerging from one direction (see insert), as observed in a-c.

Inter-sister and intra-sister interactions along arms are mediated by independent cohesin complexes that act at different genomic distances

Cohesin mediated interactions along chromosome arms have been identified by Hi-C before[7,25,26], however prior to SisterC it has not been possible to differentiate between interactions between and along sister chromatids. When we aggregate and plot all inter-sister interactions (figure 4a) and intra-sister interactions (figure 4b) at and around individual cohesin sites (far left panels), we see that cohesin sites preferentially interact with sites located at least 5kb away on either side. Interestingly, visual inspection of these aggregate interaction maps reveals that inter-sister interactions at these cohesin sites occur at shorter distances than intra-sister interactions. This becomes more evident when we calculate the difference between the two aggregate interaction maps (figure 4c). We observe in this difference plot an enrichment for inter-sister interactions over intra-sister interaction for cohesin sites separated by less than 20kb. Figure 4d illustrates this difference in another way by plotting the enrichment of inter and intra-sister signal over expected as a function of genomic distance from cohesin binding sites (figure 4d). Inter-sister interactions at cohesin sites preferentially occur at a distance up to 20kb, whereas intra-sister interactions become more abundant at distances larger than 20kb. This can also be seen at an individual cohesin site (supplemental figure 11).

To explore and quantify these differences between inter-sister and intra-sister interactions in more detail, we analyzed pairwise cohesin-cohesin site interactions at different genomic distances. Interactions of cohesin sites within 10kb from each other are preferentially inter-sister interactions (figure 4a-c, second panel from left). This difference is also observed when cohesin sites are separated by 10-20kb (figure 4a-c, middle panel). Interestingly, this preference switches for pairwise cohesin site interactions for sites separated by more than 20kb: for cohesin sites separated by 20 to 35kb we observe a slight preference for intra-sister interactions (figure 4a-c, second panel from right). This difference for intra-sister interactions becomes much more prominent for pairwise cohesin interactions for sites separated by 35 to 50kb (figure 4a-c, far right panel). Thus, cohesive cohesin enables interactions between sites at sister chromatids that are separated by 5 to 25kb, whereas extruding cohesin generates loops along sister chromatids that can be as large as 50kb.

Above we described that findings in the SisterC data suggest that a given cohesin binding site can be engaged in an inter-sister interaction in one cell and an intra-sister interaction in another cell. To investigate this further, we leveraged the fact that these two types of interactions occur at different length scales. Specifically, we ranked SisterC signal of pairs of cohesin binding sites at different distances on

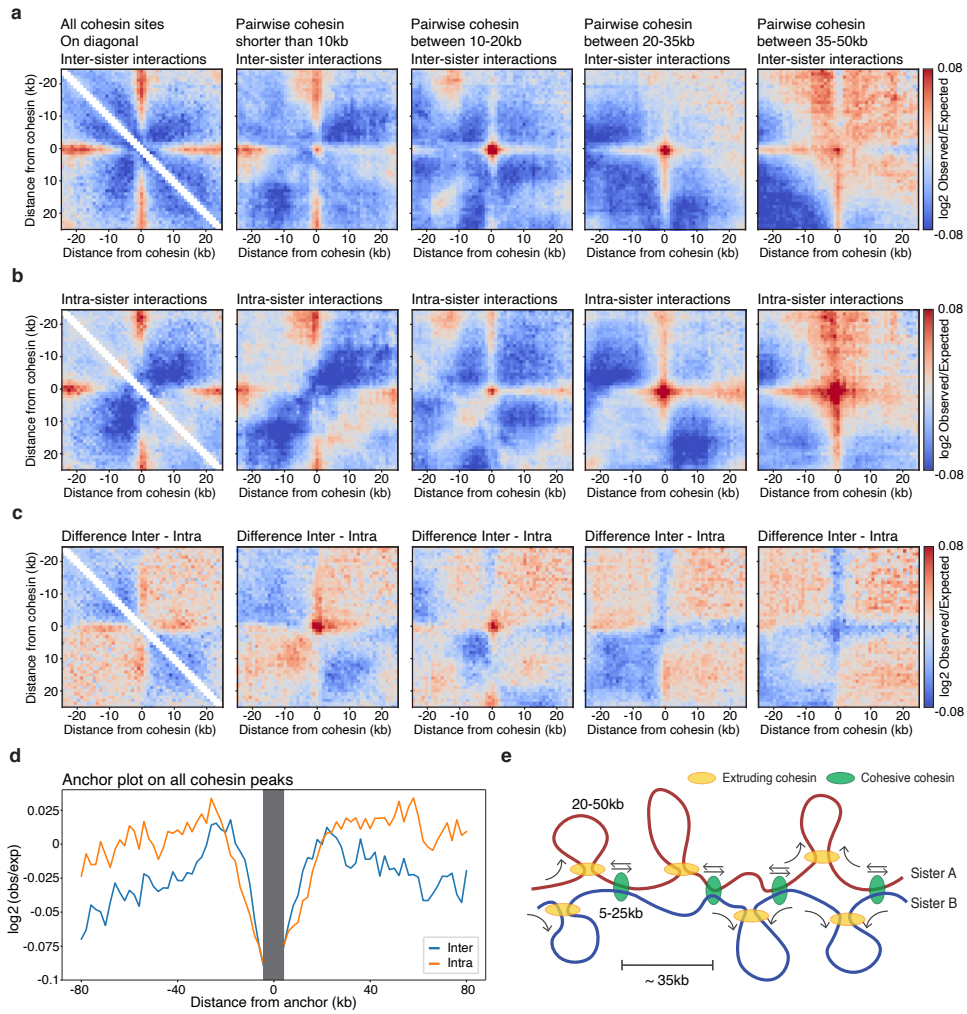


Figure 4 – SisterC shows that cohesive cohesin mediates interactions between sister chromatids at shorter genomic distances than loops formed within sister chromatids by extruding cohesin. (**a-b**) Aggregate pile up plot of inter-sister interactions (**a**), intra-sister interactions (**b**) and the difference between inter-sister and intra-sister interactions (**c**) of all cohesin sites (left panel), of all pairwise cohesin interactions for pairs separated by less than 10kb (second to left panel), all pairwise interactions for sites separated by 10-20kb (middle panel), all pairwise interactions between cohesin sites separated by 20-35kb (second to right panel), and for cohesin sites separated by 35-50kb (far right panel). (**d**) Anchor plot of all cohesin binding sites show that inter-sister interactions are preferentially formed at distances shorter than 30kb, whereas intra-sister interactions at cohesin sites predominantly interact with sites further than 25kb away. (**e**) Proposed model of intra-sister interactions formed by extruding cohesin and inter-sister interactions formed by cohesive cohesin. Loops of different sizes are made within sisters by extruding cohesin with loop sizes ranging from 10-50 kb. Inter-sister interactions are mediated by cohesive cohesin, are spaced by ~35 kb apart, and occur between sites that can be off-set by 5-25 kb.

their intensity (supplemental figure 12). We identified 284 cohesin binding sites that are engaged in the strongest inter-sister interactions (top 10 percent) with other cohesin binding sites that are located 10 to 20kb away. When we explore intra-sister interactions for this set of cohesin binding sites, we observed that these cohesin binding site pairs also display enriched intra-sister interactions 10 to 20kb away, albeit at lower frequency compared to their inter-sister cohesin-cohesin binding site interactions (supplemental figure 12a-c). Similarly, we identified 284 cohesin binding sites that mediate the strongest intra-sister interactions between cohesin sites separated by 35 to 50kb. These cohesin-cohesin binding site pairs also show enriched in inter-sister interactions, although at a lower frequency compared to intra-sister interaction signal (supplemental figure 12d-f). This suggests that there are cohesin sites that mediate both strong intra-sister interaction and strong inter-sister interactions. Further, we find an overlap of 65 cohesin binding sites) between cohesin binding sites that mediate strong inter-sister interactions with distal cohesin binding sites located at 10 to 20kb distance as well as strong intra-sister interactions with cohesin binding sites located at 35 to 50kb distance (supplemental figure 12g). This provides further support that a given cohesin binding site can engage in either inter-sister interactions or intra-sister interactions in different cells in the population.

Discussion

Here we describe SisterC, a chromosome conformation capture technique that allows for detection of interactions between and along sister chromatids separately. SisterC leverages BrdU incorporation and single strand breaks induced by UV/Hoechst treatment, to assign Hi-C interactions as inter-sister or intra-sister interactions based on read orientation after sequencing.

First, SisterC reveals the extent to which sister chromatids are aligned. At centromeres the alignment is rather precise, possibly as a result of high density of cohesin binding sites that are engaged in inter-sister interactions at centromeric regions in every cell. The alignment is more loose along chromosome arms with inter-sister interactions spaced every 35kb on average. Analysis of cohesin binding sites in cell populations shows that cohesin binding sites occur every 10 to 15kb. Combined with our data, this observation suggests that not every cohesin site will be bound in every cell or not every cohesin site will be engaged in inter-sister interactions. The alignment of sisters observed here resembles the pairing of homologues observed in *Drosophila*[27]. In that study, a similar analysis and comparison of $P_{inter}(s)$ and $P_{intra}(s)$ was used to infer the extent of alignment of homologues and regional variation in the precision of alignment of chromosomes was observed along the length of the chromosomes. Although we see a clear difference in alignment between centromeric

regions and chromosome arms in mitotic yeast sister chromatids, we do not observe different degrees of alignment along chromosome arms.

Second, we observe that inter-sister interactions likely mediated by cohesive cohesin occur mostly between cohesin binding sites separated by less than 25kb. This distance is smaller than the distance between inter-sister interaction sites, which we estimated to occur every 35kb on average (see above). One explanation for this can be that inter-sister interactions get established during S-phase every 35 kb or so and relatively close to the replication fork that generates the sister chromatid pair[28–30]. Possibly, inter-sister interactions are initially very precise, but can possibly move to the closest cohesin binding site at convergent gene pairs on both sister chromatids[24,31,32], producing a spacing that can be up to 25kb. Along these lines, we explored whether this offset of inter-sister chromatid interactions would differ from genome-wide average at and around origins of replications. However, we find no clear differences in sister chromatid interactions along or between sister chromatids near origins of replication (supplemental figure 13). Interestingly, we do see a clear boundary at origins of replication in our control G1 Hi-C libraries (supplemental figure 13d).

Third, we observe that intra-sister interactions, possibly mediated by extruding cohesin, form larger loops ranging from 25 to 50kb. These cohesin-mediated extruded loops are established during G2/M-phase[7]. Interestingly, previous polymer simulation showed that yeast Hi-C data from mitotic cells is consistent with the formation of dynamic loops of around 35kb in size, which cover about 35% of the genome[7]. This is in agreement with our SisterC observations.

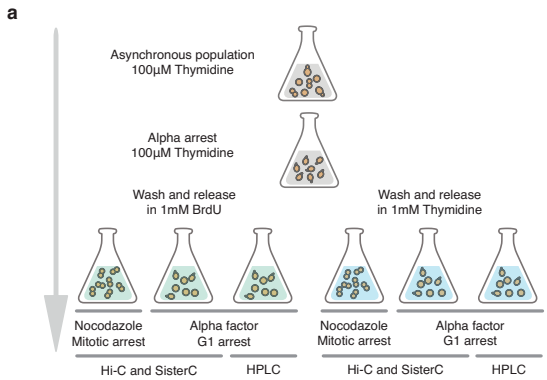
Cohesive cohesin and extruding cohesin are preferentially interacting at different genomic distances, mediate interactions independent from each other and are loaded and established in different phases in the cell cycle[7,33]. This leads us to propose that these cohesin complexes are distinct, possibly having different subunit compositions or modifications. For instance, cohesin can be bound by either Scc2/4 or Pds5[34–36], leading to cohesin complexes with different properties. Further, the acetylation status of a cohesin complex mediated by Eco1 is particularly important for establishment of cohesion[37,38]. Where in yeast both inter-sister interactions and intra-sister interactions in mitosis are mediated by cohesin complexes, in vertebrates these are formed by two different protein complexes: cohesin establishes sister chromatid cohesion and condensin I and II mediate intra-sister looping formation to compact chromosomes[7,39]. Additionally, it is important to note that in yeast a given cohesin binding site can be involved in both an inter-sister interaction and an intra-sister interaction, although most likely not occurring in the same cell at a given time. Due to a low cohesin binding density in yeast and absence of sequence specificity

of cohesin complexes, there will be large cell-to-cell variation in which cohesin sites will be bound by either cohesive or extruding cohesins.

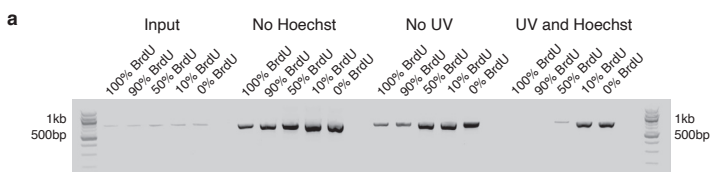
In the current SisterC procedure, selective depletion of DNA strands containing BrdU is not complete. Therefore, there is some level of cross contamination of inter- and intra-sister interactions. This level of contamination can be estimated by analyzing SisterC libraries of cells in G1-phase after one round of S-phase in the presence of BrdU. The reason why selective depletion is not complete is most likely related by the relatively low efficiency of DNA breakage after UV/Hoechst treatment. More efficient strand depletion using alternative approaches, such as enzymatic or chemical depletion of T-analogues, could result in more sensitive differentiation of inter-sister and intra-sister interactions. Importantly we did not identify particular types of molecules that are resistant to selective depletion. For instance, A-T content of the genomic site (supplemental figure 14a-b), distance from a digestion site (supplemental figure 14c-d) or regions near replication origins (supplemental figure 14e-f) do not affect assignment of interactions as being inter-sister or intra-sister. BrdU incorporation as measured by HPLC is near to complete in the newly replicated strand. In order to achieve this in budding yeast, a mutant strain was used. However, we note that in order to apply SisterC in mammalian cell, no special cell lines are required as mammalian cell lines are able to take up and incorporate BrdU without experimental manipulations[14,40], which makes application of SisterC to any mammalian cell type straightforward. Despite incomplete depletion of DNA strands containing BrdU, SisterC is able to identify interactions that are enriched in the inter-sister dataset compared to the intra-sister dataset and vice versa. This allowed the detection of the two distinct roles of cohesin; cohesive cohesin and extruding cohesin. This independent behavior of cohesin complexes was also observed in mammalian cells by another novel Hi-C based technique that can distinguish interactions between and along sister chromatids, scsHi-C [41].

SisterC allows the study of the significant topological challenge each cell faces during the cell cycle: the concordant chromosome compaction and sister chromatid separation during mitosis, particularly prophase. This process has been difficult to study by conventional Hi-C, because of its inability to distinguish between inter- and intra-sister interactions. We believe SisterC will have a broad applicability in different model organisms and during different phases of the cell cycle from late S-phase to mitosis.

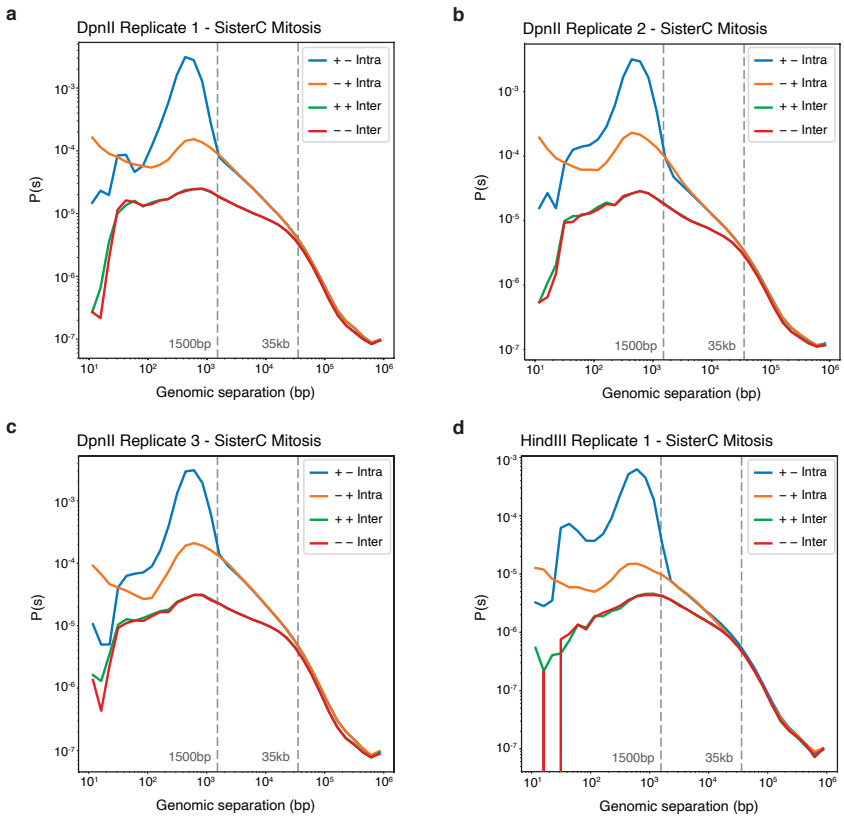
Extended data figures



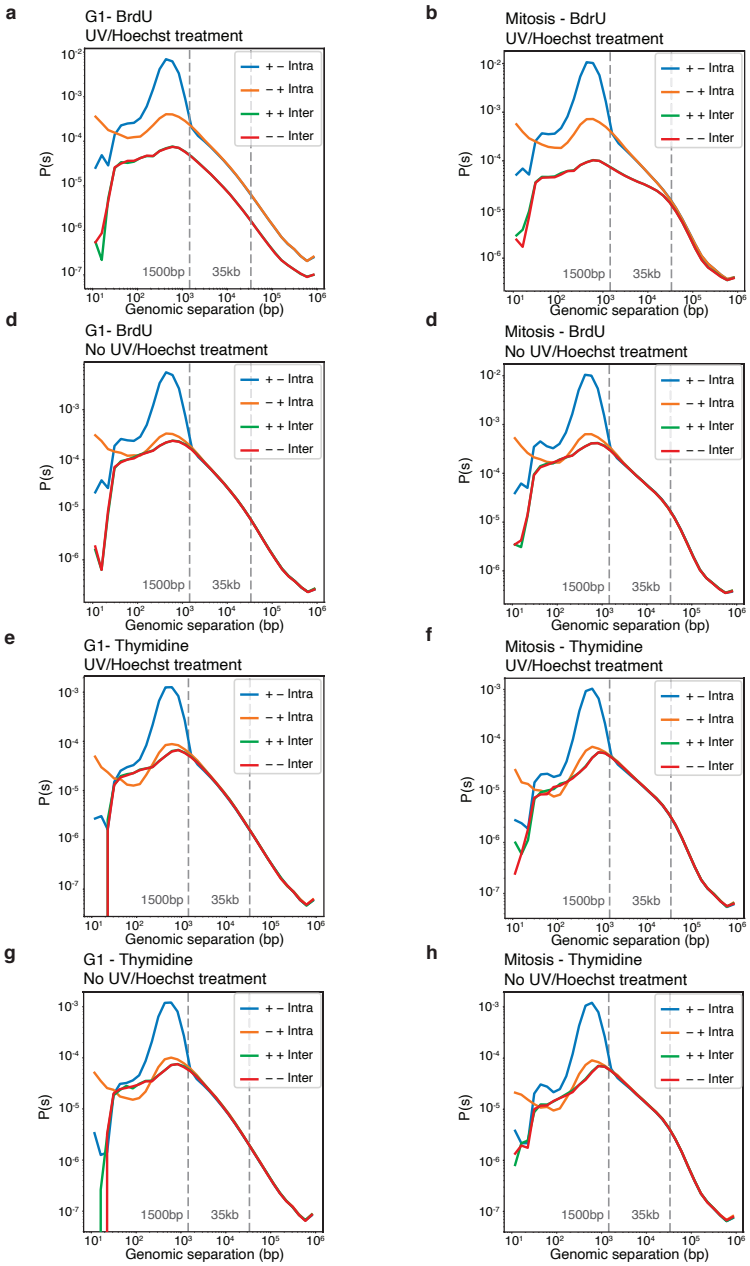
Extended Data Fig. 1 - Outline of SisterC yeast culture conditions and flow cytometry analysis. (**a**) Asynchronous yeast cultures are cultured and synchronized in late G1 using alpha factor. Cells are released in media containing BrdU or Thymidine, followed by an arrest in mitosis using nocodazole or G1 using a second alpha factor arrest. Cells are harvested and prepared for Hi-C or SisterC library production or processed for BrdU detection using HPLC. (**b**) Flow cytometry analysis of cell cycle profile and BrdU incorporation of harvested yeast cultures for preparation of SisterC libraries.



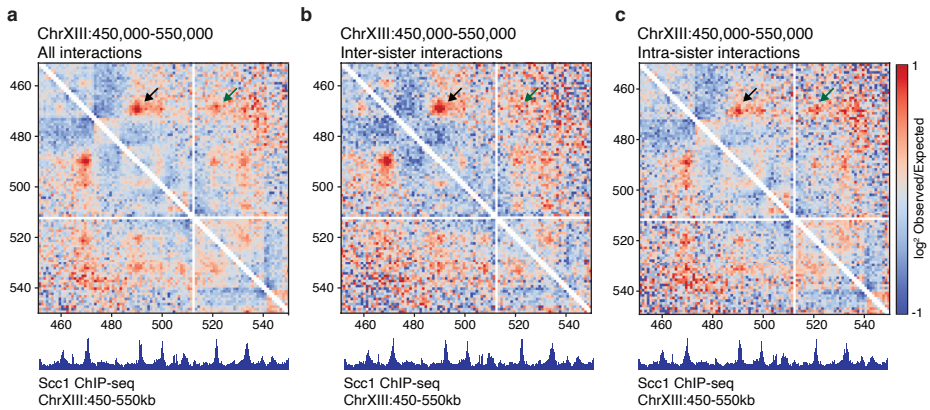
Extended Data Fig. 2 - Depletion of BrdU containing DNA molecules by PCR. DNA fragments were amplified in presence of 0%, 10%, 50%, 90% and 100% BrdU to allow for incorporation in both strands (first 5 lanes). This was followed by treatment of UV only (second 5 lanes), Hoechst only (third 5 lanes) or treatment with both UV and Hoechst (last 5 lanes). Fragments containing more than 10% BrdU did not get amplified after UV/Hoechst treatment. This experiment was repeated twice with highly similar results.



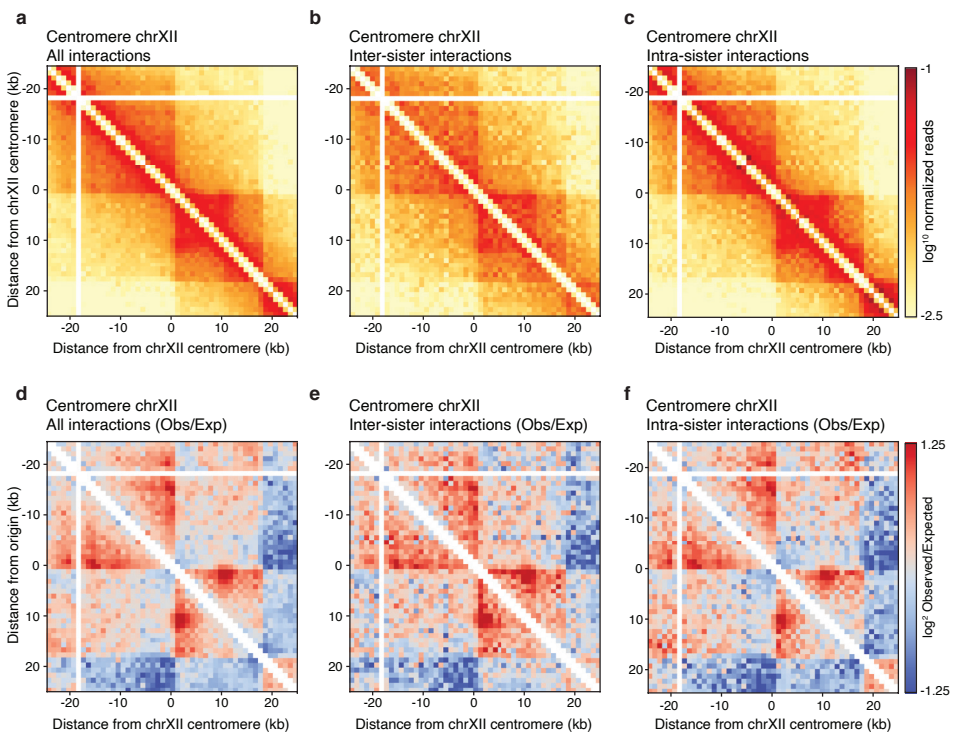
Extended Data Fig. 3 - Distance decay plots of all SisterC replicates. Distance decay plots of all SisterC mitotic libraries: DpnII replicate 1 (**a**), DpnII replicate 2 (**b**), DpnII replicate 3 (**c**), and HindIII replicate 1 (**d**)



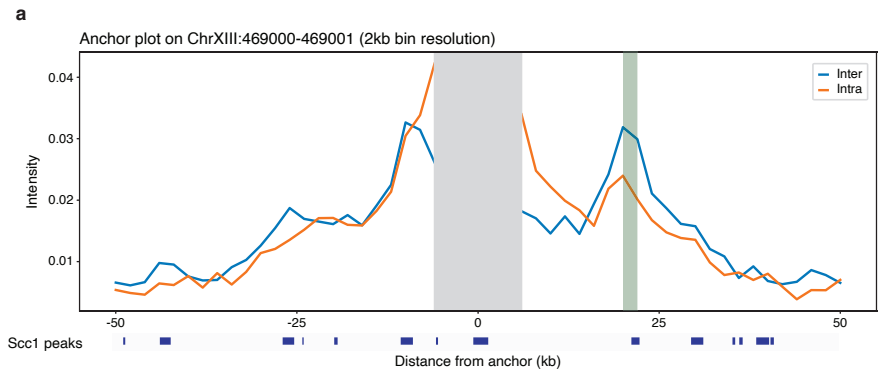
Extended Data Fig. 4 - Distance decay plots of all SisterC control experiments. Distance decay plots of all SisterC control libraries: (**a-b**) G1 (a) and mitotic (b) arrested cells grown in BrdU and treated with UV/Hoechst. (**c-d**) G1 (c) and mitotic (d) arrested cells grown in BrdU, not treated with UV/Hoechst. (**e-f**) G1 (e) and mitotic (f) arrested cells grown in Thymidine, treated with UV/Hoechst. (**g-h**) G1 (g) and mitotic (h) arrested cells grown in Thymidine, not treated with UV/Hoechst.



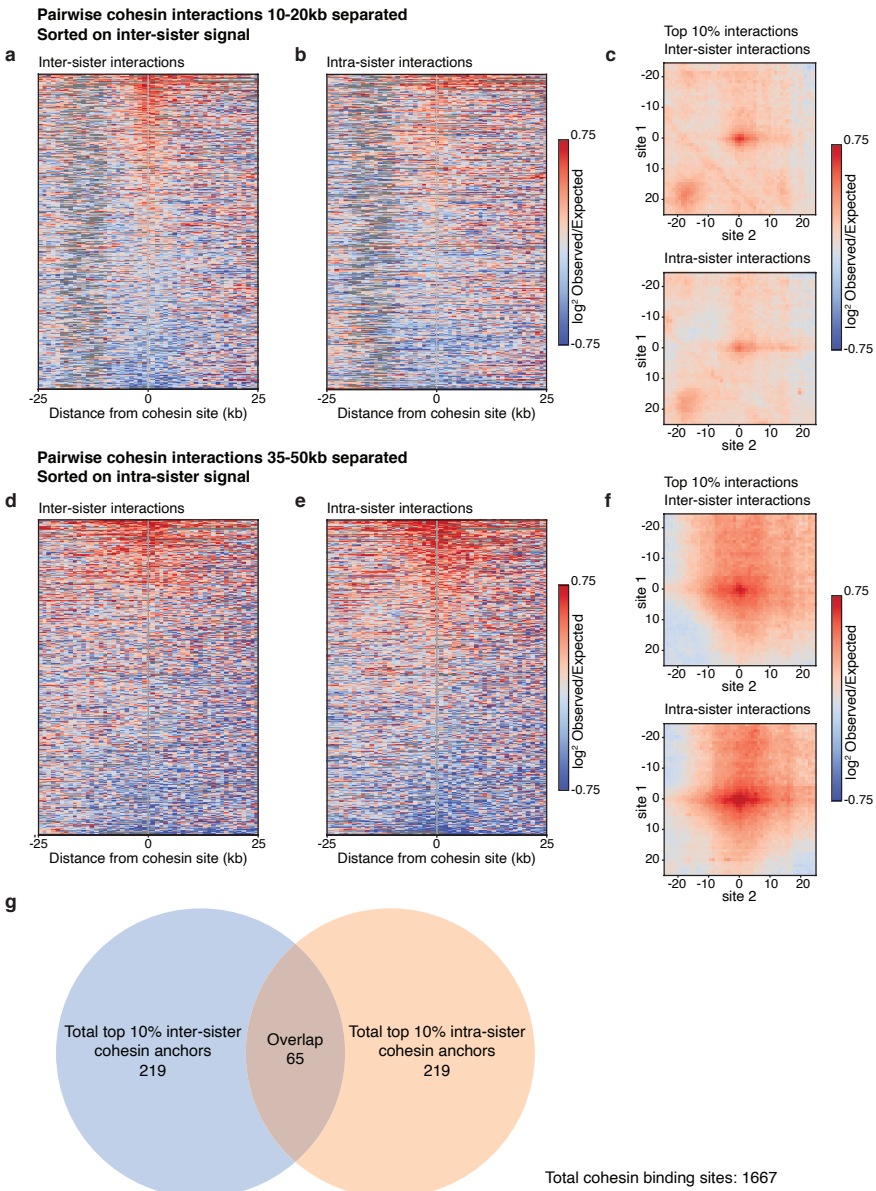
Extended Data Fig. 5 - Observed/expected interaction heatmap of chrXIII:450,000–550,000. Log2 observed over expected interaction frequency on region ChrXIII:450,000–550,000 (as shown in main Fig. 2) for all interactions (**a**) inter-sister interactions (**b**) and intra-sister interactions (**c**).



Extended Data Fig. 6 - Centromeric region of chrXII. (**a-c**) Interaction heatmaps of 50 kb window around the centromeric region of chrXII, for all interactions (a), inter-sister interactions (b) and intra-sister interactions (c). (**d-f**) Log2 observed over expected interaction heatmaps of all interactions (d), inter-sister interactions (e) and intra-sister interactions (f) on the centromeric region on chrXII.

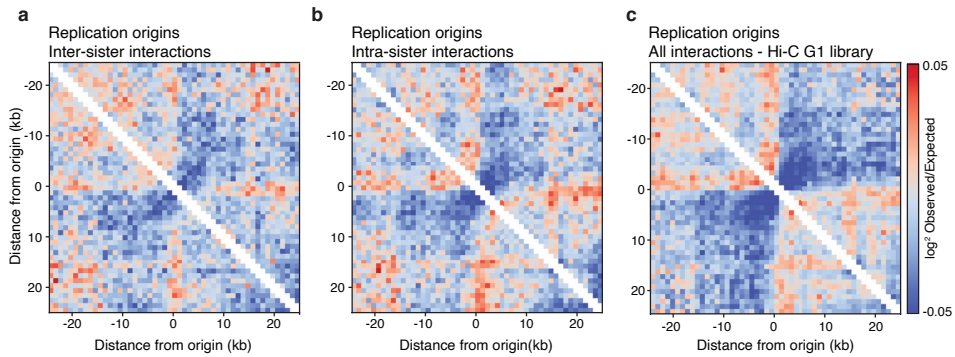


Extended Data Fig. 7 - Anchor plot on ChrXIII:469000. Interaction frequency of inter-sister and intra-sister interactions anchored on chrXIII:469,000 shows higher frequency of inter-sister interactions at 25 kb distance from the anchor (highlighted in green). Annotated below the plot are the positions of Scc1 peaks. Source data

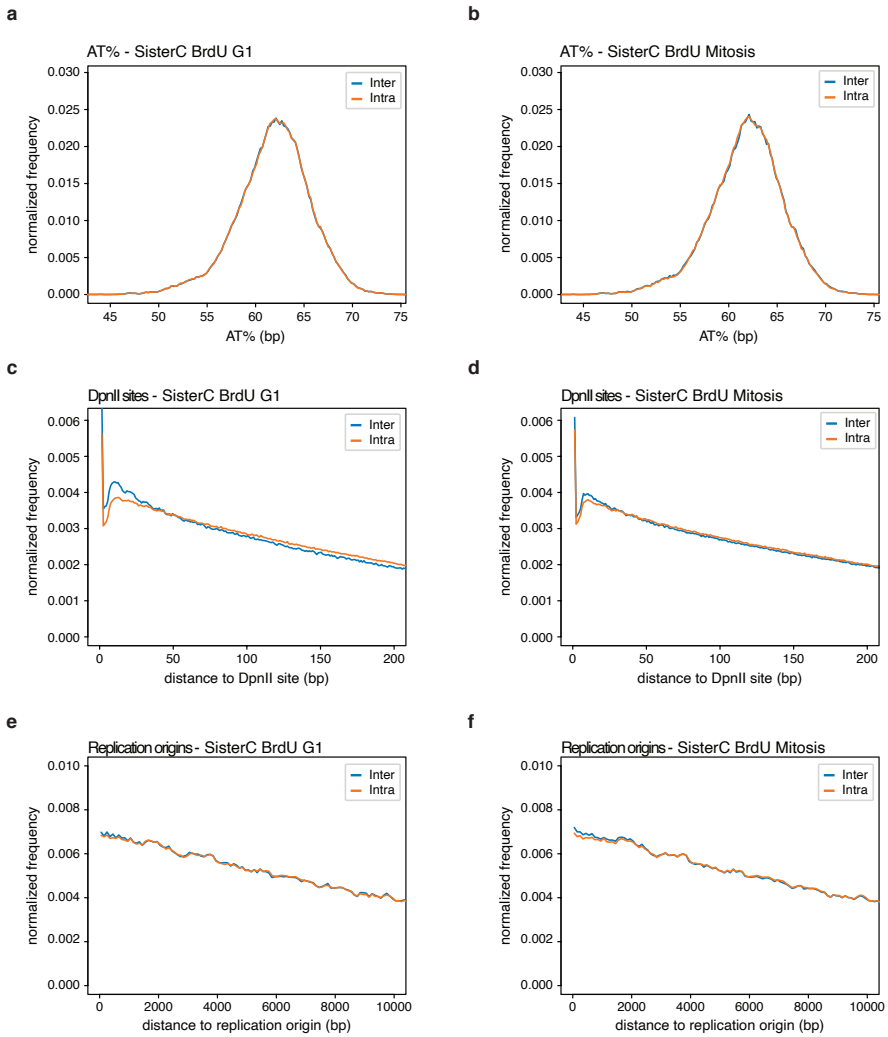


Extended Data Fig. 8 - Ranking of SisterC signal on pairs of cohesin binding sites at different genomic distances. (**a, b**) Inter-sister interaction (a) and intra-sister interaction (b) frequency of pairs of cohesin binding sites separated by 10 to 20 kb were ranked by inter-sister interaction intensity for 10 kb window around the cohesin binding site. (**c**) Interaction pile up plots of inter-sister and intra-sister interactions of the top 10 percent sites that were ranked in (a). (**d-e**) Inter-sister interaction (d) and intra-sister interaction (e) frequency of pairs of cohesin binding sites separated by 35 to 50 kb were ranked by intra-sister interaction intensity for 10 kb window around the cohesin binding site. (**f**) Interaction pile up plots of inter-sister

and intra-sister interactions of the top 10 percent sites that were ranked in (e). (**g**) 284 cohesin sites that anchor the top 10% inter-sister interacting cohesin pairs at 10 to 20 kb distance were identified as well as again 284 cohesin sites that anchor the top 10% intra-sister interacting cohesin pairs at 35 to 50 kb distance. 65 cohesin sites are found to mediate both top 10% inter-sister cohesin-cohesin interactions and top 10% intra-sister cohesin-cohesin interactions.

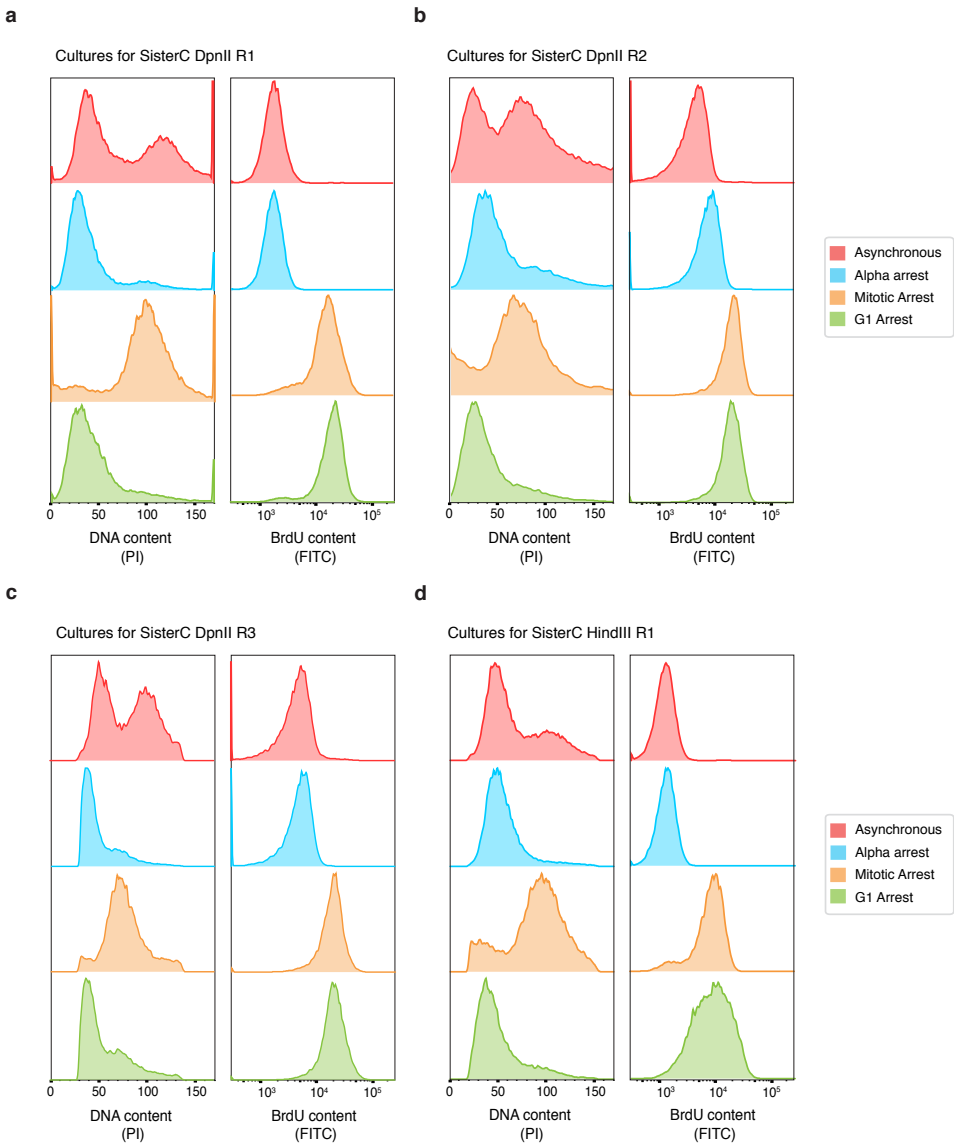


Extended Data Fig. 9 - SisterC pile up plots of origins of replication. (**a, b**) Mitotic SisterC inter-sister interactions (a) and intra-sister interactions (b) plotted on replication origins. (**c**) Pile up plot of Hi-C data of G1 alpha arrested cells on replication origins.

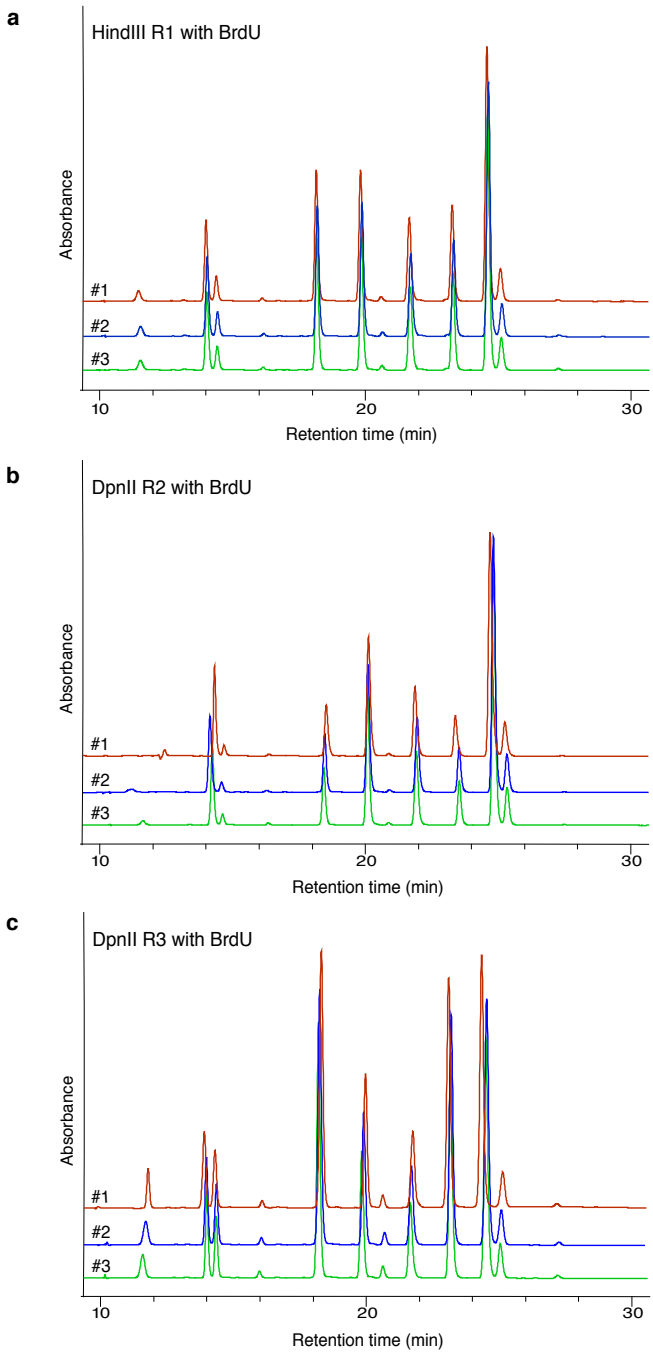


Extended Data Fig. 10 - Efficiency of SisterC inter- and intra-sister interaction assignment. (**a, b**) Normalized frequency SisterC G1 (a) and mitotic (b) inter-sister and intra-sister interactions as a function of AT percentage. (**c, d**) Normalized frequency SisterC G1 (c) and mitotic (d) inter-sister and intra-sister interactions as a function of distance to DpnII digestion site. (**e, f**) Normalized frequency SisterC G1 (e) and mitotic (f) inter-sister and intra-sister interactions as a function of distance to origins of replication.

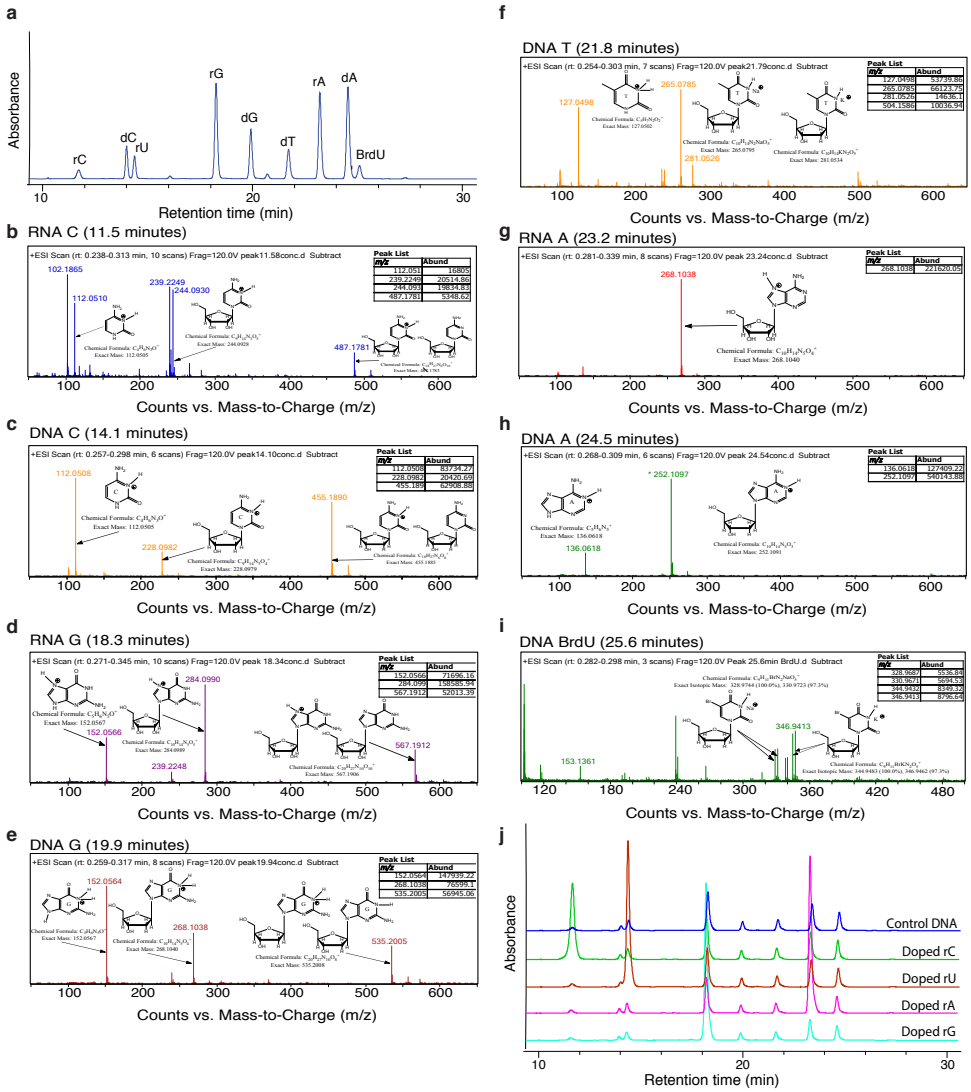
Supplemental figures and tables



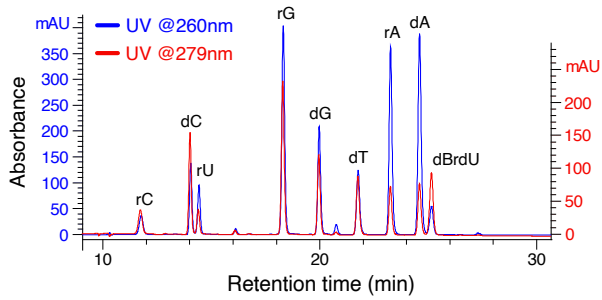
Supplemental Figure S1 - **FACS data of all SisterC experiments**. Flow cytometry data detecting cell cycle profile and BrdU content of cultures grown for SisterC DpnII replicate 1 (**a**), SisterC DpnII replicate 2 (**b**), SisterC DpnII replicate 3 (**c**) and HindIII R1 (**d**).



Supplemental Figure S2 - HPLC Nucleoside digestion data of all replicates. Stacked HPLC traces from 3 repeats of each SisterC library HindIII R1 (**a**), DpnII R2 (**b**) and DpnII R3 (**c**).



Supplemental Figure S3 - Definitive assignment of nucleoside peaks. (**a**) HPLC trace showing peaks from a digested, BrdU containing sample of DNA; provided for reference of the peaks assigned in the following sections. (**b-i**) Mass spectrometry analysis of individual peaks shown in panel a. No diagnostic m/z peaks were found for the 14.4 minute peak, however RNA U doping (brown trace, panel j) was used to unambiguously assign this peak as RNA U. RNA A has the same chemical formula and mass as dG (218.1040 g/mol). Both the fragmentation pattern of the DNA G peak (panel e) and the RNA A doping experiment (magenta trace panel j) were used to confirm the assignment of the 23.2 minute peak as RNA A (panel g). (**j**) Stacked HPLC runs of digested, non-BrdU-treated control DNA doped with eachribonucleoside individually (2 uL of 5 mM solution of rU, rC, rA or rG was added to the digested DNA immediately before HPLC analysis).



Supplemental Figure S4 - HPLC Nucleoside Digestion data at 260nm and 279nm. Overlaid HPLC traces at two different UV detector wavelengths from a single injection of a digested, BrdU-treated DNA sample. Blue trace shows absorption at 260nm and red trace at 279nm. The max of BrdU is 279nm.

Replicate	Release in	Synchronized	UV/Hoechst treatment	Reads total	Total Mapped		Total valid reads (no duplicates)		Cis all interactions		Cis all >1500bp		Cis intra-sister >1500bp		Cis inter-sister >1500bp		Trans all		Trans intra-sister		Trans inter-sister	
				#	#	% of total	#	% of	#	% of valid	#	% of cis	#	% of cis	#	% of cis	#	% of valid	#	% of trans	#	% of trans
DpnII R1	BrdU	Mitosis	Yes	2.27E+07	1.57E+07	69.11	1.42E+07	90.71	1.04E+07	73.32	5.12E+06	49.16	3.20E+06	62.42	1.92E+06	37.58	3.79E+06	26.68	1.90E+06	50.00	1.90E+06	50.00
	BrdU	Mitosis	No	2.28E+07	1.54E+07	67.33	1.57E+07	101.97	1.11E+07	70.64	6.78E+06	61.30	3.40E+06	50.11	3.38E+06	49.89	4.60E+06	29.36	2.30E+06	50.02	2.30E+06	49.98
	BrdU	G1	Yes	2.27E+07	1.57E+07	69.11	1.46E+07	92.99	9.50E+06	65.20	3.96E+06	41.71	3.12E+06	78.90	8.36E+05	21.10	5.07E+06	34.80	2.53E+06	49.99	2.53E+06	50.01
	BrdU	G1	No	2.28E+07	1.54E+07	67.33	1.46E+07	95.10	8.65E+06	59.19	4.60E+06	53.14	2.31E+06	50.24	2.29E+06	49.76	5.96E+06	40.81	2.98E+06	49.99	2.98E+06	50.01
	Thymidine	Mitosis	Yes	2.46E+07	1.66E+07	67.55	1.68E+07	101.09	1.27E+07	75.30	8.29E+06	65.55	4.16E+06	50.15	4.13E+06	49.85	4.15E+06	24.70	2.08E+06	50.03	2.07E+06	49.97
	Thymidine	Mitosis	No	2.43E+07	1.70E+07	69.93	1.60E+07	94.33	1.21E+07	75.21	8.29E+06	65.55	4.16E+06	50.15	4.13E+06	49.85	3.98E+06	24.79	1.99E+06	49.99	1.99E+06	50.01
	Thymidine	G1	Yes	2.46E+07	1.66E+07	67.55	1.56E+07	93.68	8.78E+06	56.42	4.50E+06	51.20	2.26E+06	50.16	2.24E+06	49.84	6.78E+06	43.58	3.39E+06	50.00	3.39E+06	50.00
	Thymidine	G1	No	2.04E+07	1.37E+07	67.15	1.30E+07	94.82	7.19E+06	55.36	3.82E+06	53.08	1.92E+06	50.26	1.90E+06	49.74	5.80E+06	44.64	2.90E+06	50.03	2.90E+06	49.97
DpnII R2	BrdU	Mitosis	Yes	1.63E+08	1.13E+08	69.37	7.79E+07	69.03	5.04E+07	64.67	2.22E+07	44.07	1.37E+07	61.86	8.47E+06	38.14	2.75E+07	35.33	1.38E+07	50.01	1.38E+07	49.99
	BrdU	Mitosis	No	1.64E+08	1.14E+08	69.32	8.88E+07	78.21	5.23E+07	58.84	2.95E+07	56.47	1.48E+07	50.31	1.47E+07	49.69	3.66E+07	41.16	1.83E+07	49.99	1.83E+07	50.01
	BrdU	G1	Yes	1.16E+08	8.11E+07	69.87	5.16E+07	63.64	3.19E+07	61.90	1.03E+07	32.30	8.24E+06	79.90	2.07E+06	20.10	1.97E+07	38.10	9.83E+06	50.00	9.83E+06	50.00
	BrdU	G1	No	1.23E+08	8.43E+07	68.49	6.35E+07	75.40	3.23E+07	50.83	1.48E+07	45.86	7.46E+06	50.36	7.35E+06	49.64	3.12E+07	49.17	1.56E+07	49.99	1.56E+07	50.01
	Thymidine	Mitosis	Yes	5.50E+07	3.72E+07	67.74	2.89E+07	77.52	1.67E+07	57.74	9.51E+06	57.04	4.80E+06	50.47	4.71E+06	49.53	1.22E+07	42.26	6.10E+06	50.00	6.10E+06	50.00
	Thymidine	Mitosis	No	5.45E+07	3.73E+07	68.46	2.91E+07	78.05	1.65E+07	56.58	9.97E+06	60.53	5.03E+06	50.49	4.93E+06	49.51	1.26E+07	43.42	6.32E+06	50.01	6.32E+06	49.99
	Thymidine	G1	Yes	7.33E+07	5.03E+07	68.61	3.27E+07	65.12	1.57E+07	47.86	9.51E+06	57.04	4.80E+06	50.47	4.71E+06	49.53	1.71E+07	52.14	8.53E+06	50.01	8.53E+06	49.99
	Thymidine	G1	No	5.90E+07	4.04E+07	68.47	3.22E+07	79.69	1.56E+07	48.42	9.97E+06	60.53	5.03E+06	50.49	4.93E+06	49.51	1.66E+07	51.58	8.30E+06	49.99	8.30E+06	50.01
DpnII R3	BrdU	Mitosis	Yes	1.88E+08	1.30E+08	69.13	7.57E+07	58.18	5.97E+07	78.84	3.06E+07	51.23	2.02E+07	65.94	1.04E+07	34.06	1.60E+07	21.16	8.01E+06	50.01	8.01E+06	49.99
	BrdU	Mitosis	No	1.86E+08	1.30E+08	69.80	1.07E+08	82.59	8.16E+07	76.12	5.15E+07	63.14	2.58E+07	50.15	2.57E+07	49.85	2.56E+07	23.88	1.28E+07	50.01	1.28E+07	50.01
	BrdU	G1	Yes	1.54E+08	1.06E+08	68.46	8.52E+07	80.76	5.75E+07	67.52	2.40E+07	41.70	1.90E+07	79.02	5.03E+06	20.98	2.77E+07	32.48	1.38E+07	50.01	1.38E+07	49.99
	BrdU	G1	No	1.89E+08	1.28E+08	67.85	1.10E+08	85.55	6.72E+07	61.25	3.58E+07	53.24	1.79E+07	50.19	1.78E+07	49.81	4.25E+07	38.75	2.12E+07	50.01	2.12E+07	49.99
	Thymidine	Mitosis	Yes	8.38E+07	5.57E+07	66.51	4.65E+07	83.55	3.48E+07	74.81	2.26E+07	64.82	1.13E+07	50.09	1.13E+07	49.91	1.17E+07	25.19	5.86E+06	50.01	5.86E+06	49.99
	Thymidine	Mitosis	No	8.92E+07	5.89E+07	66.05	5.05E+07	85.67	3.78E+07	74.86	2.47E+07	65.51	1.24E+07	50.05	1.24E+07	49.95	1.27E+07	25.14	6.34E+06	49.98	6.35E+06	50.02
	Thymidine	G1	Yes	8.94E+07	5.87E+07	65.64	4.95E+07	84.32	2.95E+07	59.67	1.52E+07	51.45	7.64E+06	50.29	7.55E+06	49.71	2.00E+07	40.33	9.98E+06	50.01	9.97E+06	49.99
	Thymidine	G1	No	9.65E+07	6.25E+07	64.77	5.26E+07	84.19	3.12E+07	59.34	1.68E+07	53.68	8.40E+06	50.14	8.36E+06	49.86	2.14E+07	40.66	1.07E+07	49.99	1.07E+07	50.01
HindIII R1	BrdU	Mitosis	Yes	5.81E+07	3.05E+07	52.48	1.08E+07	35.52	9.33E+06	86.20	3.45E+06	36.91	2.00E+06	58.13	1.44E+06	41.87	1.49E+06	13.80	7.47E+05	50.00	7.47E+05	50.00
	BrdU	Mitosis	No	3.90E+07	2.33E+07	59.79	1.63E+07	69.75	1.37E+07	84.20	6.46E+06	47.17	3.22E+06	49.84	3.24E+06	50.16	2.57E+06	15.80	1.29E+06	50.02	1.29E+06	49.98
	BrdU	G1	Yes	8.04E+07	5.55E+07	69.04	2.74E+07	49.37	2.17E+07	79.09	5.63E+06	25.99	4.23E+06	75.15	1.40E+06	24.85	5.73E+06	20.91	2.86E+06	49.99	2.86E+06	50.01
	BrdU	G1	No	4.29E+07	2.92E+07	68.20	2.40E+07	81.95	1.84E+07	76.93	5.68E+06	30.79	2.83E+06	49.81	2.85E+06	50.19	5.53E+06	23.07	2.76E+06	50.00	2.76E+06	50.00
	Thymidine	Mitosis	Yes	4.86E+07	2.66E+07	54.73	1.09E+07	41.10	9.31E+06	85.11	4.11E+06	44.20	2.09E+06	50.81	2.02E+06	49.19	1.63E+06	14.89	8.14E+05	50.00	8.14E+05	50.00
	Thymidine	Mitosis	No	5.34E+07	3.23E+07	60.59	1.31E+07	40.39	1.11E+07	85.03	4.89E+06	44.01	2.53E+06	51.71	2.36E+06	48.29	1.95E+06	14.97	9.78E+05	50.04	9.77E+05	49.96
	Thymidine	G1	Yes	7.64E+07	4.59E+07	60.03	2.28E+07	49.66	1.68E+07	73.72	5.62E+06	33.47	2.89E+06	51.48	2.73E+06	48.52	5.99E+06	26.28	3.00E+06	50.02	2.99E+06	49.98
	Thymidine	G1	No	8.84E+07	5.45E+07	61.66	3.39E+07	62.30	2.48E+07	73.00	8.60E+06	34.69	4.43E+06	51.55	4.16E+06	48.45	9.16E+06	27.00	4.59E+06	50.04	4.58E+06	49.96

Supplemental Table S1 - SisterC and Hi-C mapping statistics across all DpnII and HindIII experiments.

Methods

Yeast synchronization and culture conditions

The YLV11 strain[20] was used for all experiments in this study. For normal growth, cells were cultured in YP media with 2% galactose and 100μM thymidine at 30°C. Prior to synchronization, cultures were diluted to OD ~0.15 and allowed to double. To synchronize cells in late G1, 5 μM alpha factor mating pheromone (zymoresearch #Y1001) was added for 2.5-3 hours until cells started schmoo formation. SisterC DpnII R2 and R3 received a boost of 500uM thymidine or BrdU after 2 hours of alpha factor synchronization. G1 arrested cells were washed 3 times and released in prewarmed media containing 1mM BrdU or Thymidine. For mitotic arrested cells, 1% DMSO was added 15 minutes after release and 10μg/mLnocodazole was added 30 minutes later. Mitotic cells were harvested approximately 4.5 hours after alpha factor release. For G1 arrested cells, the culture was released for 2 hours, followed by a second alpha factor arrest for 3 hours. Cells were washed and pelleted for genomic DNA extraction for HPLC detection and stored at -80°C until further processing. For Hi-C and SisterC, cells were fixed with 3% formaldehyde for 20 minutes at 30°C while in shaker incubator. Fixing was quenched by adding 2.5M glycine for an additional 5 minutes at 30°C. Cells were washed twice in MilliQ and pelleted cells were stored at -80°C till further processing. Throughout synchronization protocol, cells were washed and fixed in 95% ethanol for flow cytometry analysis.

Flow cytometry

Ethanol fixed cells were resuspended and washed with 50mM NaCitrate. After mild sonication, cell walls were degraded with 10 units of zymolyase in PBS for 30 min at 30°C, followed by a lysis using 2M HCl and 0.5% Triton Tx-100 for 30 min at room temperature, followed by 30 min incubation at room temperature in 0.1M NaB_4O_7. Cells were then washed and resuspended in PBS, 1% milk, 0.2% Tween and 1:20 anti-BrdU-FITC (ThermoFisher #11-5071-42) and incubated for 30 min at RT. Cells were again washed and resuspended in PBS, 1% milk, 0.2% Tween, 0.25mg/mL RNase A and 10mg/mL propidium iodide and incubated at 37°C for 30 min. Cells were washed and resonicated before flow cytometry detection using a MACSquant flow cytometer. Data was analyzed using FlowJo software.

Amplification of DNA fragments containing after UV/Hoechst treatment

DNA fragments (686 bp length) were amplified for 15 cycles using DreamTaq (ThermoFisher # EP1701) in presence of 100%, 90%, 50%, 10% or 0% BrdUTP, supplemented with dTTP. Amplified DNA fragments were incubated in TLE with 100ng/uL Hoechst 33342 (ThermoFisher #H3570) for 15 min at room temperature

while protecting from light, followed by UV radiation at 2.7kJ/m^2. Samples were washed 3 times in 30KDa amicon columns (MilliPore # UFC5030BK) and amplified with 10 PCR cycles.

HPLC separation and LCMS analysis

Cells were harvested, pelleted and frozen at -80°C for HPLC analysis from 300mL yeast cultures at OD ~0.3 (approximately 600 million cells). Frozen cells were washed in 1mL spheroplasting buffer and lysed for 10 minutes at 35°C using 0.5% beta-mercaptoethanol and 10ug/mL zymolyase (Zymoresearch # E1005). Cells were washed in 1x NEBuffer 3.1 (NEB #B7203S) and incubated with proteinase K twice for 12 hours and an additional 2 hours at 65°C. DNA was extracted with phenol:chloroform, followed by ethanol precipitation and RNA digestion using RNase A. DNA samples were first washed with milliQ water using a 3 kDa cut-off Amicon filter. Digestion into nucleosides was carried out on 3 μg aliquots using 'nucleoside digestion enzyme mix' from New England Biolabs (NEB#M0649) and incubated overnight at 37°C. Once digested, aliquots from the same sample were pooled, filtered, and rinsed through a 3 kDa Amicon filter with milliQ water. The flow through was concentrated and quantified by A260 for subsequent HPLC injection. Digested deoxynucleosides were resolved using an Agilent 1260 Infinity HPLC with a Synergi C18 4-μm Fusion-RP 80Å 250 x 4.6 mm LC column. Nucleosides were resolved over 35 minutes using an isocratic gradient of 2-22% [95% acetonitrile, 5% 20 mM ammonium acetate pH 4.5] in [20 mM ammonium acetate] at 25°C, using a flow rate of 0.5 mL/min. This was followed by wash steps at higher eluent strengths between runs. Absorbance was recorded at 260 nm and 279 nm. Normal deoxynucleosides were quantitated using HPLC peak areas over three repeats, and published values for molar extinction coefficients at 260 nm [ε260 (M-1 cm-1) adenosine = 15400, cytidine = 7300, guanosine = 11700, thymidine = 8800[42]. The molar extinction coefficients for bromodeoxyuridine at 260 and 279 nm were empirically measured and found to be 9229 and 5003, respectively. For mass spectrometry analysis, deoxynucleosides were first resolved using the HPLC method above, and UV peaks were manually collected in separate vials. These were evaporated to dryness, resuspended in 50 μL milliQ water, and desalted on the same HPLC column using an isocratic gradient of 0-40% [98% acetonitrile, 2% water] in [90% water, 10% acetonitrile] over 15 min at 25°C, using a flow rate of 1 mL/min. The desalted peak was collected, evaporated to dryness, then resuspended in 20 μL milliQ water and analysed by high resolution liquid chromatography mass spectrometry. Mass spectrometry was carried out in positive ESI mode on an Agilent 6530 Accurate-Mass Q-TOF LC/MS linked to a pre-

injection Agilent 1260 infinity HPLC. Samples were run in milliQ water containing 0.1 % formic acid. Mass spectrometry settings to detect nucleosides were as follows; gas temperature 350°C, nebulizer gas rate 45 psig, drying gas 10 L/min, VCap 4000 V, fragmentor voltage 120 V, skimmer voltage 65 V. 5 mM stock solutions of RNA nucleosides A, U, C, and G (ChemGenes) were made up in milliQ water. HPLC injections of digested deoxynucleosides were doped with the addition of 2 μL of each of the RNA nucleosides. Nucleosides were resolved using the standard separation method above.

SisterC and Hi-C library preparation

For each mitotic SisterC or Hi-C library approximately 300 million cells or 100mL culture at OD ~0.3 was used. For each G1 SisterC or Hi-C library this was double, approximately 600 million cells or 100mL culture at OD ~0.6. There were 3 biological replicates produced using DpnII as restriction enzyme and one replicate using HindIII. SisterC and Hi-C were performed according to previously published Hi-C protocol for yeast[25], with several major modifications. Samples were split for Hi-C and SisterC library production before treatment with Hoechst/UV. Briefly, cells were fixed and stored as described above. Crosslinked cells were thawed, washed and resuspended in spheroplasting buffer (1M Sorbitol, 50mM Tris pH 7.5). Cells were lysed by addition of 0.5% beta-mercaptoethanol and 10ug/mL zymolyase (Zymoresearch # E1005) and incubated for 10 min at 35°C. Cells were washed twice with 1x NEBuffer 3.1 (for DpnII libraries) or NEBuffer 2.1 (for HindIII libraries). Chromatin was solubilized with 0.1% SDS for 10 minutes at 65°C, followed by quenching with 1% Triton X-100. Chromatin was digested with 400U HindIII or DpnII overnight at 37°C. After inactivation of restriction enzyme for 20 min at 65°C, DNA ends were filled in with nucleotides and supplemented with biotin-14-dCTP (LifeTech #19518018) for HindIII libraries and biotin-14-dATP (LifeTech #19524016) for DpnII libraries for 4 hours at 23°C. DNA fragments were ligated with T4 DNA ligase (LifeTech #15224090) for 4 hours at 16°C in reactions of 75 μL each. All ligation reactions were combined and samples were treated with proteinase K overnight at 65°C. DNA was purified using 1:1 phenol:chloroform and ethanol precipitation. Samples treated with RNAse A and biotin from unligated ends were removed using T4 DNA polymerase. DNA was sonicated and size-selected using AMpure XP beads (Bedman coulter #A63881) to obtain fragments sized 600-800bp. We performed end repair and a-tailing prior to illumina TruSeq adapter ligation. Each sample was split in two to obtain one SisterC library treated with UV and Hoechst and one Hi-C library without treatment from the same biological sample. SisterC libraries were treated in two reaction volumes of

50uL each with 100ng/uL Hoechst 33342 (ThermoFisher #H3570) for 15 min at room temperature while protecting from light, followed by UV radiation at 2.7kJ/m^2. Samples were washed with TLE 3 times in 30KDa amicon columns (MilliPore # UFC5030BK). Both SisterC and Hi-C libraries were then enriched for biotin-containing fragments by pull down with MyOne Streptavidin C1 beads (Life Tech #65-001). Libraries were amplified, cleaned from pcr primers and sequenced using paired end 50bp reads on an Illumina HiSeq4000 platform. All libraries within a set of replicates were amplified with the same number of PCR cycles.

SisterC and Hi-C analysis

Hi-C and SisterC FASTQ sequencing files were mapped to saccer3 yeast reference genome using publicly available distiller-nf mapping pipeline (https://github.com/mirnylab/distiller-nf) and downstream analysis tools pairtools (https://github.com/mirnylab/pairtools) and cooltools (https://github.com/mirnylab/cooltools). Briefly, reads were mapped with bwa-mem, deduplicated and filtered for mapping quality, resulting in only "valid reads". Reads were classified as inter-sister reads when one read end was mapped as + orientation and the other end as – read orientation. Reads were classified as intra-sister reads when both read ends mapped as + or – read orientation. For downstream analysis, interactions at a shorter distance than 1500bp were removed. Interactions were binned at 1kb, 2kb and 10kb resolution using cooler[43]. Iterative balancing was applied to all matrices, individually for inter-sister and intra-sister interactions, while ignoring two bins from the diagonal[44]. Hi-C and SisterC statistics for all samples are provided in supplemental table 1.

Distance decays were plotted from valid pairs separated by read orientation were used to calculate contact frequency (P) as a function of genomic distance (s) using cooltools code. Pile up plots on genomic loci were produced using valid interactions binned at 1kb resolution separated for inter- or intra-sister interactions and contained only interactions at distances larger than 1500bp. Observed over expected values were calculated using expected files of matching conditions (e.g. expected file of inter-sister interactions of SisterC library for matching observed interactions of inter-sister interactions of SisterC library). Heatmaps were plotted using modified cooltools code. For centromere pile up plot, the directionality of the centromere DNA elements was taken into account. Anchor plots were plotted from valid interactions binned at 2kb resolution, which were separated for inter or intra-sister interactions and contained only interactions at distances larger than 1500bp.

Publicly available datasets used in this manuscript

Scc1 calibrated ChIP-seq tracks from Hu et al[22] were used for cohesin pile up SisterC heatmaps and ChIP-seq tracks in figure 1. This dataset is available on GEO under accession number GSM1712309. Peaks were called on this dataset using MACS2. Pairwise cohesin interactions were compiled by listing all possible pairwise combinations of cohesin peak sites in cis, followed by separation on distance between cohesin pairs (smaller than 10kb, 10 to 20kb, 20 to 35kb and 35 to 50kb). Cohesin sites in a 50kb window around centromeres and on all of chrXII and chrIV were removed from the dataset. Additionally Hi-C samples from cdc45 mutant cells were used from Schalbetter et al [7] to investigate distance decay. This dataset is available on GEO under accession number GSM2327664. This data was processed identical to Hi-C libraries produced for this study. Sites of origin of replication were downloaded from OriDB (http://www.oridb.org/)[45].

Code availability

Hi-C mapping pipeline distiller-nf is available on https://github.com/mirnylab/distiller-nf. Downstream analysis tools pairtools and cooltools are available through https://github.com/mirnylab/pairtools and https://github.com/mirnylab/cooltools.

Data availability

All genomic data generated for this study are publicly available on the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE145695.

Acknowledgements

We thank all current and former Dekker lab members for helpful discussions, in particular Sergey V. Venev for advice on data analysis. We thank Jennifer Benanti, Heather Arsenault and Nick Rhind for advice on culturing and synchronizing yeast and providing the yeast strain used in this study. We thank Stephanie Schalbetter, Nicola Minchell and Matthew Naele for providing their yeast Hi-C protocol and suggestions. We also thank Bas van Steensel for advice and suggestions. This work was supported by grants from the National Institutes of Health (HG003143 to J.D. and NS111990 to J.K.W.). J.D. is an investigator of the Howard Hughes Medical Institute.

Author contributions

M.E.O. and J.D. conceived and designed the project. M.E.O. cultured yeast,

generated and analyzed SisterC and Hi-C datasets and performed flow cytometry experiments. A.K.H. and J.K.W. designed HPLC and mass spectrometry experiments, A.K.H. performed and analyzed HPLC and MS experiments. M.E.O and J.D. wrote the manuscript.

Competing interest statement

All authors declare no conflict of interest.

References

1. Yatskevich, S., Rhodes, J. & Nasmyth, K. Organization of Chromosomal DNA by SMC Complexes. 1–38 (2019).
2. Nagasaka, K., Hossain, M. J., Roberti, M. J., Ellenberg, J. & Hirota, T. Sister chromatid resolution is an intrinsic part of chromosome organization in prophase. *Nat. Cell Biol.* (2016). doi:10.1038/ncb3353
3. Falconer, E. *et al.* Identification of sister chromatids by DNA template strand sequences. *Nature* **463**, 93–97 (2010).
4. Espinosa, E., Paly, E. & Barre, F. High-resolution whole-genome analysis of sister-chromatid cohesion. *bioRxiv* (2019).
5. Haering, C. H., Farcas, A. M., Arumugam, P., Metson, J. & Nasmyth, K. The cohesin ring concatenates sister DNA molecules. *Nature* **454**, 297–301 (2008).
6. Tanaka, T., Cosma, M. P., Wirth, K. & Nasmyth, K. Identification of Cohesin Association Sites at Centromeres and along Chromosome Arms. *Cell* **98**, 847–858 (1999).
7. Schalbetter, S. A. *et al.* SMC complexes differentially compact mitotic chromosomes according to genomic context. *Nat. Cell Biol.* **19**, 1071–1080 (2017).
8. Lazar-Stefanita, L. *et al.* Cohesins and condensins orchestrate the 4D dynamics of yeast chromosomes during the cell cycle. *EMBO J.* **36**, 2684–2697 (2017).
9. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J. & Yu, H. Human cohesin compacts DNA by loop extrusion. *Science (80-.).* **366**, 1345–1349 (2019).
10. Ganji, A. M. *et al.* Title : Real-time imaging of DNA loop extrusion by condensin. **7831**, 1–9 (2018).
11. Golfier, S., Quail, T., Kimura, H. & Brugues, J. Cohesin and condensin extrude loops in a cell-cycle dependent manner. *bioRxiv* 821306 (2019). doi:10.1101/821306
12. Alves, P. & Jonasson, J. New staining method for the detection of sister-chromatid exchanges in BrdU-labelled chromosomes. *J. Cell Sci.* **Vol.32**, 185–195 (1978).
13. Falconer, E. *et al.* DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nat. Methods* **9**, (2012).
14. Claussin, C. *et al.* Genome-wide mapping of sister chromatid exchange events in single yeast cells using strand-seq. *Elife* **6**, 1–17 (2017).
15. Van Wietmarschen, N. & Lansdorp, P. M. Bromodeoxyuridine does not contribute to sister chromatid exchange events in normal or Bloom syndrome cells. *Nucleic Acids Res.* **44**, 6787–6793 (2016).
16. Chaisson, M. J. P. *et al.* Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nat. Commun.* **10**, 1–16 (2019).
17. Porubsky, D. *et al.* Dense and accurate whole-chromosome haplotyping of individual genomes. *Nat. Commun.* **8**, (2017).
18. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–93 (2009).
19. Belaghzal, H., Dekker, J. & Gibcus, J. H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* **123**, 56–65 (2017).
20. Vernis, L., Piskur, J. & Diffley, J. F. X. Reconstitution of an efficient thymidine salvage pathway in Saccharomyces cerevisiae. *Nucleic Acids Res.* **31**, 120e – 120 (2003).
21. Lajoie, B. R., Dekker, J. & Kaplan, N. The Hitchhiker ' s guide to Hi-C analysis : Practical guidelines. *Methods* **72**, 65–75 (2015).
22. Hu, B. *et al.* Biological chromodynamics: A general method for measuring protein occupancy across the genome by calibrating ChIP-seq. *Nucleic Acids Res.* **43**, 1–20 (2015).

23. Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A. & Mirny, L. A. Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harb. Symp. Quant. Biol.* **LXXXII**, 034710 (2018).
24. Lengronne, A. *et al.* Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**, 573–578 (2004).
25. Schalbetter, S. A., Neale, M. J., Fudenberg, G., Baxter, J. & Pollard, K. S. Principles of meiotic chromosome assembly revealed in S . cerevisiae. *Nat. Commun.* 1–12 (2019). doi:10.1038/s41467-019-12629-0
26. Muller, H. *et al.* Characterizing meiotic chromosomes' structure and pairing using a designer sequence optimized for Hi-C. *Mol. Syst. Biol.* **14**, 1–19 (2018).
27. AlHaj Abed, J. *et al.* Highly structured homolog pairing reflects functional organization of the Drosophila genome. *Nat. Commun.* **10**, 1–14 (2019).
28. Zheng, G., Kanchwala, M., Xing, C. & Yu, H. MCM2–7-dependent cohesin loading during s phase promotes sister-chromatid cohesion. *Elife* **7**, 1–25 (2018).
29. Lengronne, A. *et al.* Establishment of Sister Chromatid Cohesion at the S. cerevisiae Replication Fork. *Mol. Cell* (2006). doi:10.1016/j.molcel.2006.08.018
30. Moldovan, G. L., Pfander, B. & Jentsch, S. PCNA Controls Establishment of Sister Chromatid Cohesion during S Phase. *Mol. Cell* (2006). doi:10.1016/j.molcel.2006.07.007
31. Borrie, M. S., Campor, J. S., Joshi, H. & Gartenberg, M. R. Binding, sliding, and function of cohesin during transcriptional activation. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1062–E1071 (2017).
32. Bausch, C. *et al.* Transcription Alters Chromosomal Locations of Cohesin in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **27**, 8522–8532 (2007).
33. Lazar-Stefanita, L. *et al.* Cohesins and condensins orchestrate the 4D dynamics of yeast chromosomes during the cell cycle. *EMBO J.* 1–14 (2017). doi:10.15252/embj.201797342
34. Makrantoni, V. & Marston, A. L. Cohesin and chromosome segregation. *Curr. Biol.* **28**, R688–R693 (2018).
35. Kikuchi, S., Borek, D. M., Otwinowski, Z., Tomchick, D. R. & Yu, H. Crystal structure of the cohesin loader Scc2 and insight into cohesinopathy. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12444–12449 (2016).
36. Lee, B. G. *et al.* Crystal Structure of the Cohesin Gatekeeper Pds5 and in Complex with Kleisin Scc1. *Cell Rep.* **14**, 2108–2115 (2016).
37. Tóth, A. *et al.* Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication. *Genes Dev.* **13**, 320–333 (1999).
38. Skibbens, R. V., Corson, L. B., Koshland, D. & Hieter, P. Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery. *Genes Dev.* **13**, 307–319 (1999).
39. Gibcus, J. H. *et al.* A pathway for mitotic chromosome formation. *Science (80-.).* **359**, eaao6135 (2018).
40. Falconer, E. & Lansdorp, P. M. Strand-seq: A unifying tool for studies of chromosome segregation. *Semin. Cell Dev. Biol.* **24**, 643–652 (2013).
41. Mitter, M. *et al.* Sister-chromatid-sensitive Hi-C reveals the conformation of replicated human chromosomes. *bioRxiv* 2020.03.10.978148 (2020). doi:10.1101/2020.03.10.978148
42. Andrus, A. & Kuimelis, R. G. Base Composition Analysis of Nucleosides Using HPLC. *Curr. Protoc. Nucleic Acid Chem.* **1**, (2000).
43. Abdennur, N. & Mirny, L. A. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics* **36**, 311–316 (2019).
44. Imakaev, M. *et al.* Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* **9**, 999–1003 (2012).

45. Siow, C. C., Nieduszynska, S. R., Müller, C. A. & Nieduszynski, C. A. OriDB, the DNA replication origin database updated and extended. *Nucleic Acids Res.* **40**, 682–686 (2012).

Chapter 6

SisterC 2.0: Optimizing SisterC for application in mammalian cells

Marlies E. Oomen[1], Adam K. Hedger[2], Bastiaan Dekker[1], Jonathan K. Watts[2], Job Dekker[1,3]

Unpublished results

1. Program in Systems Biology, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA;
2. RNA Therapeutics Institute and Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, USA.
3. Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA.

Abstract

We recently developed SisterC, a Hi-C technique which is able to differentiate 3D interactions between and along sister chromatids by incorporating BrdU and inducing single strand nicks by UV/Hoechst treatment. Although this technique was developed in budding yeast, it is by definition applicable to any model system that can be synchronized and that can efficiently integrate BrdU during DNA replication. Here, we describe approaches to optimize SisterC efficiency to enable the study of sister chromatid organization in human cell lines. In vitro experiments suggest that both increased UV radiation and the use of *Pfu* polymerase for amplification improves the depletion of BrdU containing strands. However, when performing SisterC in budding yeast using these conditions, depletion efficiency is not improved. Additionally, we propose a synchronization protocol for the human cell line Hap1 *cdk1as* that allows for the incorporation of BrdU during S-phase, the subsequent synchronization in G2 and mitosis, and release into G1 for control experiments. Using this protocol, we observe over 80% BrdU incorporation in the newly replicated strand measured by HPLC. This promises the successful application of the SisterC protocol to Hap1 *cdk1as* cells, and will allow the study the complex interplay of sister chromatid condensation and disentanglement as cells progress from S-phase to anaphase.

Introduction

During DNA replication in S-phase, sister chromatids are formed and interactions are established between the sisters and along the same sister chromatid. As the cell progresses through the cell cycle into mitosis, chromosomes face a complicated topological problem; sister chromatids need to be disentangled and condensed, while side-to-side alignment is maintained to enable processes such as DNA damage repair[1–3]. As sister chromatids have identical DNA sequences, inter-sister and intra-sister interactions cannot be distinguished in conventional genomics techniques such as Hi-C. Therefore, we developed SisterC as described in chapter 5 of this thesis and published in Oomen et al[4]. Briefly, SisterC enables differentiation of interactions between and along sister chromatids based on strand orientation after mapping of the SisterC reads (in this thesis; chapter 5 - figure 1a-c). In order to perform SisterC, cells are first synchronized in G1 and then released into S phase in presence of BrdU, which will be incorporated in the newly replicated strands. Cells are then harvested in the cell cycle state of interest. This is followed by standard processing of cells for Hi-C, with the addition of UV/Hoechst treatment of the libraries before biotin pull-down. SisterC was established in budding yeast, using the mutant strain YLV-11 that can take up BrdU efficiently[5]. The small yeast genome allowed for optimization of the technique without requiring many sequence reads. We assessed the efficiency of BrdU depletion by quantifying inter-sister signals in G1-synchronized cells, a phase in the cell cycle when there are no sister chromatids and therefore no inter-sister signal is expected. On average 20.7% of interactions within the same chromosome are misassigned as inter-sister interactions in G1 between the replicates in the original set of SisterC libraries. This suggests that ~70-80% of BrdU containing strands are successfully depleted in a SisterC experiment. Despite suboptimal depletion of BrdU containing strands, we were able to use SisterC to gain better understanding of the spatial organization of pairs of sister chromatids[4].

An alternative C-based technique to study sister chromatid organization, called scs-HiC, was developed by the Gerlich laboratory[6]. Scs-HiC uses the thymidine analogue 4sT to create T-to-C conversion after chemical treatment of the Hi-C library. Hi-C interactions between sister chromatids can then be assigned as inter-sister or intra-sister interactions after amplification and high-throughput sequencing based on the chemically induced SNPs. Scs-HiC was established in HeLa cells, showing that sister chromatid interactions can indeed be distinguished in mammalian cell lines. However, scs-HiC depends on a complicated toxic chemical reaction and requires many sequence reads to obtain proper signal, as 4sT incorporation efficiency is very low. Because of the need for a specialized chemical safety set up and the high experimental costs, it will be challenging for other laboratories to implement scs-HiC.

We set out to optimize the SisterC technique to obtain better depletion of BrdU-containing strands by UV/Hoechst treatment, which will enable the application of SisterC to mammalian cells. We explored the role of several factors in the original SisterC protocol: UV radiation intensity, the polymerase that is used for final library amplification, and denaturing of the SisterC library to single strands by heat before biotin pull down. Additionally, we show preliminary data on BrdU incorporation efficiency in the human cell line Hap1 *cdk1as*[7]. Sufficient BrdU incorporation will be essential to adapt SisterC to mammalian cells. We propose a protocol tailored to Hap1 *cdk1as*, which enables both efficient BrdU incorporation and allows for easy synchronization throughout the cell cycle.

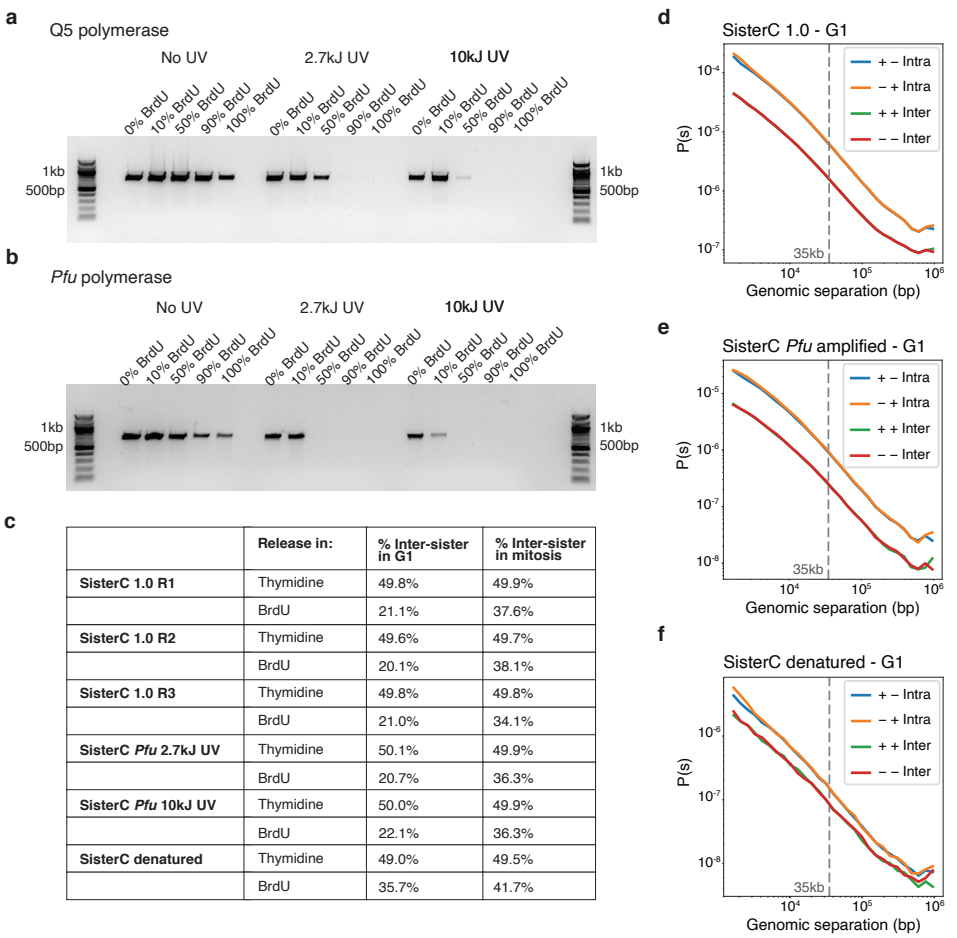
Results

SisterC protocol optimization

We aimed to optimize the original SisterC technique at three specific steps in the protocol. First, we tested different UV radiation strengths. For the development of SisterC, we were inspired by the technique Strand-seq[8]. We therefore used 2.7kJ/m^2 UV for the initial SisterC protocol, as is standard for Strand-seq[9–11]. UV treatment combined with Hoechst incubation is essential to create single-strand nicks at BrdU sites, which will deplete BrdU-containing strands upon PCR amplification. However, high levels UV radiation will destroy any DNA molecule and will result in amplification errors. UV radiation strength should be titrated to only prevent amplification of BrdU-containing strands. We tested a range of UV radiation strengths (No UV, 2.7kJ/m^2, 10kJ/m^2, 20kJ/m^2 and 40kJ/m^2) on a DNA template which was amplified in presence of 0%, 10%, 50%, 90% and 100% BrdUTP. This creates DNA molecules with BrdU in both strands. We then treated the DNA with Hoechst and radiated with UV, followed by DNA clean up and amplification using 10 PCR cycles. As BrdU is present in both strands, UV/Hoechst treatment should not be able to amplify DNA fragments with high levels of BrdU. We found that 10kJ/m^2 UV radiation improved BrdU-containing strand depletion efficiency compared to 2.7kJ/m^2, (figure 1a). When higher levels of UV radiation were used, we found that strands withouth BrdU were also depleted, suggesting that DNA was too damaged to be amplified (data not shown).

Second, we tested whether the choice of polymerase enzyme used to amplify the SisterC library effects BrdU depletion efficiency. We hypothesized different polymerases will be more or less successful in amplifying nicked DNA strands and will therefore have more or less misassigned sister chromatid interactions. In the original SisterC experiments, we used the Illumina TruSeq DNA nano kit for library preparation, which uses Q5 polymerase for final library amplification. In addition to Q5 polymerase we tested *Pfu* polymerase using the same in vitro DNA fragment

assay as described above. Figure 1b shows that *Pfu* polymerase is more successful in discriminating against amplification of BrdU-containing strands, especially when radiated with 10kJ/m^2 UV.



	Release in:	% Inter-sister in G1	% Inter-sister in mitosis
SisterC 1.0 R1	Thymidine	49.8%	49.9%
	BrdU	21.1%	37.6%
SisterC 1.0 R2	Thymidine	49.6%	49.7%
	BrdU	20.1%	38.1%
SisterC 1.0 R3	Thymidine	49.8%	49.8%
	BrdU	21.0%	34.1%
SisterC *Pfu* 2.7kJ UV	Thymidine	50.1%	49.9%
	BrdU	20.7%	36.3%
SisterC *Pfu* 10kJ UV	Thymidine	50.0%	49.9%
	BrdU	22.1%	36.3%
SisterC denatured	Thymidine	49.0%	49.5%
	BrdU	35.7%	41.7%

Figure 1 – Approaches to optimize the SisterC protocol for more efficient BrdU depletion by UV/Hoechst treatment. (a-b) DNA gel image of fragments amplified in presence of 0%, 10%, 50%, 90% and 100% BrdUTP (supplemented with dTTP). Samples were then treated with Hoechst and no radiation (first 5 lanes) or UV radiation at 2.7kJ/m^2 (middle 5 lanes) or 10kJ/m^2 (last 5 lanes), followed by 10 PCR cycles amplification using Q5 polymerase (**a**) or *Pfu* polymerase (**b**). (**c**) Percentage inter-sister interactions in G1 and mitosis on chrXIII for original SisterC 1.0 replicates and the modified SisterC protocols using *Pfu* polymerase after UV radiation with 2.7kJ and 10kJ/m^2, or denatured DNA (**d**) Distance decay of combined G1 replicates of original published SisterC 1.0 protocol. (**e**) Distance decay of G1 SisterC libraries that were radiated with 2.7 kJ/m^2 and amplified with *Pfu* polymerase. (**f**) Distance decay of G1 SisterC libraries that were denatured by heat prior to biotin pull down.

As these in vitro experiments using PCR-amplified DNA fragments suggested that the use of increased UV radiation and *Pfu* polymerase for library amplification could be promising to improve the SisterC protocol, we continued to perform SisterC in budding yeast using these optimized parameters. Unfortunately, SisterC libraries amplified using *Pfu* polymerase after radiation with 2.7kJ/m^2 UV did not show improved depletion of BrdU containing strands (figure 1c and 1e and supplemental table 1). Additionally, increased UV radiation to 10kJ/m^2 did not improve BrdU depletion either (figure 1c and supplemental table 1). Moreover, mapping statistics showed that SisterC libraries radiated with 10kJ/m^2 UV contained less valid Hi-C interactions, suggesting that higher UV radiation damages the ligation products beyond creating BrdU single strand nicks (Supplemental table 1).

In order to enrich a Hi-C or SisterC library for 3D interactions, we pull down on biotinylated nucleotides that label a ligation junction using streptavidin beads. As DNA molecules are double stranded when libraries are pulled down, strands will be in close proximity to their compliment strand on the streptavidin bead during amplification. This could potentially give rise to chimeric amplification products, which could occur when polymerase enzymes face nicks, as is the case for our SisterC libraries. When the polymerase is halted due to a Hoechst/UV-induced nick, while also in close proximity to the non-BrdU containing complementary strand, the nicked strand could be extended using the intact strand as complement. These chimeric amplification products will then become misassigned as inter-sister or intra-sister interactions. We hypothesized that G1 inter-sister interactions in the original SisterC dataset were due to chimeric amplification products. In order to prevent compliment strands being in close proximity to each other during amplification, we denatured SisterC and Hi-C budding yeast libraries prior to biotin pull-down by heating up the libraries for 15 minutes at 95°C. However, contrary to what we expected, this resulted in much less efficient BrdU depletion as G1 inter-sister signal increases to 35.7% compared to 20.7% average across replicates in the initial SisterC dataset (figure 1c and compare figure 1d to 1f and supplemental table 1).

SisterC in mammalian cells

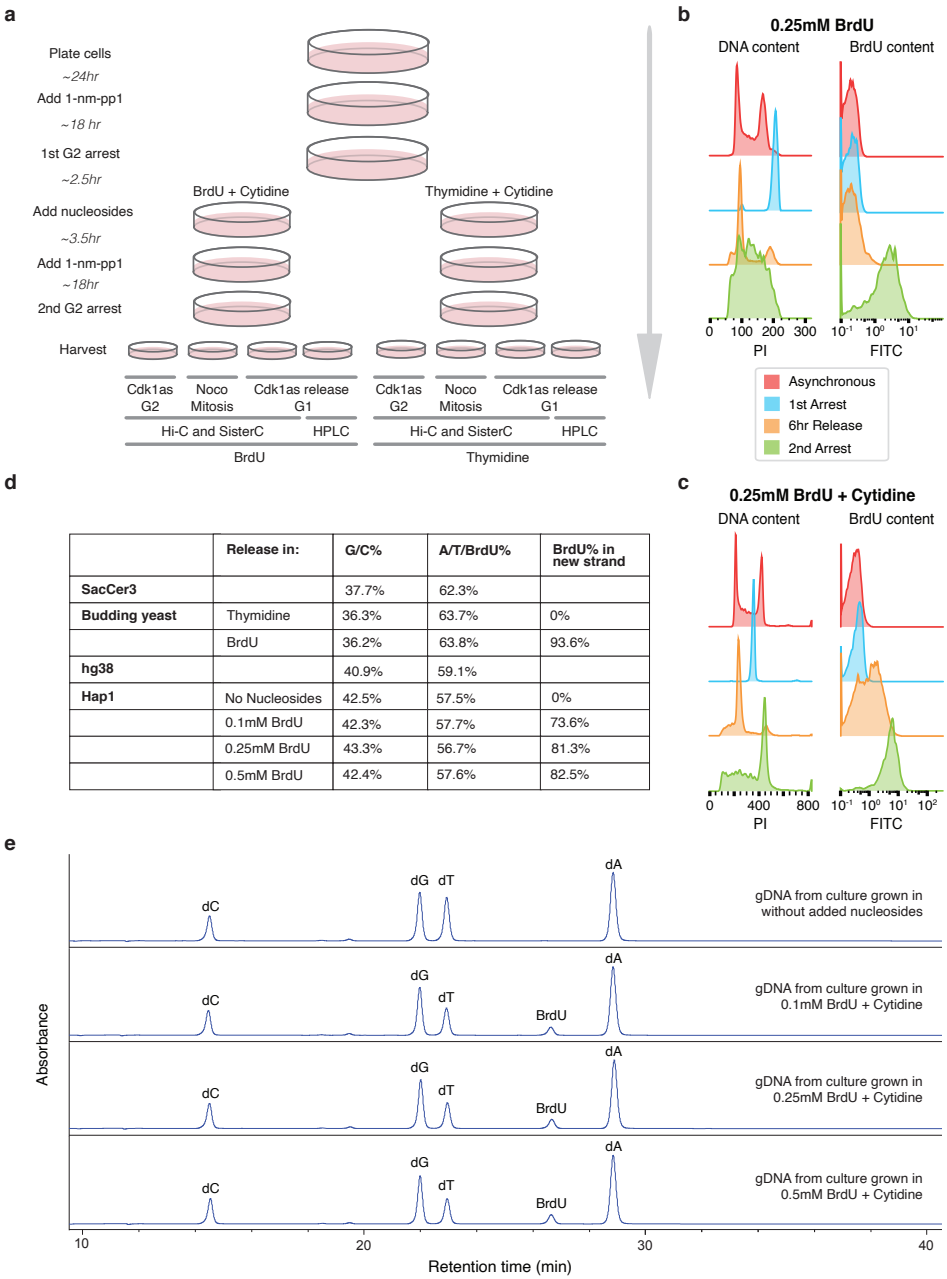
Although SisterC was originally developed and tested in budding yeast[4], SisterC is in principle applicable to all systems that can be synchronized and incorporate BrdU. One system of particular interest to many researchers is cultured human cell lines. We set out to adapt the SisterC protocol for use in the cell line Hap1. In budding yeast, we were able to use alpha-factor arrest to synchronize cells in late G1-phase prior to BrdU incorporation. Unfortunately, this is not applicable to most mammalian cells. Traditional synchronization protocols often rely on synchronization

in S-phase by creating a nucleotide imbalance using thymidine[12]. However, adding thymidine prior BrdU incorporation – a thymidine analogue – would result in low incorporation efficiency. To enable easy synchronization, a mutant Hap1 cell line was created, in which endogenous *Cdk1* has been replaced by an analogue-sensitive mutant *cdk1as*. When the ATP analogue 1-NM-PP1 is added to the media, Cdk1as will temporarily be disabled and cells become arrested in G2. When 1-NM-PP1 is washed out, cells are released from the G2-arrest and will synchronously enter prophase and progress through mitosis. This cell line design has previously been used in the chicken cell line DT40[7] and was successfully used to study chromosome conformation in chicken cells during the progression from G2-phase to prometaphase by Hi-C[13].

To synchronize a population of cells to study using SisterC, we first arrest Hap1 *cdk1as* cells in late G2 using 1-NM-PP1 (figure 2a). After a brief release to allow the cells to enter the subsequent cell cycle, BrdU is added to the media to enable strand specific labelling of the sister chromatids during S-phase. This is then followed by a second 1-NM-PP1 arrest in G2, from which cells can be synchronously released. Depending on which phase of the cell cycle the researcher wants to study using SisterC, cells can be harvested in G2, at different timepoints as the population progresses from G2 to mitosis, or in prometaphase using a nocodazole arrest (figure 2a). Similarly to SisterC experiments in budding yeast[4], it is advisable to collect a population of G1 cells after the S-phase during which BrdU was incorporated to measure misassigned 'inter-sister' interactions. This can be done by letting cells release from the 1-NM-PP1 arrest and progress though mitosis into the subsequent G1, possibly added by a block in G1 by cell cycle inhibitors such as palbociclib[14]. By performing both HPLC and SisterC on this G1 population, incorporation efficiency and the level of depletion of BrdU-containing strands after UV/Hoechst treatment can be assessed (figure 2a).

Hap1 cells are very sensitive to addition of thymidine or thymidine analogues such as BrdU. When high concentrations of BrdU are added to the culture after release from the 1st 1-NM-PP1 arrest, Hap1 *cdk1as* cells will not progress through S-phase and many cells in the population will die (figure 2b). Cells most likely arrest due to a nucleotide imbalance[15]. However, this imbalance can be resolved by the addition of cytidine in equimolar concentration[16], which both keeps cells healthier and allows for better incorporation of BrdU by Hap1 *cdk1as* cells, as shown by flow cytometry in figure 2c.

We then continued to quantify the level of BrdU incorporation in Hap1 *cdk1as* cells by HPLC (figure 2d-e). When we feed budding yeast strain YLV-11 1 mM BrdU according to the published SisterC protocol[4,17], we typically observe that



	Release in:	G/C%	A/T/BrdU%	BrdU% in new strand
SacCer3		37.7%	62.3%	
Budding yeast	Thymidine	36.3%	63.7%	0%
	BrdU	36.2%	63.8%	93.6%
hg38		40.9%	59.1%	
Hap1	No Nucleosides	42.5%	57.5%	0%
	0.1mM BrdU	42.3%	57.7%	73.6%
	0.25mM BrdU	43.3%	56.7%	81.3%
	0.5mM BrdU	42.4%	57.6%	82.5%

Figure 2 – Optimization of synchronization and BrdU incorporation by Hap1 Cdk1as cells. (a) Schematic overview of Hap1 Cdk1as synchronization protocol that enables BrdU incorporation during S-phase, followed by harvesting of SisterC/Hi-C and HPLC samples in G2, mitosis and subsequent G1-phase. (**b-c**) Flow cytometry data of cells synchronized in media supplemented with 0.25mM BrdU (**b**) and 0.25mM BrdU and 0.25mM Cytidine (**c**). (**d**) Nucleoside percentage in budding yeast genome SacCer3 and human genome hg38 compared to nucleoside content measured by HPLC in budding yeast and Hap1 Cdk1as

cells when synchronized in media supplemented with BrdU at different concentrations. The percentages of each nucleoside are calculated using the corresponding extinction coefficient. (**e**) HPLC spectra measured at 260nm of digested gDNA isolated from Hap1 Cdk1as cells synchronized in media without nucleoside (top), 0.1 mM BrdU and Cytidine (second), 0.25 mM BrdU and Cytidine (third) and 0.5 mM BrdU and Cytidine (bottom).

90-95% of thymidine has been replaced by BrdU in the newly replicated strand. We tested BrdU incorporation by Hap1 *cdk1as* using above-described protocol at 3 different concentrations of BrdU - 0.1 mM, 0.25 mM and 0.5 mM - supplemented with cytidine at equimolar concentrations. We observed that when adding 0.1 mM BrdU to Hap1 *cdk1as* cells, cells incorporate approximately 73.6% BrdU, whereas cells synchronized in presence of 0.25mM and 0.5mM BrdU incorporate 81.3 and 82.5% respectively. However, even with the addition of equimolar concentrations cytidine during synchronization, Hap1 *cdk1as* cells show an increased fraction of dead cells after addition of all concentrations of BrdU ranging from 20% up to 27%, as observed by trypan blue staining (supplemental table 2). As the incorporation efficiency does not improve dramatically between 0.25 mM and 0.5 mM, but the fraction dead cells is does increase at higher concentrations, we suggest using 0.25 mM BrdU to prepare Hap1 *cdk1as* cells for future SisterC experiments.

Discussion

Here we describe strategies to adapt the SisterC protocol for use in mammalian cells, such as the human cell line Hap1 *cdk1as*. In the published SisterC protocol, we observed 20-22% inter-sister interactions in G1[4], a cell cycle phase when these interactions are not expected as sister chromatids are segregated into the two new daughter cells during the preceding mitosis. This suggests that in mitosis similar levels of inter-sister interactions are misassigned as intra-sister interactions and vice versa. We therefore set out to decrease the levels of wrongly assigned interactions in SisterC. Although in vitro tests using a DNA template containing BrdU suggested that using higher level of UV radiation and using *Pfu* polymerase could improve SisterC specificity, when these conditions were used to perform SisterC in budding yeast no improvement was observed. Similarly, we tested whether denaturing double stranded DNA into single stranded molecules prior to biotin pull-down of SisterC interactions. Not only did this not decrease inter-sister interactions in G1, but it also resulted in low-quality SisterC libraries (figure 1 and supplemental table 1). Additionally, we set the first steps towards the application of SisterC in mammalian cells. We propose a synchronization protocol for Hap1 *cdk1as* cells, a human cell line that allows for easy synchronization in G2 and synchronous release

as the cells progress through mitosis. Supplementing BrdU and cytidine to the media during the synchronization protocol allows for optimal BrdU incorporation efficiency. HPLC measurements on DNA isolated from cells synchronized using this protocol estimate over 80% BrdU incorporation in the newly replicated strand (figure 2). This suggests that SisterC will be feasible in Hap1 *cdk1as* cells without much further optimization being necessary.

Although we did not find conditions that improve SisterC specificity, there are still several aspects of the protocol that could be worth pursuing to optimize depletion of BrdU containing strands. In the current SisterC protocol, fragments are sonicated to obtain 600-800bp size. This is larger than typically done for Hi-C libraries, for which the recommended fragment size is 150-300bp. The reasoning to maintain fragments at larger size, is to increase the chance of UV/Hoechst induced single strand nicks on both sides of the ligation junction, and therefore successful depletion of BrdU containing strands upon PCR amplification. However, using a longer fragment size comes with the risk that a small fraction of sequenced reads contains 3 or more 3D interactions instead of 2. As we use strand orientation of each end of the paired-end sequence run to assign interactions as inter- or intra-sister interactions, this could lead to misassignment. However, if SisterC fragments contain more than 2 interactions, it is also more likely that the sequenced ends of the reads are close to a DpnII restriction site. When we explored for the published SisterC data whether inter-sister interactions in G1 libraries are more prone to map closer to DpnII restriction sites, we did not observe this (chapter 5 – extended data figure 10)[4]. Nonetheless, if further optimization of the current SisterC protocol is essential for adaptation in mammalian cell lines, exploring different fragment lengths of interactions could be an avenue worth pursuing. Along these lines, a recent publication showed that different digestion enzymes and crosslinkers can improve Hi-C library quality[18]. This could also be an improvement to the SisterC protocol. Although this experiment was not repeated, when SisterC was performed using the restriction enzyme HindIII instead of DpnII for initial digestion, the level of background inter-sister interactions in G1 was substantially higher (25%) than when using DpnII (20-22%)[4]. However, coincidentally the level of BrdU incorporation in these HindIII as measured by HPLC was lower for DpnII replicates (83% compared to 90-95%), which could also have caused an increased level of misassigned inter-sister interactions in G1[4].

We were motivated to optimize the SisterC protocol as we hypothesized that mammalian cells would be much less successful at taking up and incorporating BrdU compared to the budding yeast strain YLV-11. However, HPLC data of Hap1 *cdk1as* cells synchronized in presence of 0.25mM BrdU suggest that over 80% of the thymidine residues have been replaced by BrdU in the newly replicated strand.

This is much higher than we anticipated and relies on the addition of equimolar concentration of cytidine which allows Hap1 cells to tolerate high levels of BrdU. Moreover, the Hap1 *cdk1as* cell line allows for relatively easy synchronization in G2, which can be followed by a synchronous release as cells progress into mitosis. This will allow to gain a better understanding in the interplay of cohesins and condensins that mediate sister chromatid cohesion and loop formation in mammalian cells. SisterC opens doors to study the complicated topological problem cells face, when sister chromatids need to orchestrate both compaction and segregation to enable proper cell division.

Supplemental materials

Replicate	Release in	Cell-cycle state	UV/Hoechst treatment	Reads total	Total Mapped		Total valid reads (no duplicates)		Cis all interactions		Cis all >1500bp		Cis intra-sister >1500bp		Cis inter-sister >1500bp		Trans all		Trans intra-sister			Trans inter-sister		
				#	#	% of total	#	% of mapped	#	% of valid	#	% of cis	#	% of cis >1500	#	% of cis >1500	#	% of cis >1500	#	% of valid	#	% of trans	#	% of trans
Published SisterC 1.0 DpnII R1	BrdU	Mitosis	Yes	2.27E+07	1.57E+07	69.11	1.42E+07	90.71	1.04E+07	73.32	5.12E+06	49.16	3.20E+06	62.42	1.92E+06	37.58	3.79E+06	26.68	1.90E+06	50.00	1.90E+06	50.00
	BrdU	Mitosis	No	2.28E+07	1.54E+07	67.33	1.57E+07	101.97	1.11E+07	70.64	6.78E+06	61.30	3.40E+06	50.11	3.38E+06	49.89	4.60E+06	29.36	2.30E+06	50.02	2.30E+06	49.98
	BrdU	G1	Yes	2.27E+07	1.57E+07	69.11	1.46E+07	92.99	9.50E+06	65.20	3.96E+06	41.71	3.12E+06	78.90	8.36E+05	21.10	5.07E+06	34.80	2.53E+06	49.99	2.53E+06	50.01
	BrdU	G1	No	2.28E+07	1.54E+07	67.33	1.46E+07	95.10	8.65E+06	59.19	4.60E+06	53.14	2.31E+06	50.24	2.29E+06	49.76	5.99E+06	40.81	2.98E+06	49.99	2.98E+06	50.01
	Thymidine	Mitosis	Yes	2.46E+07	1.66E+07	67.55	1.68E+07	101.09	1.27E+07	75.30	8.29E+06	65.55	4.16E+06	50.15	4.13E+06	49.85	4.15E+06	24.70	2.08E+06	50.03	2.07E+06	49.97
	Thymidine	Mitosis	No	2.43E+07	1.70E+07	69.93	1.60E+07	94.33	1.21E+07	75.21	8.29E+06	65.55	4.16E+06	50.15	4.13E+06	49.85	3.98E+06	24.79	1.99E+06	49.99	1.99E+06	50.01
	Thymidine	G1	Yes	2.46E+07	1.66E+07	67.55	1.56E+07	93.68	8.78E+06	56.42	4.50E+06	51.20	2.29E+06	50.16	2.24E+06	49.84	6.78E+06	43.58	3.39E+06	50.00	3.39E+06	50.00
	Thymidine	G1	No	2.04E+07	1.37E+07	67.15	1.30E+07	94.82	7.19E+06	55.36	3.82E+06	53.08	1.92E+06	50.28	1.90E+06	49.74	5.80E+06	44.64	2.90E+06	50.03	2.90E+06	49.97
Denatured SisterC	BrdU	Mitosis	Yes	4.54E+07	2.92E+07	64.22	1.76E+07	60.35	1.62E+07	92.08	1.22E+06	7.51	7.09E+05	58.27	5.08E+05	41.73	1.35E+06	7.66	6.76E+05	50.11	6.72E+05	49.86
	BrdU	Mitosis	No	4.76E+07	3.10E+07	65.23	1.89E+07	60.86	1.22E+07	64.67	7.29E+06	59.83	3.64E+06	50.02	3.64E+06	49.98	6.68E+06	35.33	3.34E+06	50.00	3.34E+06	50.00
	BrdU	G1	Yes	5.41E+07	2.94E+07	54.37	1.63E+07	55.28	1.41E+07	87.03	1.17E+06	8.26	7.51E+05	64.27	4.17E+05	35.73	2.11E+06	12.97	1.05E+06	50.03	1.05E+06	49.97
	BrdU	G1	No	3.63E+07	2.29E+07	63.20	1.42E+07	61.81	7.41E+06	52.28	3.32E+06	44.74	1.62E+06	48.98	1.69E+06	51.02	6.77E+06	47.72	3.38E+06	50.00	3.38E+06	50.00
	Thymidine	Mitosis	Yes	4.72E+07	3.36E+07	71.12	1.79E+07	53.34	1.42E+07	79.55	3.77E+06	26.44	1.90E+06	50.51	1.86E+06	49.49	3.66E+06	20.45	1.83E+06	49.99	1.83E+06	50.01
	Thymidine	Mitosis	No	4.25E+07	2.79E+07	65.53	1.72E+07	61.76	1.09E+07	63.57	6.63E+06	60.59	3.34E+06	50.36	3.29E+06	49.64	6.27E+06	36.43	3.13E+06	50.00	3.13E+06	50.00
	Thymidine	G1	Yes	5.17E+07	3.57E+07	69.08	1.79E+07	50.03	1.41E+07	78.87	2.12E+06	15.02	1.08E+06	51.01	1.04E+06	48.99	3.78E+06	21.13	1.89E+06	50.04	1.89E+06	49.96
	Thymidine	G1	No	5.13E+07	3.34E+07	65.01	1.91E+07	57.18	1.07E+07	55.88	4.82E+06	45.23	2.44E+06	50.70	2.38E+06	49.30	8.42E+06	44.12	4.21E+06	50.04	4.20E+06	49.96
SisterC pfu amplified 2.7kJ	BrdU	Mitosis	Yes	3.52E+07	2.20E+07	62.34	1.72E+07	78.46	1.29E+07	75.05	6.70E+06	51.77	4.27E+06	63.74	2.43E+06	36.26	4.30E+06	24.95	2.15E+06	50.02	2.15E+06	49.98
	BrdU	Mitosis	No	3.32E+07	1.98E+07	59.57	1.66E+07	85.63	1.23E+07	73.96	7.98E+06	63.70	4.00E+06	50.14	3.98E+06	49.86	4.41E+06	26.04	2.21E+06	50.03	2.20E+06	49.97
	BrdU	G1	Yes	3.58E+07	2.30E+07	64.38	1.90E+07	82.46	1.44E+07	75.67	9.63E+06	67.03	4.83E+06	50.10	4.81E+06	49.90	4.62E+06	24.33	2.31E+06	49.98	2.31E+06	50.02
	BrdU	G1	No	3.49E+07	2.24E+07	64.34	1.91E+07	85.06	1.45E+07	75.70	9.66E+06	66.82	4.84E+06	50.10	4.82E+06	49.90	4.64E+06	24.30	2.32E+06	50.02	2.32E+06	49.98
	Thymidine	Mitosis	Yes	3.49E+07	2.30E+07	65.96	1.88E+07	81.75	1.30E+07	68.82	5.61E+06	43.28	4.44E+06	79.28	1.16E+06	20.72	5.87E+06	31.18	2.94E+06	50.01	2.93E+06	49.99
	Thymidine	Mitosis	No	3.49E+07	2.23E+07	63.95	1.91E+07	85.27	1.19E+07	62.64	7.01E+06	58.77	3.51E+06	50.10	3.50E+06	49.90	7.12E+06	37.36	3.56E+06	50.03	3.56E+06	49.97
	Thymidine	G1	Yes	3.45E+07	2.28E+07	66.06	1.94E+07	84.98	1.33E+07	68.69	5.34E+06	40.10	2.66E+06	49.84	2.68E+06	50.16	6.07E+06	31.31	3.03E+06	50.00	3.03E+06	50.00
	Thymidine	G1	No	3.74E+07	2.39E+07	63.96	2.03E+07	85.06	1.23E+07	60.64	7.08E+06	57.34	3.52E+06	49.78	3.55E+06	50.22	8.01E+06	39.36	4.00E+06	49.98	4.01E+06	50.02
SisterC pfu amplified 10kJ	BrdU	Mitosis	Yes	1.26E+07	8.07E+06	63.80	6.68E+06	85.22	5.98E+06	86.90	1.35E+06	22.54	8.58E+05	63.71	4.89E+05	36.29	9.01E+05	13.10	4.50E+05	50.00	4.50E+05	50.00
	BrdU	Mitosis	No	3.32E+07	1.98E+07	59.57	1.69E+07	85.63	1.25E+07	73.98	7.98E+06	63.70	4.00E+06	50.14	3.98E+06	49.86	4.41E+06	26.04	2.21E+06	50.03	2.20E+06	49.97
	BrdU	G1	Yes	1.19E+07	7.96E+06	66.85	6.90E+06	86.67	5.72E+06	82.99	2.43E+06	42.51	1.22E+06	50.12	1.21E+06	49.88	1.17E+06	17.01	5.86E+05	49.95	5.87E+05	50.05
	BrdU	G1	No	3.49E+07	2.24E+07	64.34	1.91E+07	85.06	1.45E+07	75.70	9.66E+06	66.82	4.84E+06	50.10	4.82E+06	49.90	4.64E+06	24.30	2.32E+06	50.02	2.32E+06	49.98
	Thymidine	Mitosis	Yes	1.33E+07	9.60E+06	72.18	8.13E+06	84.60	7.43E+06	91.41	6.46E+05	8.69	5.03E+05	77.94	1.42E+05	22.06	6.98E+05	8.59	3.49E+05	50.08	3.48E+05	49.92
	Thymidine	Mitosis	No	3.49E+07	2.23E+07	63.95	1.91E+07	85.27	1.19E+07	62.64	7.01E+06	58.77	3.51E+06	50.10	3.50E+06	49.90	7.12E+06	37.36	3.56E+06	50.03	3.56E+06	49.97
	Thymidine	G1	Yes	3.43E+07	2.34E+07	68.38	2.00E+07	85.19	1.63E+07	81.80	3.16E+06	19.34	1.58E+06	50.00	1.58E+06	50.00	3.63E+06	18.20	1.82E+06	50.04	1.81E+06	49.96
	Thymidine	G1	No	3.74E+07	2.39E+07	63.96	2.03E+07	85.06	1.23E+07	60.64	7.08E+06	57.34	3.52E+06	49.78	3.55E+06	50.22	8.01E+06	39.36	4.00E+06	49.98	4.01E+06	50.02

Supplemental table 1 - Mapping statistics SisterC optimization libraries. Note that samples not radiated with UV for both *Pfu* amplified SisterC sets are identical.

	# cells harvested for HPLC	% dead cells by Trypan blue
No Nucleotide	4.23E+07	10
0.1mM BrdU	3.29E+07	20.5
0.25mM BrdU	3.28E+07	23
0.5mM BrdU	2.72E+07	27

Supplemental table 2 – Trypan blue staining to quantify % dead cells during Hap1 *cdk1as* synchronization for HPLC experiments.

Methods

Amplification of DNA fragments to test UV and polymerase conditions

DNA fragments (686bp) were amplified for 15 cycles using DreamTaq in presence of 0%, 10%, 50%, 90% or 100% BrdUTP supplemented with dTTP. Amplified fragments were incubated with 100ng/μL Hoechst 33342 in TLE for 15 minutes at room temperature while shielding from light. Samples were either not radiated or UV radiated at 2.7kJ/m^2 or 10kJ/m^2, followed by amicon clean up 3 times using 30kDa cutoff columns. Samples were amplified using Q5 polymerase (NEB # M0515S) or *Pfu* polymerase (Agilent #600674) for 10 cycles and run on 1.25% agarose ethidium bromide gels.

Yeast culture conditions and synchronization

The yeast strain YLV11[5] was used for all budding yeast experiments in this study. Yeast was cultured in yeast peptone (YP) media supplemented with 2% galactose and 100μM thymidine at 30ºC. Synchronization and harvesting for SisterC and Hi-C was performed as described in Oomen et al[4]. Briefly, cells were synchronized in late G1 using alpha-factor, followed by a boost of either 500μM Thymidine or 500μM BrdU and release in respectively 1mM Thymidine or 1mM BrdU. Cells were arrested in mitosis using nocodazole or in the subsequent G1 using alpha-factor. Cells were fixed for Hi-C and SisterC using 3% formaldehyde and stored at -80ºC until further processing.

SisterC yeast

SisterC were performed as previously published[4,17,19], with modifications to test SisterC parameters as described above. Briefly, crosslinked cells were washed in spheroplasting buffer and lysed using 0.5% beta-mercaptoethanol and zymolyase for 10 minutes at 35ºC. Cells were washed with 1x NEBuffer 3.1 and chromatin was solubilized using 0.1% SDS at 65ºC for 10 minutes, followed by quenching with 1% Triton-X100. Chromatin was then digested overnight with 400U DpnII at 37ºC, followed by inactivation of the enzyme at 65ºC for 20 minutes. Digested restriction sites were filled in with nucleotides supplemented with Biotin-14-dATP for 4 hours at 23ºC. Samples were aliquoted in reactions of 75μL each in a 96-well pcr plate and ligated with T4 DNA ligase for 4 hours at 16ºC. After ligation, reactions of each sample were combined and treated with proteinase K at 65ºC for 2 hours, followed by another proteinase K dose with incubation overnight. DNA was purified using phenol:chloroform and ethanol precipitation. Samples were treated with RNase A and biotin from unligated ends were removed. DNA was sonicated and selected using AMpure beads to obtain average fragment size of 600-800bp. Samples were

split in two for the SisterC on denatured DNA and in 3 to test different UV radiation strengths. After DNA end repair and a-tailing, DNA was ligated to barcoded Illumina TruSeq adapters. SisterC samples were treated with 100ng/μL Hoechst 33324 and incubated for 15 minutes at room temperature while shielding from light. Samples where then radiated with 2.7kJ or 10kJ/m^2 UV and washed 3 times in amicon columns. For the libraries that were denatured prior to pull down, samples were incubated at 95 ºC for 15 minutes, followed by 5 minutes on ice. All SisterC and Hi-C samples were then enriched for biotin-containing fragments by pull-down on MyOne streptavidin C1 beads. Libraries were amplified using the Illumina TruSeq master mix or *Pfu* polymerase (Agilent #600674) and TruSeq primer cocktail. All samples of the same set of SisterC libraries were amplified with the same number of PCR cycles, typically between 7-10 cycles. PCR primers were removed using AMpure beads and libraries were sequenced using 50bp paired-end reads on an HiSeq 4000 system (Illumina).

SisterC mapping and downstream analysis

SisterC data was mapped and processed as previously published[4]. FASTQ sequencing files were mapped to the budding yeast reference genome SacCer3 using the publicly available mapping pipeline distiller-nf (https://github.com/mirnylab/distiller-nf) and downstream analysis tools pairtools (https://github.com/mirnylab/pairtools) and cooltools (https://github.com/mirnylab/cooltools). Interactions shorter than 1500bp were removed from downstream analysis. Reads with read orientation + – and – + were classified as intra-sister interactions, while reads with read orientation + + and – – were classified as inter-sister interactions.

Hap1 Cdk1as culture conditions and synchronization

Hap1 Cdk1as cells were grown in IMDM supplemented with 10% FBS and1x pen/strep (Gibco #P4333). To start synchronization, ~0.75 million cells were plated per 10cm dish. After 24 hours, 1-NM-PP1 (1μM final concentration) was added to media for 16 hours to arrest cells in late G2-phase. Cells were realized from arrest by washing 2x with warm PBS and 1x media, followed by adding warm media for 2.5 hours. BrdU or Thymidine with equimolar concentration Cytindine were added to the medium for another 2.5 hours. 1-NM-PP1 (1μM final concentration) was added to media for an additional 15 hours to block cells in the 2nd G2 arrest. Cells can then be released again by washing with warm PBS (2x) and media (1x). For mitotic arrest, cells were released from the 2nd 1-nm-pp1 arrest for ~30 minutes, followed by a nocodazole arrest (50ng/mL) for ~2 hours. Obtain cells arrested in G1, cells were released from 2nd 1-NM-PP1 arrest for ~2 hours, followed by at least 3 hours

incubation with Palbociclib.

Flow cytometry

Hap1 cells were fixed with 95% ethanol and stored at -20ºC until processed for flow cytometry. Cells were pelleted for 5 min at 500xg and washed three times in incubation buffer (PBS + 0.5% BSA). Cells were lysed by incubation with PBS + 1.5M HCL for 30 minutes at room temperature, followed by 2 washes in incubation buffer. Samples were immunostained for BrdU in 95uL incubation buffer with 5uL anti-BrdU-FITC and incubated for 45min at room temperature while shielded from light. Cells were washed once with incubation buffer, followed by cell cyle staining in 185μL PBS + 5μL RNase A + 10uL propidium iodide. Samples were incubated for at least 30min at room temperature (while shielding from light). Samples were diluted with PBS if necessary. BrdU and cell cycle data was collected using a MACSQuant flow cytometer and analyzed using FlowJo software.

HPLC Hap1 cdk1as cells

3x15cm plates with Hap1 *cdk1as* cells were harvested by accutase after their second 1-nm-pp1 arrest in G2, pelleted and washed with PBS. Cells were then resuspended in 1mL Hi-C lysis buffer for 5 minutes at RT. Cells were split in two tubes and washed in 1mL PBS each. 50uL proteinase K was added and incubated for 2 hours at 65ºC, followed by a second dose of 50uL proteinase K for overnight incubation at 65ºC. Each tube was split in 2x 2mL phase lock tubes and samples were cleaned up using double phenol chloroform, followed by ethanol precipitation. DNA pellets were resuspended in TLE and RNA was removed by RNase A treatment at 37ºC for 30 minutes. Samples were cleaned up and washed using 3kDa amicon filter columns. DNA was prepared for HPLC as described in Oomen et al[4]. HPLC spectra were measured at 260 and 279nm. Nucleosides were quantified over three repeats and corrected using published values for molar extinction coefficients at 260 nm (A, T, C, G) and 279nm (BrdU).

Code availability

Hi-C mapping pipeline distiller-nf used to map SisterC data is publicly available at https://github.com/mirnylab/distiller-nf. Downstream analysis tool pairtools and cooltools are available at https://github.com/mirnylab/pairtools and https://github.com/mirnylab/cooltools.

Author contributions

M.E.O. and J.D. conceived and designed the project. M.E.O cultured yeast and

Hap1 cells, generated and analyzed SisterC data and performed flow cytometry experiments. A.K.H. and J.K.W. designed HPLC experiments, A.K.H. performed and analyzed HPLC experiments. B.D. generated the Hap1 Cdk1as cell line and gave input on optimization of the synchronization protocol. M.E.O wrote the manuscript with input from A.K.H. and J.D.

Acknowledgements

We thank all current and former members of the Dekker lab for helpful discussions and suggestions. This work was supported by grants from the National Institutes of Health (HG003143 to J.D.). J.D. is an investigator of the Howard Hughes Medical Institute.

References

1. Nagasaka, K., Hossain, M. J., Roberti, M. J., Ellenberg, J. & Hirota, T. Sister chromatid resolution is an intrinsic part of chromosome organization in prophase. *Nat. Cell Biol.* (2016). doi:10.1038/ncb3353
2. Goloborodko, A., Imakaev, M. V., Marko, J. F. & Mirny, L. Compaction and segregation of sister chromatids via active loop extrusion. *Elife* **5**, 1–20 (2016).
3. Peters, J. M., Tedeschi, A. & Schmitz, J. The cohesin complex and its roles in chromosome biology. *Genes Dev.* **22**, 3089–3114 (2008).
4. Oomen, M., Hedger, A., Watts, J. & Dekker, J. Detecting chromatin interactions along and between sister chromatids with SisterC. *Nat. Methods* (2020). doi:10.1101/2020.03.10.986208
5. Vernis, L., Piskur, J. & Diffley, J. F. X. Reconstitution of an efficient thymidine salvage pathway in Saccharomyces cerevisiae. *Nucleic Acids Res.* **31**, 120e – 120 (2003).
6. Mitter, M. *et al.* Sister-chromatid-sensitive Hi-C reveals the conformation of replicated human chromosomes. *bioRxiv* 2020.03.10.978148 (2020). doi:10.1101/2020.03.10.978148
7. Hochegger, H. *et al.* An essential role for Cdk1 in S phase control is revealed via chemical genetics in vertebrate cells. *J. Cell Biol.* **178**, 257–68 (2007).
8. Falconer, E. *et al.* Identification of sister chromatids by DNA template strand sequences. *Nature* **463**, 93–97 (2010).
9. Falconer, E. *et al.* DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nat. Methods* **9**, (2012).
10. Claussin, C., Porubský, D., Spierings, D. C. J., Halsema, N. & Rentas, S. Genome-wide mapping of sister chromatid exchange events in single yeast cells using Strand-seq. *Elife* (2017).
11. Sanders, A. D., Falconer, E., Hills, M., Spierings, D. C. J. & Lansdorp, P. M. Single-cell template strand sequencing by Strand-seq enables the characterization of individual homologs. *Nat. Protoc.* **12**, 1151 (2017).
12. Jackman, J. & O'Connor, P. M. Methods for Synchronizing Cells at

Specific Stages of the Cell Cycle. *Curr. Protoc. Cell Biol.* 1–20 (2001). doi:10.1002/0471143030.cb0803s00

13. Gibcus, J. H. *et al.* A pathway for mitotic chromosome formation. *Science (80-.).* **359**, eaao6135 (2018).
14. Schmidt, M. & Sebastian, M. Palbociclib—The first of a new class of cell cycle inhibitors. in *Recent Results in Cancer Research* **211**, 153–175 (Springer New York LLC, 2018).
15. Reichard, P. Interactions Between Deoxyribonucleotide And DNA Synthesis. *Annu. Rev. Biochem.* **57**, 349–374 (1988).
16. Bjursell, G. & Reichard, P. Effects of thymidine on deoxyribonucleoside triphosphate pools and deoxyribonucleic acid synthesis in Chinese hamster ovary cells. *J. Biol. Chem.* **248**, 3904–3909 (1973).
17. Oomen, M. E., Hedger, A. K., Watts, J. K. & Dekker, J. SisterC: A novel 3C-technique to detect chromatin interactions between and along sister chromatids. *Res. Sq. Protoc. Exch.* 1–20
18. Oksuz, B. A., Yang, L., Abraham, S., Venev, S. V & Krietenstein, N. Systematic evaluation of chromosome conformation capture assays. 0–42 (2020).
19. Schalbetter, S. A., Neale, M. J., Fudenberg, G., Baxter, J. & Pollard, K. S. Principles of meiotic chromosome assembly revealed in S . cerevisiae. *Nat. Commun.* 1–12 (2019). doi:10.1038/s41467-019-12629-0

Chapter 7

General Discussion

Overview of our findings

The scope of this thesis has been to study the characteristics of mitotic chromosomes. Although perhaps often perceived as condensed and transcriptionally static rod-shaped structures, further exploration unveiled a much more intricate and dynamic organization of mitotic chromosomes. In **chapter 2 and 3**, we describe that mitotic chromosomes maintain certain site-specific characteristics, although many interphase chromatin proteins lose binding during mitosis[1]. For example, many genomic sites such as promoters maintain chromatin accessibility, even when transcription is halted during mitosis[2,3]. Similarly, although CTCF binding itself is lost in mitosis in human differentiated cell lines, CTCF binding sites continue to be marked by histone modifications such as H3K4Me1 and H3K4Me3, and histone variant H2A.z[4]. Along these lines, in **chapter 4** we describe that chromatin characteristics in mitosis can be cell type specific as well as species specific. In contrast to somatic cell lines where CTCF binding is lost entirely, we find that 30-50% of CTCF remains bound in mitosis in mouse embryonic stem cells (mESCs). Additionally, when we investigated mitotic Hi-C data of mouse, chicken and human cell lines, we found that the size of mitotic loops can be highly variable between species. Furthermore, we found that the size of mitotic loops is correlated with the average arm length of the chromosomes. This suggests that dimensions of mitotic chromosomes can be adjusted as a result of loop length regulation. Lastly, in **chapter 5 and 6**, we investigated the interactions between and along sister chromatids with the novel Hi-C technique SisterC. This technique allows us to observe the interplay of dynamic loop extrusion along sister chromatids and the cohesive bonds between sister chromatids formed in mitotic budding yeast[5]. This process requires complex orchestration of extruding cohesin and cohesive cohesin to enable the proper formation of mitotic chromosomes, while simultaneously disentangling the sister chromatids. This process suggests that these mitotic interactions are not static structures but are constantly disassembled and assembled as the cell progresses through the phases of mitosis until the sister chromatids are finally separated in anaphase.

Combining techniques to find new insights

Most observations described in this thesis were made using genomics techniques, such as Hi-C, Cut&Run and ATAC-seq. Although the rise of sequencing and genomics techniques in the past decades opened many doors to new approaches in chromosome biology, it is important to note that they should not be stand-alone techniques. It can be very informative to confirm observations made by genomics techniques using other techniques, for example by microscopy or biochemical approaches. Additionally, the genomics techniques that were used in this thesis

have several shortcomings. First, many genomics techniques such as ChIP-seq and Hi-C require cells to be fixed using formaldehyde[6–8]. As has been published for ChIP-seq and immunofluorescence microscopy, different fixation conditions can change observations that are made using these techniques[9,10]. It is therefore important to confirm findings with techniques that do not require fixed cells, as we did for our study to CTCF binding dynamics in mitosis as described in chapter 3 and 4 using genomics techniques ATAC-seq[11] and Cut&Run[12], and live-cell imaging using super resolution microscopy[4]. Second and possibly more importantly, all genomics techniques used in this thesis research make observations of populations of thousands and even millions of cells. Therefore, cell-to-cell variability cannot be observed directly using these techniques. A less intense ATAC-seq peak or a weaker TAD boundary observed by Hi-C at a given CTCF site suggests that the CTCF site is occupied in only a fraction of cells within the population and remains unbound other cells, however this cannot be concluded with absolute certainty from population-based genomics techniques. Furthermore, any genomics technique is by definition a single snapshot in time and does not allow to track the same cell over a longer period.

In chapter 3, we use live-cell imaging to confirm our finding that CTCF binding is lost in mitosis in human somatic cell lines. Live-cell imaging, such as FRAP and single particle tracking, allowed us to follow cells over time and to directly observe the binding dynamics of CTCF in the cell[4,13]. This would not have been possible without the great improvement of microscopy resolution[14] and the increased ease of making mutant cell lines with fluorescently labelled proteins after the development of CRISPR-Cas9 and other genome engineering tools[15,16]. Using fluorescently labelled CTCF, we were able to confirm our hypothesis that somatic cells lose CTCF binding to CTCF motifs specifically and the entire chromosome in general during mitosis. The finding that CTCF loses binding over the entire chromosome in mitosis could not have been made without the microscopy observations.

In addition to imaging, observations made by Hi-C techniques can greatly benefit from modelling studies[17]. Approaching questions in DNA folding biology with the principles of polymer physics creates new ways to interpret large genome-wide Hi-C datasets. The work described in this thesis was heavily inspired by collaborative work between the Dekker and Mirny laboratories[18,19]. Combining polymer simulations and Hi-C data, these studies proposed that mitotic chromosomes are organized as helical loop arrays[18]. Similarly, polymer modelling following Hi-C data of mitotic yeast chromosomes had suggested that sister chromatids form loops with an average size of ~35kb[20]. However, as conventional Hi-C cannot differentiate interactions between and along sister chromatids, it was not possible to confirm this finding at the time. As

we describe in chapter 5 of this thesis, SisterC and its application to mitotic budding yeast allows to differentiate cohesive cohesin and loop forming extruding cohesin. Our SisterC observations confirm the predictions that were done by polymer simulations years prior to the development of SisterC[5]. Finally, in chapter 4 we describe the scenario where CTCF remains bound on mitotic chromosomes in mouse stem cells, although no TADs and CTCF-CTCF dependent loops are observed by Hi-C. This strongly suggests that vertebrate mitotic loop extruders condensin I and II are not blocked by CTCF, unlike their interphase counterpart cohesin. Although this is a new observation by Hi-C, polymer modeling studies had already suggested that it is unlikely that mitotic loop extruders are blocked site-specifically by a boundary element such as CTCF[21,22]. These examples illustrate the importance of collaboration and that by combining techniques new scientific insights can be acquired, which might not be accessible when limited by a certain technology alone.

Function of mitotic bookmarks

In chapter 3, we describe the maintenance of histone modifications H3K4Me1/3 and histone variant H2A.z at sites that lose CTCF binding sites during mitosis[4]. We hypothesize that these mitotic bookmarks help CTCF to regain binding upon mitotic exit and forming of interphase structures such as TADs and CTCF dependent loops. Similarly, mitotic maintenance of histone modification H3K9Me3 has been proposed to aid in the re-binding of heterochromatin protein HP1 in the next G1[23]. Although correlations of histone modifications and mitotically lost chromatin binding proteins have been observed, it is challenging to systematically prove that these histone modifications are indeed essential to restore binding of proteins such as CTCF or HP1 to the chromatin as cells enter G1. First, in order to study the functionality of mitotic bookmarks upon G1 entry by genomics techniques, highly synchronous release from prometaphase arrest is essential. Although recent reports make use of chemical inhibitors or sorting of cells by flow cytometry to obtain a time course of cells exiting mitosis[24,25], it remains a big challenge to study cells as they transition from mitosis to G1. Second, studying the role of histone modifications and variants can be very complex, as these marks are known to be present at different regulatory regions throughout the genome. For example, H3K4Me1 not only marks CTCF sites, but has also been described to be present at enhancer sites[26]. Disrupting marks such as H3K4Me1 genome-wide, will not only remove the mitotic bookmarks at CTCF sites, but will also have a more dramatic effect on general chromatin characteristics, making it difficult to interpret the results. However, the fast development of CRISPR-cas9 technology also led to tools that can be used to locally modify epigenetic marks. By fusing dCas9 to either gene activating domains

(CRISPRa) or gene interference domains (CRISPRi), transcriptional activity of specific genomic regions can be modulated[16]. Although these are still a very challenging techniques and not yet possible at the timescale needed to study mitotic exit, the use of dCas9 guided histone modifying proteins might aid the study to the functionality of mitotic bookmarks at specific genomic locations in the near future[27].

Along these lines, we describe in chapter 4 that mouse embryonic stem cells maintain CTCF binding at 30-50% of CTCF sites during mitosis. Similar observations were made in mouse pluripotent cells[28] and mouse progenitor cells[25]. However, we show using Hi-C that there are no remaining CTCF mediated loops or TAD boundaries in mitosis. As extruding cohesin is mostly dissociated from chromosomes by late prometaphase[18,29], this results suggests that bound CTCF cannot block loop extrusion by condensin I or II. It is still unclear why partial binding of CTCF is maintained during mitosis and what the function of mitotically bound CTCF could be. We and others hypothesize that bookmarked CTCF sites can mediate transcription of genes in early G1[25,28]. This could be studied by temporarily removing CTCF from the cell entirely, for example by using the auxin-based inducible degron (AID) system[30] during mitotic exit followed by RNA-seq experiments to observe any changes in gene expression. However, as CTCF is an essential architectural protein[31], it will be very difficult to separate the specific function of bookmarked CTCF on gene expression of early transcribing genes from the genome wide effects due to lack of TADs and loop formation as the cells progress into G1.

Studying sister chromatid conformation

In chapter 5 we introduce our novel Hi-C technique, SisterC[5]. In addition to our technique, the Gerlich laboratory also released a new sister chromatid sensitive Hi-C technique called scs-HiC[32]. Scs-HiC is based on 4sT incorporation, followed by the chemical conversion of 4sT into C during regular Hi-C library preparation. Both SisterC and scs-HiC will open up the possibility to study sister chromatid organization from G2 cells up to sister separation in anaphase. However, in order to adapt and apply these techniques, several optimizations will be necessary. First, in the case of SisterC the depletion of BrdU-containing strands after UV/Hoechst in the SisterC protocol is not complete. We address this in chapter 6, although no improvement to the protocol was made as of now. In contrast to the SisterC protocol, scs-HiC does not require depletion of strands, but instead benefits from the chemical conversion of 4sT in the newly replicated strands. Although the differentiation of inter- and intra-sister interaction is more accurate as it relies on a positive read-out, it is very inefficient and therefore costly as it requires many sequencing reads. This highlights the importance of the second step that requires optimization for both SisterC and scs-

HiC. In order for both SisterC and scs-HiC to work efficiently, cells need to efficiently incorporate thymidine analogues BrdU or 4sT. As we show in chapter 5 and 6, both yeast and mammalian cell line Hap1 incorporate BrdU at high efficiency when cultured in the right conditions. Contrary, 4sT is much less efficiently incorporated in mammalian cells[32]. This is most likely due to suboptimal uptake and phosphorylation of 4sT nucleoside prior to the incorporation during S-phase. It has been shown that 4sTTP is readily incorporated by polymerases in vitro, suggesting that in vivo cells can indeed successfully incorporate 4sTTP as a thymidine analogue[33]. An approach to optimize 4sT incorporation is by the design of prodrugs, which can improve the uptake by the cell and circumvent the initial phosphorylation reaction by the cell that converts 4sT to 4sTMP[33]. Furthermore, increased efficiency of 4sT incorporation will allow for the exploration of sister specific chromosome organization during S-phase, when sister chromatids are newly formed. As SisterC relies on efficient BrdU incorporation, followed by the depletion of the BrdU containing strands, this cannot be used to study S-phase, when only replicated regions will contain BrdU. In contrast, scs-HiC would be able to study sister chromatid interacions in replicated regions during S-phase, however the need to obtaining sufficient sequence depth makes this approach too costly due to low efficiency.

Future applications of SisterC

It will be very interesting to apply SisterC to mammalian cells. In chapter 6 we suggest a protocol for synchronization and BrdU incorporation in the human cell line Hap1 which has modified *cdk1as* to allow for easy synchronization in G2. This will allow to study the orchestration of chromosome condensation and sister chromatid separation as cells progress from G2 to prometaphase. Additionally, it will be possible to study the role of proteins that are known to be essential in mitosis, in the context of sister chromatid organization. An example of such an essential mitotic protein is topoisomerase. Topoisomerase II is known to play an important role in sister chromatid separation[34,35]. For example, when topoisomerase II is chemically inhibited, an increased rate of sister chromatid exchanges (SCEs) and chromosomal aberrations can be observed, suggesting unresolved entanglement of sister chromatids during mitosis[36,37]. Although this function of topoisomerase II has been known for several years, SisterC will allow to detect the sister chromatid interactions and entanglements that eventually lead to increased SCEs in a more direct fashion. Similarly, it has been described that condensin aids the decatenation of sister chromatids by topoisomerase[38]. Studying these phenomena using SisterC after genetic manipulation or chemical disruption of proteins will help understand the complex interplay of condensation and segregation of sister chromatids in mitosis.

Along these lines, as both described in chapter 5[5] and in the study by Mitter et al[32], distinct cohesin complexes form both the cohesive bonds between sister chromatids and extruding loops along sister chromatids in mitotic yeast and G2 mammalian cells. It has been proposed that these cohesin subunits can differ in subunit composition and posttranslational modifications. It is however still unknown how loop formation by extruding cohesin in G2 eventually is handed over to mitotic extruding factors condensin I and II in vertebrates[18], while simultaneously maintaining the function of cohesive cohesin until anaphase. Adapting the SisterC protocol to vertebrate cell lines will be essential to study these phenomena and gain a better understanding of the distinct behavior of cohesins and condensins.

Finally, in addition to mitosis, regulation of sister chromatid interactions and alignment are of great importance during meiosis as well. Similarly to mitosis, it has been proposed that there are distinct cohesin complexes that mediate the extruding loops and the bonds between sister chromatids in meiosis[39,40]. Although it can be challenging to study meiosis in mammalian systems, in particular using genomic approaches, this is much more feasible in budding yeast[41]. An additional challenge of studying chromosome folding in meiosis is that interactions not only occur between sister chromatids, but also between chromosome homologues[39,41]. In previous studies, it was shown that chromosome homologues can be distinguished in Hi-C data by crossing genetic distinct strains, which allows to separate Hi-C reads from maternal or paternal alleles based on their unique SNPs[42,43]. Although this data will be highly complex to interpret, it will be very informative to observe chromosome organization in meiotic cells using sister chromatid sensitive Hi-C techniques such as scs-HiC and sisterC.

Concluding remarks

Although chromosomes seem to have very unique characteristics and organization in mitosis, we can use this phase of the cell cycle to gain knowledge about the role of chromatin marks in cell type identity. As chromosomes undergo dramatic conformational and epigenetic changes in mitosis, bookmarks remain on the chromatin, which aids quick reestablishment of a functional interphase nucleus. In the coming years, it will be interesting to learn which mitotic bookmarks are essential for the cell and how exactly these bookmarks mediate rebinding of chromatin proteins and refolding of interphase chromosome organization upon mitotic exit. Additionally, the challenges chromosomes face as they disentangle in preparation of sister chromatid separation in anaphase will help us understand similar topological stresses chromosomes are exposes to throughout the cell cycle, for example during transcription or DNA repair. Combined, this will bring better understanding on the

importance of epigenetic marks and chromatin regulating proteins, and how they influence cell characteristics, not just throughout the cell cycle, but also between different cell types and cell states in general.

References

1. Oomen, M. E. & Dekker, J. Epigenetic characteristics of the mitotic chromosome in 1D and 3D. *Crit. Rev. Biochem. Mol. Biol.* **0**, 1–20 (2017).
2. Martinez-Balbas, M. A., Dey, A., Rabindran, S. K., Ozato, K. & Wu, C. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* **83**, 29–38 (1995).
3. Hsiung, C. C. *et al.* Genome accessibility is widely preserved and locally modulated during mitosis. 1–29 (2015). doi:10.1101/gr.180646.114
4. Oomen, M. E., Hansen, A. S., Liu, Y., Darzacq, X. & Dekker, J. CTCF sites display cell cycle–dependent dynamics in factor binding and nucleosome positioning. *Genome Res.* 1–14 (2019). doi:10.1101/gr.241547.118.
5. Oomen, M., Hedger, A., Watts, J. & Dekker, J. Detecting chromatin interactions along and between sister chromatids with SisterC. *Nat. Methods* (2020). doi:10.1101/2020.03.10.986208
6. Belaghzal, H., Dekker, J. & Gibcus, J. H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* **123**, 56–65 (2017).
7. Oksuz, B. A., Yang, L., Abraham, S., Venev, S. V & Krietenstein, N. Systematic evaluation of chromosome conformation capture assays. 0–42 (2020).
8. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–93 (2009).
9. Teves, S. S. *et al.* A Dynamic Mode of Mitotic Bookmarking by Transcription Factors. *Elife* 066464 (2016). doi:10.1101/066464
10. Festuccia, N. *et al.* Transcription factor activity and nucleosome organization in mitosis. *Genome Res.* **29**, 250–260 (2019).
11. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–8 (2013).
12. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* 1–35 (2017). doi:10.1101/097188
13. Hansen, A. S. *et al.* Robust model-based analysis of single-particle tracking experiments with Spot-On. *Elife* **7**, 1–33 (2018).
14. Schermelleh, L. *et al.* Super-resolution microscopy demystified. *Nat. Cell Biol.* **21**, 72–84 (2019).
15. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science (80-.).* **346**, (2014).
16. Wang, F. & Qi, L. S. Applications of CRISPR Genome Engineering in Cell Biology. *Trends Cell Biol.* **26**, 875–888 (2016).
17. Dekker, J., Marti-Renom, M. a & Mirny, L. a. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat. Rev.*

Genet. **14**, 390–403 (2013).
18. Gibcus, J. H. *et al.* A pathway for mitotic chromosome formation. *Science (80-.).* **359**, eaao6135 (2018).
19. Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342**, 948–53 (2013).
20. Schalbetter, S. A. *et al.* SMC complexes differentially compact mitotic chromosomes according to genomic context. *Nat. Cell Biol.* **19**, 1071–1080 (2017).
21. Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 1–12 (2016). doi:10.1101/024620
22. Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A. & Mirny, L. A. Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harb. Symp. Quant. Biol.* **LXXXII**, 034710 (2018).
23. Hirota, T., Lipp, J. J., Toh, B.-H. & Peters, J.-M. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438**, 1176–80 (2005).
24. Abramo, K. *et al.* A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. *Nat. Cell Biol.* **21**, 1393–1402 (2019).
25. Zhang, H. *et al.* Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* **576**, 158–162 (2019).
26. Local, A. *et al.* Identification of H3K4me1-associated proteins at mammalian enhancers. *Nat. Genet.* **50**, 73–82 (2018).
27. Thakore, P. I. *et al.* Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* **12**, 1143–1149 (2015).
28. Pelham-Webb, B. *et al.* Mitotic retention of H3K27 acetylation promotes rapid topological and transcriptional resetting of stem cell-related genes and enhancers upon G1 entry. *bioRxiv* (2020). doi:10.1101/2020.06.02.130104
29. Ohta, S. *et al.* The Protein Composition of Mitotic Chromosomes Determined Using Multiclassifier Combinatorial Proteomics. *Cell* **142**, 810–821 (2010).
30. Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. & Kanemaki, M. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods* **6**, 917–922 (2009).
31. Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-944.e22 (2017).
32. Mitter, M. *et al.* Conformation of sister chromatids in the replicated human genome. *Nature* **586**, 139–144 (2020).
33. Hedger, A. K. *et al.* Progress toward an amplifiable metabolic label for DNA: Conversion of 4-thiothymidine (4sT) to 5-methyl-2′-deoxycytidine and synthesis of a 4sT phosphorodiamidate prodrug. *Can. J. Chem.* **96**, 636–645 (2018).
34. Vagnarelli, P. Mitotic chromosome condensation in vertebrates. *Exp. Cell Res.* **318**, 1435–1441 (2012).
35. Bar-Ziv, R., Voichek, Y. & Barkai, N. Chromatin dynamics during DNA replication. *Genome Res.* gr.201244.115- (2016). doi:10.1101/gr.201244.115
36. Cortés, F., Piñero, J. & Ortiz, T. Importance of replication fork progression

for the induction of chromosome damage and SCE by inhibitors of DNA topoisomerases. *Mutat. Res. Lett.* **303**, 71–76 (1993).
37. Dillehay, L. E., Jacobson-Kram, D. & Williams, J. R. DNA topoisomerases and models of sister-chromatid exchange. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **215**, 15–23 (1989).
38. Charbin, A., Bouchoux, C. & Uhlmann, F. Condensin aids sister chromatid decatenation by topoisomerase II. *Nucleic Acids Res.* **42**, 340–348 (2014).
39. Silva, M. C. C. *et al.* Wapl releases Scc1-cohesin and regulates chromosome structure and segregation in mouse oocytes. *J. Cell Biol.* **219**, (2020).
40. Gassler, J. *et al.* A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* **36**, 3600–3618 (2017).
41. Schalbetter, S. A., Neale, M. J., Fudenberg, G., Baxter, J. & Pollard, K. S. Principles of meiotic chromosome assembly revealed in S . cerevisiae. *Nat. Commun.* 1–12 (2019). doi:10.1038/s41467-019-12629-0
42. AlHaj Abed, J. *et al.* Highly structured homolog pairing reflects functional organization of the Drosophila genome. *Nat. Commun.* **10**, 1–14 (2019).
43. Giorgetti, L. *et al.* Structural organization of the inactive X chromosome in the mouse. *Nature* **535**, 575–579 (2016).

Addendum

Summary
Nederlandse samenvatting
Abbreviations
Curriculum Vitae
List of Publications
PhD portfolio
Acknowledgements

Summary

Chromosomes are the structural units that contain the information of our genome. Therefore, chromosome folding is tightly modulated in order to enable transcriptional regulation, protection from DNA damage and maintenance of epigenetic characteristics. Seemingly counterintuitive, chromosomes undergo dramatic conformational changes as cells progress through the cell cycle. After DNA replication, sister chromatids need to detangle and ultimately condense in order to enable proper separation into the two new daughter cells in mitosis. Even more striking, iconic interphase structures such as TADs and compartments are dissolved in mitosis and mitotic chromosomes are organized in a helical loop array mediated by loop extruding machineries. However, mitotic bookmarks such as histone modifications, histone variants and even certain chromatin binding factors are maintained, which could enable the fast reestablishment of the interphase chromosome organization upon mitotic exit.

The scope of this thesis has been to explore the characteristics of mitotic chromosomes. On a small scale, we have studied the binding dynamics of CTCF, a key architectural protein in interphase chromosomes, and the mitotic bookmarks around CTCF binding sites in differentiated cells and stem cells. On a larger genomic scale, we observed interactions between and along sister chromatids in budding yeast, for which we developed a novel Hi-C technique, SisterC.

Chapter 1 presents a general introduction to this thesis and the scientific context in which this research has been conducted. Additionally, it sets out the scope of the thesis and its individual chapters.

Chapter 2 gives a literature overview of relevant work at the start of this PhD research. We describe the epigenetic characteristics of mitotic chromosomes. As cells progress into mitosis, transcription stalls, many chromatin factors lose binding and chromosome conformation rearranges into a randomly positioned array of loops. However, chromosomes can reestablish their interphase identity upon mitotic exit very rapidly. In this chapter we explore different ways cells maintain their epigenetic memory during mitosis by bookmarks such as histone modifications and variants, chromatin binding factors, histone remodelers and non-coding RNA.

In **chapter 3** we describe our study to the binding dynamics of CTCF in differentiated cell lines during mitosis using 4 different experimental approaches. First, using 5C

we observe that TADs and CTCF mediated loops are absent in mitotic cells. Next, we show loss of chromatin accessibility and binding footprint at CTCF motifs using ATAC-seq. We confirm loss of CTCF binding in a more direct fashion by performing Cut&Run. Although CTCF binding is lost in mitosis, we find that histone modifications such as H3K4me1/3 and histone variant H2A.z are maintained. Lastly, we confirm our genomics observations using live-cell imaging. FRAP and single-particle tracking show that CTCF chromatin binding is lost in prometaphase.

Chapter 4 dives deeper into the characteristics of mitotic chromosomes and the differences between stem cells and differentated cells and between species. By reanalyzing previously published ATAC-seq data, we find that mouse stem cells can maintain partial CTCF binding in mitosis. However, no TADs and CTCF meditated loops can be observed by Hi-C, both at individual genomic locations and by aggregating Hi-C signal of many loci as pile ups plots. This suggests that eventhough a fraction of the CTCF sites maintain binding, mitotic loop extrusion by condensin I and II cannot be blocked by CTCF. Additionally, when investigating mitotic chromosomes in mouse, human and chicken Hi-C data, we find that the average size of mitotic loops can differ between species. This is correlated with the average length of the q-arm of the chromosomes in these species. We interpret this result such that when a species has longer chromosome arms, the level of condensation is greater, which could aid the proper seperation of sister chromatids in mitosis.

In **chapter 5** we introduce a novel Hi-C method called SisterC. By conventional Hi-C, it is impossible to differentiate whether interactions between the same genomic locations occur within the same sister chromatid or between sister chromatids. SisterC uses BrdU incorporation followed by UV/Hoechst treatment to introduce single strand nicks, which allows to separate inter-sister and intra-sister interactions based on read orientation after sequencing and mapping of reads. We apply this technique to budding yeast synchronized in mitosis. Both inter-sister connections and intra-sister loops are mediated by cohesin in yeast. We find that sister chromatids are not perfectly alligned in mitotic yeast chromosomes. SisterC observes that along the chromosome arms, cohesin mediates a bond between the two sister chromatids every 35kb on average. Additionally extruding cohesin forms loops along sister chromatids up to 50kb size. We find that these loops are not necessarily positioned at the same genomic location on both sister chromatids. This creates an offset of 5-25kb at the positions where cohesive cohesin mediates the interactions between the two sister chromatids.

In the current SisterC protocol, the depletion of strands containing BrdU is sufficient, but not complete. **Chapter 6** describes different ways our SisterC protocol can be improved. Futhermore, we adapt the SisterC protocol to allow for usuage in mammalian cell line Hap1 *cdk1as*. We test a synchronization protocol for Hap1 *cdk1as* cells, which enables efficient BrdU incorporation during S-phase, as well as synchronization in G2-phase, mitosis and the subsequent G1. Using this protocol, we find over 80% BrdU incorporation by HPLC, which promises the succesful adaptations of SisterC in mammalian cells.

Finally, **Chapter 7** places our results in context of the latest developments in the field of epigenetic characteristics of mitotic chromosomes. Future directions and novel research paths are discussed that can contribute to a better understanding of mitotic bookmarking and the conformational changes which mitotic chromosomes undergo.

Samenvatting

Chromosomen zijn de structurele eenheden die onze genomen bevatten. Het proces van chromosoomvouwing wordt derhalve nauwlettend gemoduleerd om de regulatie van transcriptie, bescherming tegen DNA-schade en het behoud van epigenetische kenmerken mogelijk te maken. Ogenschijnlijk tegenstrijdig hieraan ondergaan chromosomen dramatische conformatieveranderingen tijdens de celcyclus. Nadat het DNA gerepliceerd is moeten de twee kopieën (de zusterchromatiden) ontward worden en uiteindelijk condenseren om een goede splitsing naar de twee nieuwe dochtercellen tijdens de mitose mogelijk te maken. Opvallend is dat typische interfasestructuren zoals TADs en compartimenten verdwijnen tijdens mitose, en dat chromosomen worden gereorganiseerd in een spiraalvormige reeks van lussen die wordt gemedieerd door lusvormende machines. Mitotische markeringen zoals histonmodificaties, histonvarianten en zelfs bepaalde chromatinebindende factoren blijven echter behouden, wat mogelijk een rol speelt in het snelle herstel van de organisatie van interfasechromosomen na mitose.

Het doel van dit proefschrift was om de kenmerken van mitotische chromosomen te onderzoeken. Op een gedetailleerd niveau hebben wij de bindingsdynamiek van CTCF, een belangrijk architecturaal eiwit in interfasechromosomen, bestudeerd. Daarnaast onderzochten we de mitotische markeringen rond CTCF-bindingsplaatsen in zowel gedifferentieerde cellen als stamcellen. Op hoger genomisch niveau hebben we interacties waargenomen tussen en binnen zusterchromatiden in gistcellen. Hiervoor ontwikkelden we een nieuwe Hi-C-techniek: SisterC.

Hoofdstuk 1 geeft een korte inleiding tot dit proefschrift en de context waarbinnen dit onderzoek is uitgevoerd. Daarnaast worden de strekking van het proefschrift en de afzonderlijke hoofdstukken ervan beschreven.

Hoofdstuk 2 geeft een literatuuroverzicht van relevant werk aan het begin van dit promotieonderzoek. We beschrijven de epigenetische kenmerken van mitotische chromosomen. Wanneer cellen de eerste fases van mitose doorlopen stagneert de gentranscriptie, ontbinden veel chromatinefactoren en herschikt de chromosoomorganisatie zich van een celtype-specifieke conformatie naar een willekeurig gepositioneerde reeks lussen. Chromosomen kunnen hun interfase-identiteit echter zeer snel herstellen na de mitose. In dit hoofdstuk beschrijven we verschillende manieren waarop cellen hun epigenetisch geheugen behouden tijdens mitose door middel van epigenetische markeringen, zoals histonmodificaties en

-varianten, chromatinebindende factoren, histon *remodelers* en niet-coderend RNA.

In **hoofdstuk 3** beschrijven we onze studie van de bindingsdynamiek van CTCF in gedifferentieerde cellijnen tijdens mitose. We doen dit met behulp van 4 verschillende experimentele benaderingen. Als eerste observeren we met behulp van 5C dat TADs en CTCF-gemedieerde lussen afwezig zijn in mitotische cellen. Vervolgens laten we met behulp van ATAC-seq het verlies van de chromatinetoegankelijkheid en bindingssignatuur van gebonden CTCF-motieven zien. Daarnaast bevestigen we het verlies van CTCF-binding op een directe manier door middel van Cut&Run. Hoewel CTCF-binding verloren gaat in mitose, vinden we dat histonmodificaties zoals H3K4me1/3 en histonvariant H2A.z behouden blijven. Tenslotte bevestigen we onze *genomics* observaties met behulp van *live-cell imaging*. FRAP en *single-particle tracking* bevestigen in levende cellen dat de binding van CTCF-chromatine verloren gaat tijdens mitosis.

Hoofdstuk 4 gaat dieper in op de kenmerken van mitotische chromosomen en de verschillen tussen stamcellen en gedifferentieerde cellen, en de verschillen tussen gewervelde diersoorten. Met behulp van een nieuwe analyse van gepubliceerde ATAC-seq data observeren we dat muizenstamcellen deels CTCF-binding behouden tijdens mitose. Ondanks dit gedeeltelijk behoud worden er geen TADs en CTCF-gemediteerde lussen waargenomen in Hi-C-data, zowel op individuele genomische locaties als in Hi-C-aggregaties. Dit suggereert dat mitotische lusvorming door condensin I en II niet kan worden verhinderd door CTCF. Bovendien vinden we door het vergelijken van mitotische chromosomen in Hi-C-data van muizen, mensen en kippen dat de gemiddelde grootte van de lussen verschilt tussen soorten. Dit correleert met de gemiddelde lengte van de q-arm van de chromosomen van deze soorten. We interpreteren dit resultaat zo dat naarmate een organisme langere chromosoomarmen heeft, het niveau van condensatie groter is. Dit zou mogelijk kunnen helpen bij de correcte splitsing van zusterchromatiden bij mitose.

In **hoofdstuk 5** introduceren we een nieuwe Hi-C-methode, genaamd SisterC. Bij het gebruik van de conventionele Hi-C techniek is het niet mogelijk om onderscheid te maken tussen genomische interacties binnen dezelfde zusterchromatide (intrazuster) of tussen de twee verschillende zusterchromatiden (interzuster). SisterC gebruikt BrdU-incorporatie, gevolgd door UV/Hoechst-behandeling om enkelstrengs breuken te introduceren, waardoor inter- en intrazusterinteracties kunnen worden gescheiden op basis van strengoriëntatie na *high-throughput DNA sequencing*. We passen deze techniek toe op mitotisch gesynchroniseerd brouwersgist. In gist

worden zowel interzuster connecties als intrazuster chromosoomlussen gemedieerd door het eiwit cohesin. Met behulp van SisterC observeren wij dat cohesin gemiddeld elke 35 kb langs de chromosoomarmen een band tussen de twee zusterchromatiden vormt langs de chromosoom armen. Daarnaast vormt cohesin lussen langs de zusterchromatiden, tot een grootte van ten hoogste 50 kb. We merken op dat deze lussen zich niet per se op dezelfde genomische locatie op beide zusterchromatiden bevinden. Hierdoor zijn de zusterchromatiden niet perfect uitgelijnd, hetgeen resulteert in een gemiddelde afstand van 5-25 kb tussen de plaatsen waar interacties tussen de twee zusterchromatiden gemedieerd worden door bindend cohesin.

In het huidige SisterC-protocol is de verwijdering van strengen die BrdU bevatten voldoende, maar niet volledig. **Hoofdstuk 6** beschrijft verschillende manieren waarop ons SisterC-protocol kan worden verbeterd. Daarnaast passen we het SisterC-protocol voor gebruik in de menselijk cellijn Hap1 *cdk1as*. We testen een synchronisatieprotocol voor Hap1 *cdk1as*-cellen, dat efficiënte BrdU-opname tijdens de S-fase mogelijk maakt, naast synchronisatie in de G2-fase, mitose en de daaropvolgende G1. Met behulp van dit protocol vinden we ruim 80% BrdU-opname met HPLC, wat succesvolle toepassing van SisterC in zoogdiercellen voorspelt.

Hoofdstuk 7, tenslotte, plaatst onze resultaten in de context van de laatste ontwikkelingen op het gebied van epigenetische kenmerken van mitotische chromosomen. Toekomstige onderzoeksrichtingen en -vragen worden besproken, die mogelijk kunnen bijdragen tot een beter begrip van mitotische markeringen en de conformatieveranderingen die mitotische chromosomen ondergaan.

List of Abbreviations

3C	Chromosome conformation capture
5C	Chromosome conformation capture carbon copy
A	Adenosine
ATAC-seq	Assay for transposase-accessible chromatin using sequencing
Bp	Base pair
BRD4	Bromodomain 4 protein
BrdU	Bromodeoxyuridine
C	Cytidine
CENP-A	Centromere protein A
ChIP-seq	Chromatin Immunoprecipitation sequencing
CPC	Chromosomal passenger complex
CTCF	CCCTC-binding factor)
CUT&RUN	Cleavage under targets and release using nuclease
DamID	DNA adenine methyltransferase identification
DHS	DNase I hypersensistive site
DNA	Deoxyribonucleic acid
EMANIC	EM-assisted nucleosome interaction capture
FISH	Fluorescence in situ hybridization
FRAP	Fluorescence recovery after photobleaching
G	Guanosine
Gb	Gigabase (1,000,000,000 basepairs)
gDNA	genomic DNA
H3K27ac	Histone 3 lysine 27 acetylation
H3K36me3	Histone 3 Lysine 36 trimethylation
H3K4me1	Histone 3 Lysine 4 monomethylation
H3K4me3	Histone 3 Lysine 4 trimethylation
H3K9me3	Histone 3 Lysine 9 trimethylation
H3S10ph	Histone 3 Serine 10 phosphorylation
H3T3ph	Histone 3 Threonine 3 phosphorylation
HAT	Histone acetyl transferase
HeLa	Henrietta Lacks cell line
HFF	Human foreskin fibroblast
HP1	Heterochromatin protein 1
IF	Immuno fluorescence
Kb	Kilobase (1,000 basepairs)
LADs	Lamina interacting domains
Mb	Megabase (1,000,000 basepairs)
mESC	mouse embryonic stem cells
mPCs	mouse pluripotent cells
ncRNA	non-coding RNA
PCR	polymerase chain reaction
RNA	Ribonucleic acid
SCEs	Sister chromatids exchanges
scs-HiC	Sister chromatid sensitive Hi-C
SMC	Structural maintenance of chromosomes protein
SPT	Single particle tracking
T	Thymidine
TAD	Topologically associated domain
TF	Transcription factor
TSS	Transcription start site

Curriculum Vitae

Personal details

Name	Marlies E. Oomen
Date of birth	05-04-1993
Place of birth	Utrecht (The Netherlands)

Education

2015-2021	PhD program at Erasmus University Rotterdam under supervision of prof. dr. Job Dekker and prof. dr. Bas van Steensel
2013-2015	Master's degree Medical Pharmaceutical Drug Innovation at Rijksuniversiteit Groningen
2010-2013	Bachelor's degree Molecular Life Sciences at Rijksuniversiteit Groningen

Research

2015-2021	PhD student in the laboratory of prof. dr. Job Dekker at University of Massachusetts Medical School, Worcester MA, USA. "Exploring mitotic chromosomes: from epigenetic bookmarking to sister chromatid conformation"
2014-2015	Visiting student in the laboratory of prof. dr. Job Dekker at University of Massachusetts Medical School, Worcester MA, USA. "Studying the characteristics of the mitotic chromosome"
2014	Research assistant in the laboratory of dr. Christian Riedel at Rijksuniversiteit Groningen and University Medical Center Groningen
2014	Master student intern in the laboratory of dr. Christian Riedel at Rijksuniversiteit Groningen and University Medical Center Groningen "The role of histone modifications and variants in aging"
2013	Bachelor student intern in the laboratory of prof. dr. Jan-Willem Veening at Rijksuniversiteit Groningen "Exploring the grey area of the effect of antibiotics in *S. pneumonia*"

Publications

Betul Akgol Oksuz, Liyan Yang, Sameer Abraham, Sergey V. Venev, Nils Krietenstein, Krishna Mohan Parsi, Hakan Ozadam, **Marlies E. Oomen**, Ankita Nand, Hui Mao, Ryan MJ Genga, Rene Maehr, Oliver J. Rando, Leonid A. Mirny, Johan Harmen Gibcus, Job Dekker. **Systematic evaluation of chromosome conformation capture assays,** BioRxiv (2021).

Marlies E Oomen, Adam K Hedger, Jonathan K Watts, Job Dekker. **Detecting chromatin interactions between and along sister chromatids with SisterC.** Nature Methods (2020).

Marlies E Oomen, Anders S Hansen, Yu Liu, Xavier Darzacq, Job Dekker. **CTCF sites display cell cycle–dependent dynamics in factor binding and nucleosome positioning**, Genome Research (2019).

Elizabeth H Finn, Gianluca Pegoraro, Hugo B Brandão, Anne-Laure Valton, **Marlies E Oomen**, Job Dekker, Leonid Mirny, Tom Misteli. **Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization**, Cell (2019).

Adam K Hedger, **Marlies E Oomen**, Victor Liu, Michael P Moazami, Nicholas Rhind, Job Dekker, Jonathan K Watts. **Progress toward an amplifiable metabolic label for DNA: conversion of 4-thiothymidine (4sT) to 5-methyl-2′-deoxycytidine and synthesis of a 4sT phosphorodiamidate prodrug**, Canadian Journal of Chemistry (2018).

Marlies E Oomen, Job Dekker. **Epigenetic characteristics of the mitotic chromosome in 1D and 3D**, Critical Reviews in Biochemistry and Molecular Biology (2017).

PhD portfolio

Name of the PhD student	Maria Elisabeth Oomen
Research location	University of Massachusetts Medical School
Affiliation	Programs in Systems Biology
Graduate School	Erasmus University Rotterdam
PhD period	2015 – 2021
Promotors	Prof. dr. Job Dekker and prof. dr. Bas van Steensel

Training credits

2020	Virtual Conference “The Four-Dimensional Genome” (EMBL)
2016-2020	Monthly seminar series “Epigenetics club” (UMass Medical School)
2019	Statistical Methods for Functional Genomics (CSHL)
2019	Conference “Evolution, Structure and Function of Chromosomes High Order Structure” (Institut Pasteur)
2019	In-house Python/Jupyter class by dr. Sergey Venev (Dekker Lab)
2018	Conference “Chromatin Architecture and Chromosome Organization” (Keystone)
2017	In-house R and tidyverse class by dr. Filipe Cadete (Dekker Lab)
2017	Supervision of PhD student Jack Huey (UMass Medical School)
2016	Molecular Biophysics (UMass Medical School)

Talks

Detecting chromatin interactions along and between sister chromatids with SisterC
JRNLclub. March 16, 2021. https://jrnlclub.org/research-films/chromatin-interactions-sisterc

Detecting chromatin interactions along and between sister chromatids with SisterC
EMBO Meeting “The Four-Dimensional Genome”. March 30-31, 2020, virtual due to COVID-19.

Mapping interactions between sister A and sister B with SisterC
Departmental seminar series “Science on Tap”. January 17, 2020, UMass Medical School.

Seeing sister chromatids with SisterC
UMass Seminar Series “Epigenetics Club”. January 8, 2020, UMass Medical School.

Cell cycle dynamics of CTCF binding and its relation to chromosome organization
François Jacob Conference: Evolution, Structure and Function of Chromosomes High Order Structure (Institut Pasteur). June 4-7, 2019, Paris, France.

Cell cycle dynamics of CTCF binding and its relation to chromosome organization
CCSB-PSB 2018 Retreat. Sept 17-19, 2018, Gloucester, MA.

Cell cycle dynamics of CTCF binding and its relation to chromosome organization
Keystone Symposium on Chromatin Architecture and Chromosome Organization (X5). March 23-27, 2018, Whistler, BC, Canada.

Cell cycle dynamics of CTCF binding and its relation to chromosome organization
UMass Seminar Series “Epigenetics Club”. January 10, 2018, UMass Medical School.

Cell cycle dynamics of CTCF binding and its relation to chromosome organization
4DN Center Meeting. October 26, 2018. MIT, Boston.

Epigenetic Characteristics of the Mitotic Chromosome
Departmental seminar series “Science on Tap”. February 3, 2017, UMass Medical School.

Posters
Cell cycle dynamics of CTCF binding and its relation to chromosome organization
HHMI science meeting November 13-15 2018, Janelia Research Campus, Ashburn, VA.

Cell cycle dynamics of CTCF binding and its relation to chromosome organization
Keystone Symposium on Chromatin Architecture and Chromosome Organization (X5). 23-27 March 2018, Whistler, BC, Canada.

Additional activities
Co-organizer of UMass Medical School Seminar Series ‘Epigenetics Club’ (2016-2017)

Acknowledgements

I can barely believe it, but I am writing the last and most important section of my thesis, the acknowledgments. Although research can feel lonely from time to time, I have been incredibly fortunate to have been surrounded by so many wonderful, smart and kind people throughout my PhD. With the risk of forgetting people (and listing some people twice), I want to thank several people for supporting me during my PhD in many different ways.

First and foremost, my promotors prof. dr. **Job Dekker** and prof. dr. **Bas van Steensel**. Job, thank you for welcoming me back to your lab as a PhD student (even after I told everyone that I would never, *ever* return to Worcester after my cold and snowy master's internship in 2014-15). You have created an environment where I could follow my scientific interests independently, while also being available for guidance and advice when I needed it. I am so happy I was able to do my PhD in your lab, giving me the opportunity to learn from you and the team of scientists you brought together in your lab and in collaborations. Among many other skills I've learned during my PhD, I'm thankful for the science communication skills I've gained from writing papers with you. I (almost) forgive you for the seemingly endless process. You'll finally be correct when you call me "doctor" when we cross each other in the hallway.

Bas, I am very grateful that you agreed to be my promotor, which allowed me to perform my PhD research in Worcester, while maintaining a connection with the Netherlands and defending my thesis at Erasmus university. I really appreciate that you were not simply a name on the university forms but were really involved throughout my PhD. Our quarterly skype and zoom meetings helped me stay on track and you've given me many helpful suggestions on data and paper manuscripts. Plus, these meetings forced me to talk 'scientific dutch' at least a couple of times a year. I would also like to thank **all members of the Van Steensel lab** for their comments, suggestions and discussions during my visits to the NKI whenever I was in the Netherlands. Furthermore, I would like to thank **Marike van Geest**. Marike has been incredibly kind in helping me navigate the steps towards my defense. I have emailed you confused and slightly panicky regarding scheduling a defense date, filling out information on the hora finita system and many other things. You've made this process a lot less stressful.

I would like to thank all our collaborators throughout my PhD work. Outside of UMass

we collaborated with the lab of dr. **Pablo Navarro-Gil** in Institute Pasteur in Paris and dr. **Anders Hansen**, who recently started his own lab at MIT. Additionally, I have received much help from scientists within UMass Medical School. The collaborative environment at UMass has been invaluable to me. In particular, I would like to thank dr. **Sarah Hainer** (now at the University of Pittsburgh) and prof. dr. **Nick Rhind** for their help and advice during my PhD. Most importantly, I would like to thank dr. **Jon Watts** and **Adam Hedger**, who have been great collaborators throughout the many attempts and versions of SisterC. I am very grateful that both of you continued to bring your expertise and positivity to our project, even when results were sometimes discouraging.

During my years in the Dekker lab, I have also seen a lot of people join (and leave) the lab. Despite changes in lab members, the atmosphere always stayed the same. An environment where we can ask each other critical questions and have scientific (dis-) agreements, but also have long lunch breaks together and tell each other not to stay in lab too late. In no particular order; dr. **Johan Gibcus**, **Ye Zhan**, dr. **Bryan Lajoie**, dr. **Jon-Matthew Belton**, dr. **Emily Smith**, dr. **Rachel Patton McCord**, dr. **Noam Kaplan**, dr. **Filipe Tavares-Cadete**, dr. **Hakan Ozadam**, dr. **Mihir Metkar**, **Ankita Nand**, dr. **Houda Belaghzal**, dr. **Nicki Fox**, dr. **Sergey Venev**, **Liyan Yang**, dr. **Erica Hildebrand**, dr. **Yu (Sunny) Liu**, dr. **Anne-Laure Valton**, dr. **Allana Schooley**, dr. **Kristin Abramo**, **Betul Akgul Oksuz**, **Bastiaan Dekker**, **Denis Lafontaine**, **Snehal Sambare**, dr. **Davood Norouzi**, dr. **George Spracklin**, dr. **Nicola Minchell**, **Ozgun Uyan** and **Jiangyuan Liu**. Thank you all for supporting me in so many different ways during my PhD. From patiently teaching me how to analyze my own data (and how to google my questions first), to going on conferences together or synchronizing way too many plates of HFF cells. Thank you for all the practice talks, proof reading and helpful suggestions during lab meetings. And most of all, thank you for laughing at my stupid jokes and listening to my stories about hiking and farm animals.

Over the course of my PhD, our department Program in Systems Biology has grown from just 3 labs to a lively department of 7 labs. Despite the rapid growth of PSB, we maintained the friendly atmosphere, with many interactions between labs and always someone to talk to about science and non-science topics. There are too many people to name, but I do want to highlight a couple of people; dr. **Ryan Richards** and dr. **Peter Cruz-Gordillo** from the Lee lab and **Sunil Guharajan** from the Brewster lab and dr. **Cedric Diot** and **Brent Horowitz** from the Walhout lab. I also want to thank the amazing PSB support team. In particular **Missy Gardella** and **Kristen Yanick** who are always available to assist anyone in PSB. Whether I needed Missy

to help me navigate Job's busy schedule or Kristen to order a last-minute item I forgot for an important experiment, they were always happy to help. I also want to thank **Edna Froio**. Although she is no longer part of the housekeeping team on our floor, her cheerfulness spread through PSB for many years. Not only did you take care of our workspace, but you also took care of us. I will bring the tradition of telling people "Happy Friday!" to my new workplace.

I never expected this when I started my PhD, but I grew to love Worcester; from the petting zoo at Green Hill Park and the PowWow mural festivals in summer to the many brunches at Miss Worcester and the snowstorms in winter. I would like to thank the **city of Worcester** and in particular **Mayor Joseph Petty**. The city changed for the better over the past years, while staying committed to its community and its history. Worcester, the heart of Massachusetts, conquered a spot in my heart. And although I'm looking forward to moving back to Europe, I will miss being a Worcesterite.

My PhD would have been a lot less pleasant without the many hiking trips in all seasons all across New England. Through UMass hiking group **Plantation2Peaks**, I've found many of my friends. While I started as a novice hiker - coming from a country without any hills - I managed to become a P2P hiking leader and being able to give other new hikers the same introduction to the White Mountains of New Hampshire. Although I've hiked with many people from the UMass community, I would like to thank in particular dr. **Nate Erskine**, **Brent Horowitz,** dr. **Melvys Valledor**, dr. **Claudine Mapa** and **Noah Bennet**, dr. **Jessica Feldman**, dr. **Maeve Tischbein**, dr. **Bernadette Nera** and dr. **Ralph Epstein**. Thank you for going on adventures with me and getting me out of lab and into nature. I hope you can all come visit me in Germany and go on hikes in the Alps! In addition to hiking, I also found other ways to be active and in the mountains by skiing and rock climbing. Thanks to dr. **Elisa Donnard**, **Yuming Cao** and the entire climbing and skiing community that I met through **Central Rock Gym Worcester** for pushing me to ski black diamonds and belaying me on out- and indoor climbs.

I've gained many friends during my time in Worcester, but I want to mention a few people in particular. Dr. **Tyler Bormann,** dr. **Filipe Tavares-Cadete**, dr. **Hannah Simon-Girard**, dr. **Allana Schooley** and **Nisha Kini**. You are my Worcester family. Allana, you know my 'dutch love' doesn't make it easy to express feelings, but I can't imagine my PhD years without you. Thank you for being my twisted sister.

I also want to thank my friends in the Netherlands; **Jitske Loonen-Marcelis**, **Anouk Baars**, **Gerda Kamsma** and my cousin **Puck Oomen**. Despite time difference and the big distance, I'm so happy we managed to stay in touch. Hopefully we will be able to see each other more often when I live in Munich. I also want to thank dr. **Annemarie Perez-Boerema** and (soon-to-be dr.) **Kitty Hendriks**. I am so happy we became close friends during our bachelor's program. Although we moved to 3 different countries and we only saw each other a handful of times over the past years, I love that we regularly have long phone calls about our PhD experiences and life in general. I hope 2022 will be the year we will all live in the same country again!

This thesis is dedicated to my grandmothers **Trijntje Maria Vogel-Schaap** and **Marijke Oomen-Simons**. I am very aware that the opportunities I have had over the past several years were not given to my grandmothers when they were my age. Although they have passed away, I thank them for the values they instilled in me; take care of others, be creative, never stop learning, and cook and eat good food with the people you love.

I want to thank all other members of my family, in particular my parents, **Toon Oomen** and **Margreet Vogel**, and their respective partners, **Sjoukje van Weert** and **René Postulart**. Although they might not have always liked the fact that I chose to do my PhD in the US, they supported me throughout. I am very thankful for the foundation they gave me and my brothers during our childhood, which gave me the confidence I needed to pursue my scientific career in a foreign country. I also thank my brothers, dr. **Pieter Oomen** and **Joost Oomen**, and their partners, **Sophie Bous** and **Helena Hoogenkamp**. I am so happy that you agreed to be my paranymphs. Although the three of us have quite different characters, I know I have both of you in my corner to help me defend my thesis or whatever else life might throw my way. Special thanks to Pieter for helping me write the translation of my thesis summary, as my Dutch writing skills are apparently not much better than Google Translate.

Lastly and most importantly, I want to thank dr. **Eugenio Mattei**. Who would have expected that this Italian boy could melt the heart of the icy Dutch girl? Thank you for going on big and small adventures with me. Thank you for letting me be a strong and independent scientist when I want to and for taking care of me when I need. And thank you for all the meals that were ready for me right as I walked through the door after a long day in lab. I am both nervous and excited to see what life will bring next, but I know I want to share it with you.